TUM School of Computation, Information and Technology
Technische Universität München

Learnable personalization of brain tumor growth models

Ivan Ezhov

Vollständiger Abdruck der von der TUM School of Computation, Information and Technology der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitz:
Prof. Dr. Julia Schnabel

Prüfer*innen der Dissertation:
1. Prof. Dr. Bjoern H. Menze
2. Prof. Dr. John Lowengrub
3. Prof. Dr. Andreas Mang

Die Dissertation wurde am 17.04.2023 bei der Technischen Universität München eingereicht und durch die TUM School of Computation, Information and Technology am 08.11.2023 angenommen.

Abstract

Optimal treatment planning for patients diagnosed with brain tumors requires an understanding of the disease dynamics. Mathematical formalization of the tumor progression typically implies resorting to a system of differential equations. Such a system of equations describes the spatial and temporal evolution of the tumor profile through parameterized domain-specific knowledge. To calibrate the mathematical model with respect to the empirical data, one has to utilize parametric inference schemes. Currently existing optimization schemes can provide both point and probabilistic estimates of the model parameters. However, their prohibitively long compute times hinder large-scale validation and, thus, a translation of such formulations to clinical practice.

This dissertation addresses the computational bottleneck of brain tumor model personalization by developing efficient algorithms for parametric inference. First, we provide an introduction to the brain tumor modeling field, discussing tumor biology, its diagnostics, and treatment, Chapter 2. Then, we discuss the fundamentals of and state-of-the-art techniques for forward and inverse tumor growth models, Chapters 3 and 4. Chapter 5 lays out our first study within this thesis, where we introduce a fast learnable tumor growth solver as an efficient substitution to traditional numerical forward model simulators. In Chapter 6, we show how deep learning can also be used to significantly reduce the computational cost for the inverse models by reducing the number of simulations required to achieve convergence in parametric estimation. Finally, Chapter 7 demonstrates an alternative optimization scheme to obtain patient-specific tumor model parameters within clinically acceptable computational time. Concluding remarks complete the dissertation.

Zusammenfassung

Eine optimale Behandlungsplanung für Patienten mit Hirntumor erfordert ein Verständnis der Krankheitsdynamik. Die mathematische Formalisierung der Tumorprogression greift in der Regel auf ein System von Differentialgleichungen zurück. Ein solches Gleichungssystem beschreibt die räumliche und zeitliche Entwicklung des Tumorprofils durch parametrisiertes domänenspezifisches Wissen. Zur Kalibrierung des mathematischen Modells mittels der empirischen Daten werden parametrische Inferenzverfahren verwendet. Derzeitige Optimierungsverfahren sind in der Lage, sowohl Punktschätzungen, als auch probabilistische Schätzungen der Modellparameter zu liefern. Ihre langen Rechenzeiten (Tage oder sogar Wochen) verhindern jedoch eine Validierung auf größeren Patientenkohorten und somit eine Übertragung dieser Modelle in die klinische Praxis. Um den limiterenden Faktor der Rechenzeit zu eliminieren, befasst sich diese Dissertation mit der Entwicklung effizienter Algorithmen für die parametrische Inferenz zur Personalisierung von Hirntumormodellen. Die Dissertation beginnt mit einer Einführung in das Gebiet der Hirntumormodellierung. In Kapitel 2 werden Tumorbiologie, Tumordiagnostik und Tumorbehandlung erörtert. Kapitel 3 behandelt die Grundlagen und zeitgemäße Techniken zur vorwärtsgerichteten (*forward*) Tumorwachstumsmodellierung. Inverse Tumorwachstumsmodelle werden in Kapitel 4 erläutert. In Kapitel 5 wird ein schnell erlernbarer Tumorwachstumslöser als effizienter Ersatz für traditionelle numerische Vorwärtsmodellsimulatoren vorgestelt. Kapitel 6 illustriert wie Deep Learning zur erheblichen Reduktion der Rechenzeit eingesetzt werden kann. Hierzu wird die Anzahl der zur Konvergenz der parametrischen Schätzung benötigten Simulationen verringert. Darauffolgend wird in Kapitel 7 ein alternatives Optimierungsverfahren zur raschen Ermittlung patientenspezifischer Tumormodellparameter vorgestellt. Abschlußbemerkungen komplettieren die Dissertation.

Acknowledgements

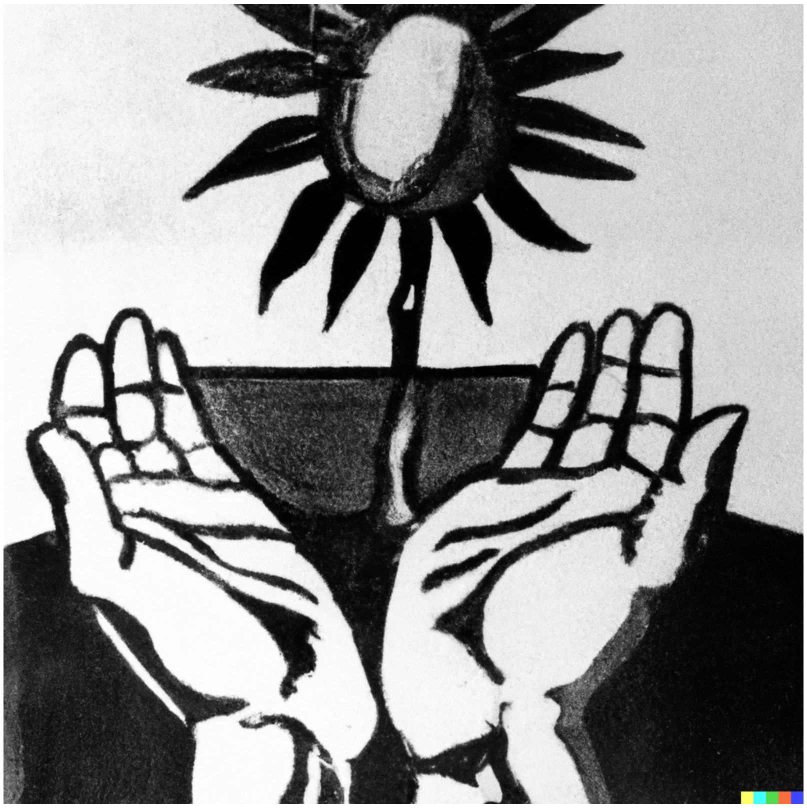

Figure 1: A sketch of a sunflower seed growing in hands. The figure is generated by the DALLE [1].

Back then, when I was sitting in a squat flipping a sunflower seed, if somebody told me that in 15 years I would be writing a Ph.D. thesis in medical artificial intelligence in a highly reputable university, I would... But here we go.

First and foremost, I would like to express gratitude to my supervisor Björn Menze, for his guidance throughout my Ph.D. I thank Bjoern for providing me with absolute research freedom - the only setting I can deliver under. His encouragement and appreciation of my color codes in conference posters and paper figures were crucial to my progress over the last years. Without his patience and time spent on me, I would still be stuck in the library trying to figure out where the exit is.

I also want to thank Benedikt Wiestler, my supervisor from the medical side, for his insights and support. His expertise in neuroradiology was instrumental in helping me focus my research. I especially thank him for keeping high the enthusiasm for translating our tumor modeling ideas into clinical practice despite my constant skepticism about it.

I am thankful to my colleagues for readily offering moral assistance over the years. Working with such talented and warm-hearted people has been necessary to get the Ph.D. accomplished. Finally, I thank my friends, who have been with me every step of the way.

Thank you all once again for your support and guidance in growing a sunflower seed in my hands into a doctoral dissertation.

Contents

Abstract . . . . **i**

Zusammenfassung . . . . **iii**

Acknowledgements . . . . **v**

Contents . . . . **vii**

List of Figures . . . . **ix**

Publication List . . . . **xi**

1 Introduction **1**

2 Biomedical background **5**
- 2.1 Glioma pathophysiology 5
- 2.2 Glioma imaging 7
- 2.3 Glioma treatment 8

3 Tumor modeling **11**
- 3.1 ODE-based tumor models 11
- 3.2 PDE-based tumor models 12
  - 3.2.1 Reaction-diffusion equation 13
  - 3.2.2 Reaction-advection-diffusion equation 14

4 Inverse problem in tumor modeling **17**
- 4.1 PDE-constrained optimisation 17
- 4.2 Monte Carlo sampling methods 18

5 Chapter 5: Geometry-Aware Neural Solver for Fast Bayesian Calibration of Brain Tumor Models **21**

6 Chapter 6: Neural parameters estimation for brain tumor growth modeling 33

7 Chapter 7: Learn-Morph-Infer: A new way of solving the inverse problem for brain tumor modeling 45

8 Concluding Remarks 73

Appendices 77

Bibliography 77

List of Figures

1 A sketch of a sunflower seed growing in hands. The figure is generated by the DALLE [1]. v

1.1 An illustration of a brain tumor. A typical pathology feature is a higher vascularization compared to healthy brain volume. The figure is generated by the stable diffusion [6]. 2

2.1 An illustration of cancerous cells infiltrating the membrane of healthy tissue. Infiltrating glioma growth is a stark manifestation of pathology malignancy. The figure is generated by the DALLE [1]. 6

2.2 Examples of MRI (T1, FLAIR) and PET images of brain tumor. P1-P5 labels indicate different patients. Copyright © 2022, IEEE. 7

3.1 A sketch of brain tumor evolving in time. The figure is generated by the DALLE [1]. 11

4.1 A sketch of the distributions (target $p(x)$ and sampling ones $q(x)$) which are involved in the importance sampling algorithm. During compute of the expectation, each sample is weighted by a value w_i, $w_i = p(x_i)/q(x_i)$. 19

Publication List

This dissertation is based on the following peer-reviewed publications:

[1] **I. Ezhov**, J. Lipkova, S. Shit, F. Kofler, N. Collomb, B. Lemasson, E. Barbier, and B. Menze. “Neural parameters estimation for brain tumor growth modeling.” In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2019, pp. 787–795.

[2] **I. Ezhov**, T. Mot, S. Shit, J. Lipkova, J. C. Paetzold, F. Kofler, C. Pellegrini, M. Kollovieh, F. Navarro, H. Li, M. Metz, B. Wiestler, and B. Menze. “Geometry-aware neural solver for fast Bayesian calibration of brain tumor models.” In: *IEEE Transactions on Medical Imaging* 41.5 (2021), pp. 1269–1278.

[3] **I. Ezhov**, K. Scibilia, K. Franitza, F. Steinbauer, S. Shit, L. Zimmer, J. Lipkova, F. Kofler, J. Paetzold, L. Canalini, D. Waldmannstetter, M. J. Menten, M. Metz, B. Wiestler, and B. Menze. “Learn-Morph-Infer: a new way of solving the inverse problem for brain tumor modeling.” In: *Medical Image Analysis* (2022), p. 102672.

Introduction

Glioma is an aggressive primary brain tumor associated with a poor prognosis [2] and a median survival time of 15 months [3]. The pathology is characterized by a highly infiltrative spread of tumor cells into surrounding healthy tissues. Such spread complicates the surgery procedure and thus lowers its effectiveness. This, in turn, necessitates follow-up treatments such as chemo- and radiotherapy.

For radiotherapy, every new patient undergoing the treatment requires customized planning. The planning includes identifying a three-dimensional volume inside a patient's brain to be irradiated with X-rays. To identify an abnormal tumor area, expert neuroradiologists investigate medical scans obtained by imaging diagnostic modalities, such as magnetic resonance imaging (MRI). The area which is visibly distinguishable in terms of image contrast from healthy brain tissue could be defined as the area for tumor irradiation. Years of clinical practice, however, revealed that such irradiation volume definition leads to a high rate of tumor recurrence after treatment. Despite the advances in the MRI technique providing high spatial resolution, the MRI modality does not yet portray complete information about the tumor spread. Below certain tumor cell concentrations, tumor spread is not sufficiently distinguishable on MRI scans. Thus, irradiating only the visible tumor volume can lead to post-treatment tumor recurrence since the low-concentration tumor area is left untreated.

To address this imaging limitation, current radiotherapy guidelines suggest irradiating a volume extended by a few centimeters margin around the MRI-derived tumor boundaries [4], [5]. Despite the improved prognosis brought by such planning, it is still far from optimal. First, such a definition implies targeting potentially healthy brain tissue confined within the margin. Furthermore, it can still miss targeting tumor cells outside of this margin in the case of highly anisotropic tumor distribution. Finally, it is unclear what the exact distance of this margin should be [1]. Thus, there is a clear need for personalized radiotherapy planning.

Due to a lack of empirical support for identifying the full spatial distribution of tumor cells, a potential direction for radiotherapy personalization is based on

---

[1] This explains notably varying margin definition over official radiotherapy guidelines [4], [5]

Figure 1.1: An illustration of a brain tumor. A typical pathology feature is a higher vascularization compared to healthy brain volume. The figure is generated by the stable diffusion [6].

computational modeling. The overarching goal of this approach is to utilize a mathematical model to simulate the progression of glioma. The simulated tumor profile can then be used for defining the irradiation volume.

In more technical terms, the main objective of the modeling-based approach is to solve the inverse problem - find tumor model parameters resulting in a simulation best describing the empirical tumor information. In this thesis, such information is obtained from medical imaging modalities. Various methods exist to calibrate a brain tumor forward model with respect to medical scans [7]. The methods based on PDE-constrained optimization typically provide point estimates of the tumor model parameters. Another class of Bayesian methods provides probabilistic estimation for the best matching parameters. The unifying drawback of the methods is, however, the fact that they require a large number of simulations to reach convergence of the parametric estimation. In the most computationally demanding cases of probabilistic approaches, the convergence requires up to tens of thousands of simulations. At the same time, a single simulation, even for simplistic mathematical models, can take a few minutes of computation. These constraints result in prohibitively long computing times for solving the inverse problem (up to multiple days on above-average hardware), questioning the utility of the modeling-based approach for clinical needs. The objective of this doctoral dissertation is to develop algorithmic solutions that would allow efficient solving of the inverse problem, opening the doors for its clinical translation.

Contribution

This thesis contributes to the biomedical modeling field by providing several techniques to infer biophysical model parameters from medical imaging data. Paper-by-paper, we addressed the computational bottlenecks for both forward and inverse problems, steadily reducing the time for tumor model personalization. In more detail, we made the following contributions:

1. Development of a learnable forward tumor growth model that serves as a surrogate for numerical tumor model solvers. The surrogate represents an end-to-end deep learning pipeline that takes input model parameters and simulation geometry and outputs the corresponding simulation. The learnable surrogate achieves a 50x speedup compared to a highly efficient numerical forward solver at the error cost acceptable for the downstream personalization task.

2. Development of a learnable inverse scheme to substitute traditional Markov Chain Monte Carlo sampling for model personalization. The proposed inverse scheme bypasses the Monte Carlo method's necessity to define an analytical form for the likelihood function. The likelihood is instead learned from simulated data. This procedure allows, in turn, to reduce by one order the number of parameters required for convergence in parametric estimation.

3. Development of a novel scheme for tumor model inversion. The method eliminates the need to simulate thousands of samples for every new patient-specific geometry. As an alternative to sampling methods, the presented technique implies learning inverse mapping – from simulation to model parameters – in a reference brain atlas. The network is trained only once in the atlas space on

a large set of simulations. For every new patient, we obtain patient-specific model parameters by registering patient information to the atlas space and performing inference with the trained beforehand network. Such a scheme allows performing model personalization within four to seven minutes instead of days or even weeks using traditional parametric inference approaches.

Biomedical background

This chapter is dedicated to the fundamentals of tumor pathophysiology, its clinical diagnostics, and treatment. After discussing biological features differentiating cancerous and normal physiology, we describe imaging diagnostics modalities allowing to monitor the progression of the disease. The chapter concludes with an introduction to existing treatment strategies to eradicate tumor progression.

2.1 Glioma pathophysiology

A distinguishing hallmark of all cancer types is abnormal cell proliferation. According to medical vocabulary, an agglomerate of excessively dividing cells is referred to as a neoplasm. A neoplasm occurring inside the brain – glioma – is the focus of this dissertation.

Glioma can be classified into benign and malignant types. The latter is characterized by high infiltration into surrounding healthy tissues. In turn, infiltrative growth results from the loss of intercellular adhesion molecules, leading to detachment from the gross tumor volume. Due to built-in genetic programs, malignant glioma cells gain, in addition to high proliferation, high motility facilitating penetration through the volume of the neighboring tissue [8].

Rapidly growing tumor demands an increased blood supply, delivering nutrients and oxygen to the cell population. As a result, excessive blood vessel growth is induced to supply the normal metabolic needs, known as angiogenesis. During angiogenesis, new vessels are formed from existing ones by processes of sprouting and splitting [9]. The sprouting manifests as an active proliferation of blood vessel (endothelial) cells to form sprouts that link neighboring existing vessels. In contrast, splitting angiogenesis implies the formation of new vessels directly from the existing ones through the reorganization of cells.

Despite breaking control mechanisms through angiogenesis, after a certain increase in the cancerous cell population, the increase in blood supply reaches its limits. Consequently, the cells in the core of the tumor get deprived of nutrient supply, resulting

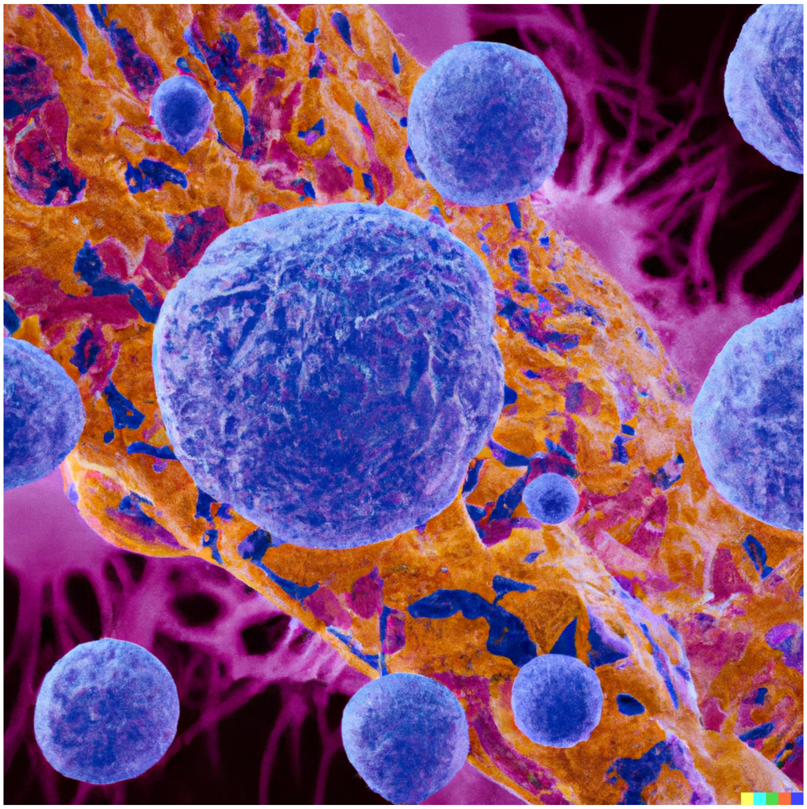

Figure 2.1: An illustration of cancerous cells infiltrating the membrane of healthy tissue. Infiltrating glioma growth is a stark manifestation of pathology malignancy. The figure is generated by the DALLE [1].

in the formation of a so-called necrotic core. Therefore, gliomas typically appear as ring-enhancing structures – necrotic core in the middle and actively proliferating enhancing part along the circumference of the core.

Another typical brain tumor characteristic is the formation of edema. Edema represents a fluid build-up around the tumor ring. The cause for its appearance can be understood as a body response to the presence of a foreign entity (glioma) inside the brain. Edema is commonly present for both benign and malignant gliomas and occupies a notably large area around the tumor. Thus, edema alone can cause increased intracranial pressure and impaired function of nerves and vessels [10].

2.2 Glioma imaging

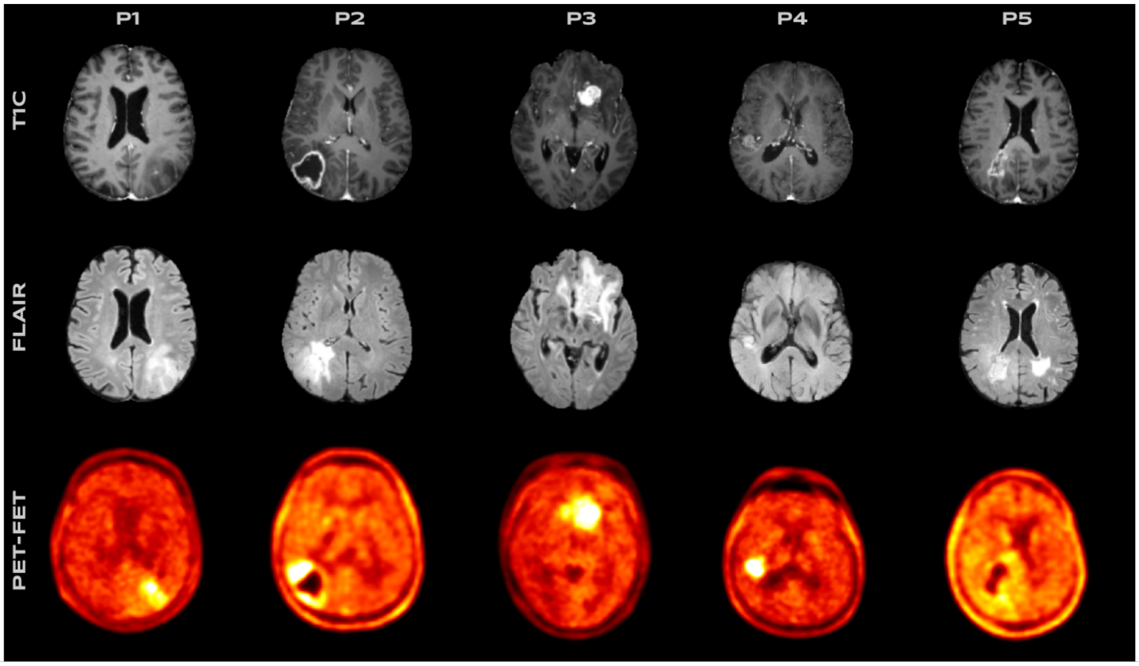

Figure 2.2: Examples of MRI (T1, FLAIR) and PET images of brain tumor. P1-P5 labels indicate different patients. Copyright © 2022, IEEE.

Magnetic resonance imaging

Advances in the 20th century in understanding the fundamentals of atomic physics led to several practical applications. For example, the seminal works of Rabi [11], Purcell [12], and Bloch [13] on nuclear magnetic resonance resulted in the development of MRI technology. Today, MRI is a widely used imaging diagnostic system, especially for medical needs. It can portray both anatomical as well as functional information about the body. In contrast to other modalities, such as computed tomography, MRI provides richer visualization capabilities for soft tissues, like brain tissues. By controlling the sequence between the external magnetization of the atomic nucleus constituting the brain matter and the inner relaxation of induced magnetic moments, a particular visualization focus can be brought to the tissue of interest. The primary MRI sequences are referred to as T1 and T2 [14]. T1 MRI sequence highlights fatty tissue and is thus valuable for contrasting white and gray matter as they vary in fat content. T2 sequence enhances areas rich in water content and is, therefore, preferable for detecting edema. Another frequently used MRI sequence, fluid-attenuated inversion recovery (FLAIR), allows suppression of the fluid signal. FLAIR can be used in brain imaging to suppress, e.g., the cerebrospinal fluid signal.

Positron-emission tomography

Positron-emission tomography (PET) is another example of atomic physics translation to medical routine practice [15]. PET is an imaging modality that utilizes radioactive tracers to quantify local changes in metabolism. Radiotracers such as FET (Fluoroethyl-L-tyrosine) injected in the body can be taken up and trapped by cells, especially malignant ones demanding abnormal nutrient supply. Gamma rays emitted by the tracers are detected by cameras, allowing spatial visualization of metabolic activities. The PET-derived metabolic activity provides more quantitative information about tumor cell concentration than MRI. However, PET imaging provides worse resolution - typically 4mm vs. 1mm (MRI) per voxel.

Computed tomography

Computed tomography (CT) is a medical imaging modality that uses X-rays to create detailed cross-sectional images of the body [16]. During a CT scanning, a patient lies on a table that moves through a circular machine that produces X-rays. The X-rays are directed at the head from different angles, and reconstruction algorithms combine the resulting 2D X-ray absorption maps to create detailed 3D brain images. CT scans can help identify morphological properties such as the tumor's size, shape, and location.

2.3 Glioma treatment

Treatment options for glioma include surgery, radiation- and chemotherapy. Due to the disease's severity and fast progression, a combination of these treatments is often applied. In some cases, immunotherapy can be recommended - a more recent treatment that stimulates a patient organism's immune system to attack tumor cells [17].

Surgery is typically the first treatment for glioma. During the surgical operation, cancerous tissue is mechanically resected while trying to preserve as much healthy brain tissue as possible. In the case of the tumor's proximity to functionally important brain regions, complete resection of the tumor might not be an option. Then partial removal can be carried out to reduce cranial pressure and the tumor area to be treated by follow-up radio- and chemotherapies.

Radiation therapy utilizes high-energy beams of X-rays to kill cancer cells and shrink the tumor size. Typically it is applied externally, using a machine delivering the radiation to the glioma area. In some cases, though, it can be applied internally using so-called brachytherapy. During brachytherapy, radioactive material is placed into or near the lesion. Several studies showed that such treatment could reduce radiotherapy side effects by reducing damage to neighboring healthy tissue [18], [19].

Finally, chemotherapy is a treatment that uses specific medicine mainly to inhibit tumor cell proliferation. The medicine is typically injected intravenously but can also be given orally. The therapy can be applied prior to surgery to reduce tumor size as well as after surgery to target any remaining tumor cells [20].

Tumor modeling

Brain tumors are a severe health concern, and understanding the causes of their progression is essential for improving treatment strategies. In this section, a reader will be introduced to mathematical means to describe the development of tumors. The main focus will be laid on the usage of differential equations for the pathology description. We will start with the simplest type of differential equations - the ordinary differential equations (ODEs). Then the partial differential equations (PDEs) will be discussed on a concrete example of reaction-diffusion PDE formalism.

Figure 3.1: A sketch of brain tumor evolving in time. The figure is generated by the DALLE [1].

3.1 ODE-based tumor models

ODEs are a type of differential equations that involve an independent variable (usually time), an unknown function of the independent variable, and one (or more) derivatives of the unknown function with respect to this variable. ODEs describe the behavior of dynamic systems in many natural science and engineering fields that change over time.

These differential equations can be classified into first-order and higher-order equations depending on the order of the derivative of the unknown function. Depending on the equation's complexity, ODEs are solved analytically or numerically.

ODEs can also serve to model the growth of brain tumors. The simplest ODE used to model tumor growth is based on the exponential model [21]. It belongs to the first-order ODE type and assumes that cells proliferate at a linear rate, leading to exponential growth:

$$\frac{dy}{dt} = \alpha \cdot y \tag{3.1}$$

where y models the size of the tumor population, and α is a growth rate.

Another ODE applied to describe brain tumor growth is based on the logistic growth model. This model evolves the tumor cell population in time with a rate of change proportional to the current population size. The growth of the population is, however, limited by the brain's capacity - the capacity to support the tumor's growth, e.g., with nutrients [21], [22]. The logistic growth model is often written as follows:

$$\frac{dy}{dt} = \rho \cdot y \cdot (1 - \frac{y}{K}) \tag{3.2}$$

where ρ is the intrinsic growth rate of the tumor, and K is the carrying capacity of the brain.

Yet another Gompertz model is widely used to describe tumor evolution [23], [24]:

$$\frac{dy}{dt} = \alpha \cdot exp(-\beta t) \cdot y \tag{3.3}$$

here, the coefficient α is the initial cell division rate, and β can be thought of as the rate of exponential decay of α.

Other ODE models have been developed to describe the tumor treatment, including models that consider the effects of chemotherapy [25] and radiation therapy [26]

3.2 PDE-based tumor models

In contrast to ODEs, PDEs are differential equations that involve partial derivatives of an unknown function with respect to two (or more) independent variables. Thus, PDEs describe physical systems that involve multiple variables changing simultaneously. Analogously to ODEs, PDEs can be classified based on the order of their derivatives.

PDEs are widely utilized to model the growth of brain tumors and their treatment [27], as well as other phenomena that occur in the brain, such as the blood flow [28] or the electrical activity of neurons [29].

3.2.1 Reaction-diffusion equation

The reaction-diffusion equation is a typical PDE model that describes a wide range of biological and chemical phenomena, including brain tumor growth [30]–[32]. In the case of glioma modeling, it can describe how the tumor cells' concentration changes in time due to cell diffusion and proliferation, i.e., the production of new tumor cells through the division process. The equation can be written as:

$$\frac{\partial c}{\partial t} = D\frac{\partial^2 c}{\partial x^2} + f(\rho, c) \tag{3.4}$$

where c is the concentration of tumor cells, D is the diffusion coefficient, and $f(\rho, c)$ is a function that describes the reaction process with ρ denoting the reaction parameter (e.g., cell proliferation rate).

The diffusion coefficient D defines how easily glioma cells migrate through the medium of brain tissues. Cell migration can depend on various factors, such as the cell's size and shape or the brain tissue's viscosity. The diffusion coefficient D is crucial in understanding and modeling how a tumor grows and spreads within the brain. For example, a tumor with a high D likely invades nearby tissues, while a tumor with a low D forms a more contained tumor mass. For gliomas, the former is assumed to be more common.

Similarly to the diffusion term, the reaction term $f(\rho, c)$ in the reaction-diffusion equation determines the overall dynamics of the system, influencing the trajectory of the cell concentration c evolution. The form of the reaction term can vary depending on the specific chemical reactions being modeled. A common approach to model brain tumor growth is by using a reaction term representing the proliferation of tumor cells:

$$f(\rho, c) = \rho c(1 - c)$$

The term $f(\rho, c)$ can also include the angiogenesis effect [33], or the death of tumor cells through radiotherapy or chemotherapy [27], [34]–[36].

3.2.2 Reaction-advection-diffusion equation

The reaction-advection-diffusion equation describes how a spatial distribution of a substance's concentration changes over time in a system that involves chemical reactions, advection (the movement of the substance by a fluid or other carrying medium), and diffusion. The equation can model the transport of substances in a medium, for example, the transport of gases or particles [37].

The equation is typically written as a PDE of the form:

$$\frac{\partial c}{\partial t} = \nabla(D\nabla c) + \nabla(c\mathbf{v}) + f(c, \mathbf{v}) \tag{3.5}$$

where $\mathbf{v}$ is the fluid's or carrier's velocity, D is the diffusion coefficient, and ∇ is a nabla operator. The function $f(c, \mathbf{v})$ represents the chemical reactions that are taking place in the system.

In brain tumor growth modeling, the reaction-advection-diffusion equation can model the movement of various substances within the brain tissue, such as oxygen or nutrients. The equation can also describe the mass effect of a growing tumor. The mass effect here refers to the physical pressure a brain tumor exerts on the surrounding tissue as the lesion grows. A system of partial differential equations to describe this process could be written as:

$$\frac{\partial c}{\partial t} = \nabla(D\nabla c) - \nabla(c\mathbf{v}) + \rho c(1-c) \tag{3.6}$$

$$\frac{\partial \mathbf{m}}{\partial t} + \nabla \cdot (\mathbf{m} \otimes \mathbf{v}) = 0 \tag{3.7}$$

$$\nabla \cdot (\lambda \nabla \mathbf{u} + \mu(\nabla \mathbf{u} + \nabla \mathbf{u}^\top)) = \gamma \nabla c \tag{3.8}$$

$$\frac{\partial \mathbf{u}}{\partial t} = \mathbf{v} \tag{3.9}$$

$$\nabla c \cdot \mathbf{n} = 0 \quad \textit{boundary condition} \tag{3.10}$$

$$\mathbf{m} = 0 \quad \textit{boundary condition} \tag{3.11}$$

$$\mathbf{u} = 0 \quad \textit{boundary condition} \tag{3.12}$$

$$\mathbf{v} = 0 \quad \textit{boundary condition} \tag{3.13}$$

here $\mathbf{m} = (m_{WM}(\mathbf{i}, t),\ m_{GM}(\mathbf{i}, t),\ m_{CSF}(\mathbf{i}, t))$ denotes the brain tissue maps for voxel $\mathbf{i}$ and time t, ρ is a tumor cell proliferation rate, $\mathbf{u}$ is a displacement vector and $\mathbf{v}$ is a displacement velocity. The driver for the displacement in this model is the mass effect resulting in tissue displacement. The linear elasticity model, Eq. 3.8, describes the interaction between the tumor and normal parenchyma. The elasticity model is parameterized through the Lame coefficients (μ, λ). Finally, the mass effect depends on the choice of the mass effect parameter γ. Choosing γ=0 would result in the simplistic reaction-diffusion PDE with $\mathbf{u}$=0 and $\mathbf{v}$=0 [38].

Inverse problem in tumor modeling

The numerical modeling of complex systems discussed in the previous chapter plays a vital role in many scientific disciplines. However, viewed from a practical perspective, the foremost objective is establishing a link between a mathematical model and empirical observations. This chapter will introduce several formalisms used to connect a physical (forward) model with empirical brain tumor data. In other words, this chapter is dedicated to solutions to the inverse problem.

In what follows, we first discuss the methods of PDE-constrained optimization. We then discuss another family of methods to address the inverse problem based on sampling, such as Monte Carlo sampling.

4.1 PDE-constrained optimisation

In PDE-constrained optimization, the goal is to minimize a discrepancy between real data $\tilde{y}$ under a constraint that simulated data y satisfies a PDE of interest. First, solving PDE-constrained optimization problems involves using numerical techniques to approximate the solution of the PDE. In addition, one has to construct an optimization algorithm to find the solution to the objective function subject to these approximations [39].

The optimization problem can be formulated as follows:

$$\min_{y,u} J(y,u) \qquad\qquad \text{subject to } F(y,u) = 0, \tag{4.1}$$

where u represents a control optimisation variable, $J(y,u)$ is the objective function to be minimized, and $F(y,u) = 0$ represents the PDE constraint. This optimization problem seeks to find the values of y and u that minimize the objective function $J(y,u)$ while satisfying the constraint $F(y,u) = 0$.

The objective function is typically represented as:

$$J(y,u) = ||\tilde{y} - y||^2 + \beta||u||^2 \tag{4.2}$$

where β is a coefficient controlling the so-called Tikhonov regularization term [40], which is used to mitigate the ill-posed character of the problem.

Various techniques can be used to solve PDE-constrained optimization problems, including the penalty, primal and dual, or adjoint-state methods. The overarching idea is to cast the PDE into the minimization objective alongside the $J(y, u)$ function. For example, the penalty method solves the following problem:

$$\min_{y,u} J(y, u) = ||\tilde{y} - y||^2 + \beta||u||^2 + \lambda||F(y, u)||^2 \tag{4.3}$$

here λ is the penalty coefficient controlling that the solution y corresponds to the PDE, $F(y, u) = 0$.

Gradient-based algorithms, such as gradient descent or Newton's method, or global optimization schemes, such as genetic or simulated annealing, are typically used to solve the minimization objective [41].

In the context of glioma modeling, several works successfully explored the PDE-constrained formalism [7], [42]–[47]. Under this formalism, the variable y would represent the cell concentration while u can represent variables, e.g., related to tumor-induced tissue deformation [44]. However, the typical drawback of PDE-constrained optimization is the computational complexity, especially for complex systems with many variables and constraints. Some of these methods, such as gradient-based optimization algorithms, may be relatively fast but may not always find the globally optimal solution. Other methods, such as a family of global optimization methods, can be slower but more likely to find the optimal solution.

4.2 Monte Carlo sampling methods

Monte Carlo sampling is a technique for approximating a distribution or expected value by generating a large number of random samples and analyzing the statistical properties of the obtained sample population. This method can estimate the value of complex distributions or expectations that cannot be computed analytically.

There are many different versions of the Monte Carlo sampling algorithm, each having specific characteristics [48]. Some of the most common versions of the Monte Carlo methods are:

Importance sampling. The importance sampling algorithm generates samples from a distribution $q(x)$ (typically Gaussian or a mixture of Gaussians) different from

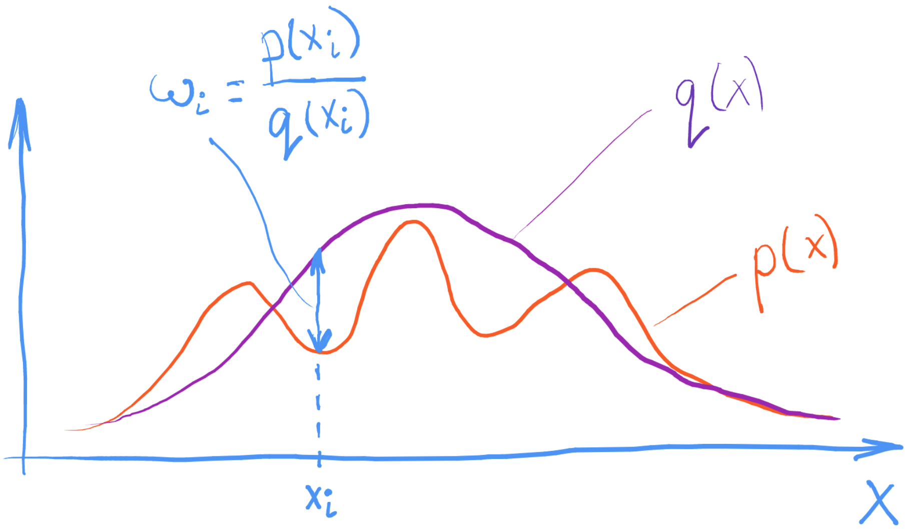

Figure 4.1: A sketch of the distributions (target $p(x)$ and sampling ones $q(x)$) which are involved in the importance sampling algorithm. During compute of the expectation, each sample is weighted by a value w_i, $w_i = p(x_i)/q(x_i)$.

the target distribution $p(x)$ but easier to sample from [49]. The samples are then weighted based on the ratio between the target and the sampling distributions, Fig. 4.1. Finally, the weighted samples are used to estimate the desired expectation value (or distribution):

$$E(f(x)) = \int f(x)p(x) = \int f(x)\frac{p(x)}{q(x)}q(x) \approx \frac{1}{n}\sum_i f(x_i)p(x_i)/q(x_i) \qquad (4.4)$$

where $E(f(x))$ is the expectation value of the function of interest $f(x)$.

Rejection sampling. Similarly to importance sampling, the rejection sampling algorithm implies generating samples from a different distribution that is easy to sample from to approximate the desired statistical quantity. But instead of weighting each sample, it rejects samples not from the target distribution.

Stratified sampling. This version of the Monte Carlo sampling algorithm involves dividing the domain of the target distribution into subregions and generating samples from each subregion separately. Stratified sampling can be more efficient than simple Monte Carlo sampling from a variance reduction point of view if the target distribution is nonuniform.

Markov chain Monte Carlo (MCMC) sampling. MCMC implies generating a sequence of samples that form a Markov chain [50], with every new sample being dependent on the previous sample in the chain [51]. In addition to estimating complex distributions, MCMC sampling is often used to perform Bayesian inference. A typical MCMC algorithm works as follows:

1. Begin with a point x_0 at a random position in the sample space.
2. Propose a new sample using a conditional distribution $g(x_1|x_0)$, where $g(|)$ is commonly a Gaussian distribution.
3. Calculate the acceptance ratio α of the new position based on the distribution we want to sample from (or a distribution that is proportional to the one) - $\alpha = p(x_{i+1})/p(x_i)$.
4. With probability equal to α move to the new position. Otherwise, stay in the current position.
5. Repeat steps 2 to 4 for a predefined number of iterations or till convergence.

By repeating these steps, we obtain a sequence of samples that can be used to approximate the desired distribution.

In relation to PDE, Monte Carlo methods sample free model parameters, which are then propagated through the forward model to obtain corresponding physical simulations. By relating the distribution of the simulations with empirical data, one can infer statistical properties, e.g., the posterior probability of the model parameters given the data.

Existing literature exploring variations of Monte Carlo sampling for the inverse brain tumor modeling [52]–[54] shows promising results in linking a tumor model with medical imaging data. However, similarly to the PDE-constrained optimization, the drawback of the sampling methods is the convergence rate. To obtain convergence of parametric estimation, such techniques require thousands or tens of thousands of generated samples. This can result in multiple hours or days of computation depending on the complexity of the forward model.

To address these limitations was the primary goal of this Ph.D. dissertation. Starting from the next chapter, we introduce our attempts to accelerate the parametric inference.

Chapter 5: Geometry-Aware Neural Solver for Fast Bayesian Calibration of Brain Tumor Models

Ivan Ezhov, Tudor Mot, Suprosanna Shit, Jana Lipkova, Johannes C. Paetzold, Florian Kofler, Chantal Pellegrini, Marcel Kollovieh, Fernando Navarro, Hongwei Li, Marie Metz, Benedikt Wiestler, and Bjoern Menze

Journal: IEEE Transactions on Medical Imaging, Volume 41, No. 5, May 2022.
To avoid violation of copyrights, the final manuscript submitted to IEEE for typesetting is reproduced below.

Synopsis: Modeling of brain tumor dynamics has the potential to advance therapeutic planning. Current modeling approaches resort to numerical solvers that simulate the tumor progression according to a given differential equation. Using highly-efficient numerical solvers, a single forward simulation takes up to a few minutes of compute. At the same time, clinical applications of tumor modeling often imply solving an inverse problem, requiring up to tens of thousands of forward model evaluations when used for a Bayesian model personalization via sampling. This results in a total inference time prohibitively expensive for clinical translation. While recent data-driven approaches become capable of emulating physics simulation, they tend to fail in generalizing over the variability of the boundary conditions imposed by the patient-specific anatomy. In this paper, we propose a learnable surrogate for simulating tumor growth which maps the biophysical model parameters directly to simulation outputs, i.e. the local tumor cell densities, whilst respecting patient geometry. We test the neural solver in a Bayesian model personalization task for a cohort of glioma patients. Bayesian inference using the proposed surrogate yields estimates analogous to those obtained by solving the forward model with a regular numerical solver. The near real-time computation cost renders the proposed method suitable for clinical settings. The code is available at https://github.com/IvanEz/tumor-surrogate.

5. Chapter 5: Geometry-Aware Neural Solver for Fast Bayesian Calibration of Brain Tumor Models

Contributions of thesis author: data processing, designing and implementing the algorithmic solution, manuscript writing.

Geometry-aware neural solver for fast Bayesian calibration of brain tumor models

Ivan Ezhov, Tudor Mot, Suprosanna Shit, Jana Lipkova, Johannes C. Paetzold, Florian Kofler, Chantal Pellegrini, Marcel Kollovieh, Fernando Navarro, Hongwei Li, Marie Metz, Benedikt Wiestler and Bjoern Menze, *Member, IEEE*

***Abstract*—Modeling of brain tumor dynamics has the potential to advance therapeutic planning. Current modeling approaches resort to numerical solvers that simulate the tumor progression according to a given differential equation. Using highly-efficient numerical solvers, a single forward simulation takes up to a few minutes of compute. At the same time, clinical applications of tumor modeling often imply solving an inverse problem, requiring up to tens of thousands of forward model evaluations when used for a Bayesian model personalization via sampling. This results in a total inference time prohibitively expensive for clinical translation. While recent data-driven approaches become capable of emulating physics simulation, they tend to fail in generalizing over the variability of the boundary conditions imposed by the patient-specific anatomy. In this paper, we propose a learnable surrogate for simulating tumor growth which maps the biophysical model parameters directly to simulation outputs, i.e. the local tumor cell densities, whilst respecting patient geometry. We test the neural solver in a Bayesian model personalization task for a cohort of glioma patients. Bayesian inference using the proposed surrogate yields estimates analogous to those obtained by solving the forward model with a regular numerical solver. The near real-time computation cost renders the proposed method suitable for clinical settings. The code is available at https://github.com/IvanEz/tumor-surrogate.**

***Index Terms*—Bayesian inference, physics-based deep learning, glioma, model personalization, tumor modeling, MRI, FET-PET**

I. INTRODUCTION

SIMULATION of brain tumor progression can provide complementary information to medical imaging for radiotherapy planning. As shown in [1]–[13], tumor modeling can be employed to define a personalized radio-treatment area using biophysical models to estimate the most likely directions of tumor cell infiltration instead of solely targeting tumor area visible on a scan. These methodologies mainly imply solving an inverse problem: finding the parameters of the biophysical tumor growth model resulting in a simulation output that best matches an empirical observation outlining the pathology.

Existing approaches for inverse tumor modeling resort to deterministic [2], [4], [14] as well as probabilistic Bayesian [15]–[17] formalisms. All these approaches rely on excessive amount of forward simulations required for inferring patient-specific parameters. The number of simulations ranges from several thousand for approximate methods [5], [15], [17] to tens of thousands in case of fully Bayesian analysis [3]. The forward brain tumor models are often based on the reaction-diffusion equation and are implemented using highly-efficient numerical solvers. In [16], authors employ the Lattice Boltzmann method which allows parallelized computing and takes ca. 80 seconds on a 60 core machine for simulating the pathology growth. In [3], the forward model is implemented by means of a multi-resolution adapted grid solver with a simulation time of 1-3 minutes using 2 cores. Despite the computational advances of the solvers, the minutes of a single forward model evaluation multiplied by thousands of forward integrations necessary for the inverse problem can result in weeks of total computing time. This constrains the testing of more elaborate tumor models (e.g., considering cell mixtures or multiple competing patho-physiological processes [18]), and translation of the personalized radiotherapy planning into clinical practice [1]–[4].

As recent years showed, speeding up heavy conventional computation becomes feasible using end-to-end learning methods. The data-driven methodology has also penetrated the field of numerical computing [19]–[29]. Learnable surrogates were proposed for various scientific computing tasks in the natural sciences by exploiting fully-connected [30]–[32], convolutional [33]–[35], and hybrid [36]–[38] neural architectures. Among them are two methods that proved capable of learning even a direct mapping from the space of parameters driving a simulator to the space of the simulator solutions in a static geometry [34], [35]. Unfortunately, these promising methods are incapable of dealing with inference in arbitrary complex geometries, such as those dictated by patient-specific anatomy. This limits their transfer to model personalization that is crucially dependant on an adaption to the patient specific simulation domain.

The contribution of the paper is the following: we introduce a learnable method emulating a numerical tumor growth

---

The first two authors contributed equally.

I. Ezhov and S. Shit are supported by the Translational Brain Imaging Training Network under the EU Marie Sklodowska-Curie programme (Grant ID: 765148). B. Menze, B. Wiestler and F. Kofler are supported through the SFB 824, subproject B12, by DFG through TUM International Graduate School of Science and Engineering, GSC 81, and by the Institute for Advanced Study, funded by the German Excellence Initiative. Fernando Navarro is supported by DFG-GRK 2274.

I. Ezhov, T. Mot, S. Shit, J. Paetzold, F. Navarro, F. Kofler, C. Pellegrini and M. Kollovieh are with the Department of Informatics, and with TranslaTUM - Central Institute for Translational Cancer Research, TUM, Munich, Germany (e-mail: ivan.ezhov@tum.de).

B. Wiestler and M. Metz are with the Neuroradiology Department of Klinikum Rechts der Isar, TUM, Munich, Germany

J. Lipkova is with Harvard Medical School, Brigham and Women's Hospital, Boston, United States

H. Li, B. Menze are with the Department of Informatics, TranslaTUM, TUM, Munich, Germany, and Department of Quantitative Biomedicine of UZH, Zurich, Switzerland

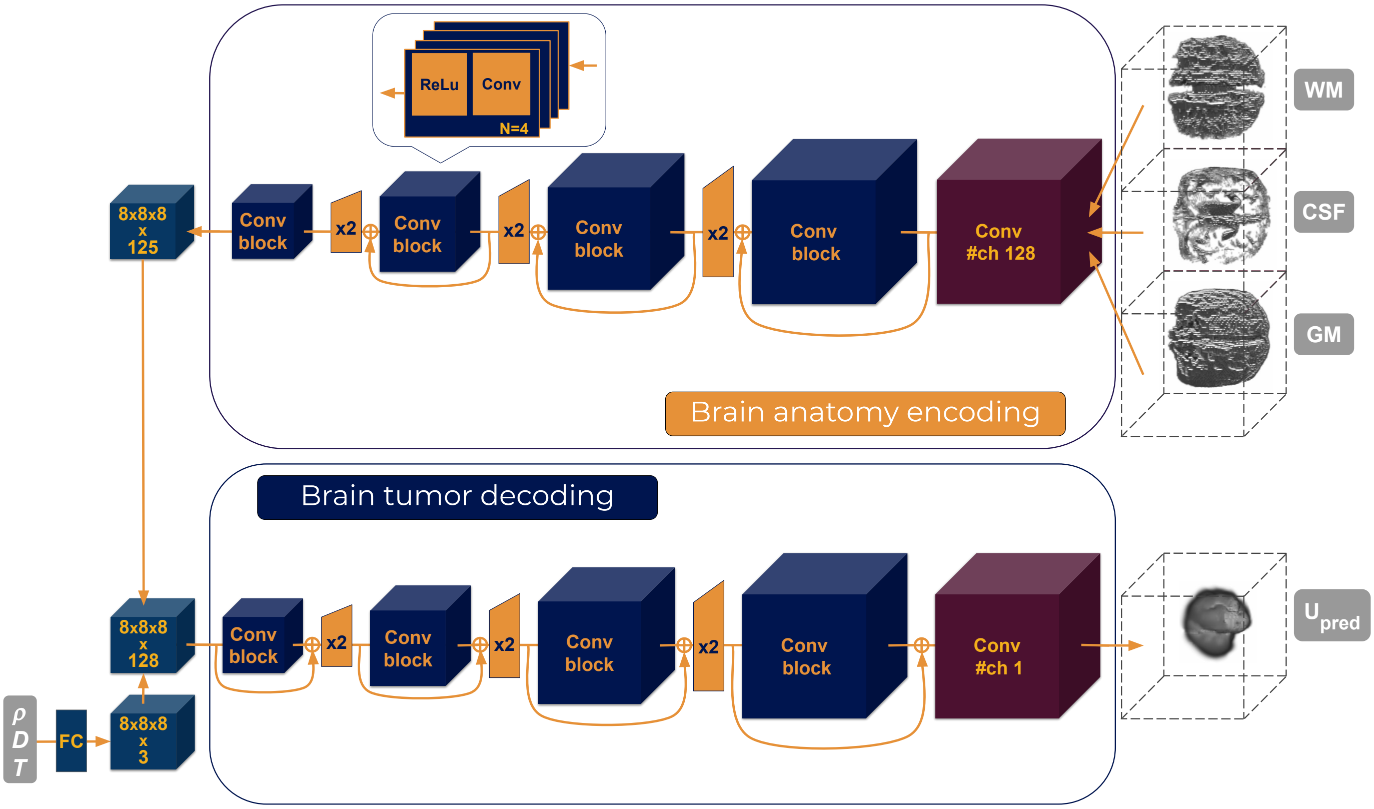

Fig. 1: Geometry-aware neural solver. The network is composed of two main parts: a) brain anatomy encoder that maps the anatomy volumes (WM, GM, CSF) to a latent representation, b) brain tumor decoder that takes as input an embedding (via a fully-connected (FC) layer) of the parameters $\{D, \rho, T\}$, concatenated with the latent representation from (a), and maps the resulting tensor to the 3D tumor simulation volume. Downsampling in the encoder is implemented as a convolution with stride 2 (other convolutions in the network have kernel size 3 and stride 1). Upsampling in the decoder is via the nearest-neighbor interpolation followed by convolutional operations. The convolutional block is composed of N repetitions of convolutional operation (with number of channels #ch 128) and ReLu non-linearity.

forward solver. Specifically, we introduce a learnable anatomy encoder that enables a patient-specific simulation of the tumor growth process. To the best of our knowledge, this is the first network-based approach in the computational pathology field that maps parameters of the biophysical model directly to the simulation outputs while generalizing over the simulation geometry. To illustrate the power of this approach, we use reaction-diffusion tumor growth model, since this is most used model and it also serves as a base for many more complex models. We achieve a 50× speed-up compared to an advanced numerical solver by employing the tumor model surrogate with an anatomy encoder that enforces patient-specific boundary conditions. This enables a fast Bayesian model personalization that is consistent with the baseline numerical solver.

II. METHOD

1) Forward tumor model: The simulations that we aim to emulate are generated by a 3D numerical solver relying on the Fisher-Kolmogorov type of partial-differential equations (PDE). The equation describes the evolution of the pathology by considering diffusion and proliferation of the tumor cells under the Neumann boundary condition (B.C.):

$$\frac{\partial \mathbf{u}}{\partial t} = \nabla(\mathbf{D}\nabla\mathbf{u}) + \rho\mathbf{u}(1-\mathbf{u}), \tag{1}$$

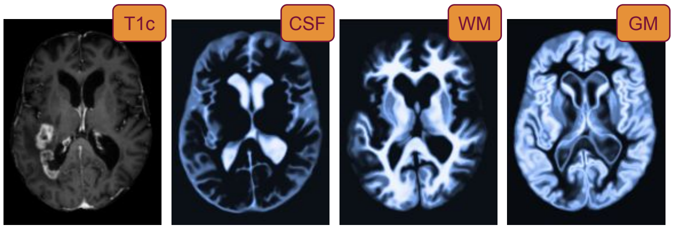

Fig. 2: An example of an MRI T1c scan from the dataset and corresponding probabilistic segmentations offering information on the expected tissues underneath and nearby the tumor (obtained by registering the brain atlas to patient space).

$$\nabla\mathbf{u}\cdot\mathbf{n} = 0 \quad B.C. \tag{2}$$

Here, $\mathbf{u}$ is the normalized 3D tumor cell density, $\mathbf{D}$ denotes the diffusion tensor, ρ is the rate of cell proliferation, and $\mathbf{n}$ is the normal vector to the boundary. We assume the infiltration of the tumor cells to occur only in white matter (WM) and grey matter (GM), considering isotropic diffusion with $\mathbf{D} = D \cdot \mathbb{I}$ (where $D \in \{D_w, D_g\}$ is a diffusion coefficient in white or grey matter, and $\mathbb{I}$ is an identity matrix). Both WM and GM constitute the simulation domain while cerebrospinal fluid (CSF) and skull determine the patient-specific boundary. The segmentations are extracted from medical scans of patients diagnosed with the tumor. Given the probabilistic nature of the WM and GM segmentation maps used here, the diffusion coefficient is defined as $D_i = p_{w_i} D_w + p_{g_i} D_g$, where p_{w_i} and p_{g_i} denote the percentage of the WM and GM at the voxel i. The diffusion coefficient in white matter D_w is considered to be greater than in grey matter. We considered two models with the ratio D_w/D_g equal 10 and 100. The input to the solver is a set of parameters $\theta_P = \{D_w, \rho, x, y, z, T\}$, where x, y, z define the position of the function $\mathbf{u}$ at time $T = 0$, which is initialized as a point source.

2) Learnable forward model surrogate with anatomy encoder: Our goal is to learn a tumor model surrogate which maps parameters of the pathology model θ_P to corresponding simulations $\mathbf{u}(\theta_P)$ given a patient-specific anatomy. We base our method on [35] which is designed to emulate the mapping for fluid simulations in a static spatial domain. Different from [35], we need to consider patient-specific boundary conditions. To this end, we introduce an anatomy encoder that imposes anatomical boundary conditions, Fig. 1.

The numerical solver's output $\mathbf{u}(\theta_P)$ has a size of $128 \times 128 \times 128$ voxels. However, to provide a greater anatomical variability to the dataset on which we train the surrogate, we crop all simulated outputs and corresponding brain anatomies to $64 \times 64 \times 64$ volumes, centered at the initialization location x, y, z. The $64 \times 64 \times 64$ is greater than half the brain size and tumors bigger than this are incompatible with life.

The architecture of the tumor model surrogate consists of the following parts:

- *Brain anatomy encoder* which encodes non-overlapping 3D volumes of the brain tissues WM, GM, and CSF through a series of convolutional blocks. The blocks are composed of alternating convolution operations (with fixed parameters of kernel size 3, stride 1, and the number of channels 128) and a non-linearity in the form of a linear rectifier. Each block is equipped with a skip connection linking the input and output of the block via an element-wise sum. Downsampling between the blocks is achieved by a convolutional operation with a stride of two,

- *Brain tumor decoder* that takes a 1D vector of the parameters $\{D, \rho, T\}$ alongside with the encoded anatomy as input. Note that we do not condition the decoder on the initialization location x, y, z, since as mentioned above we crop all training volumes exactly at this location. Thus, the network is taught to reproduce the tumor in the center for any volume. Before being passed to the decoder, the 1D vector of the model parameters is mapped via a fully connected layer to a tensor of size $8\times 8\times 8\times 3$ and is concatenated with the tensor of the encoded brain anatomy. The resulting tensor is gradually upsampled through a series of convolutional blocks analogous to the encoder and nearest-neighbor upsampling (we refer the reader to [39] for a detailed discussion on why such type of upsampling is preferred over the deconvolution operation). At the end of the series, a 4D tensor $64 \times 64 \times 64 \times 128$ is convolved to the output prediction - a 3D tumor simulation volume (height 64, width 64, depth 64). The decoder design is adopted from [35], [40].

For constructing the *loss function*, we can consider several semantically different sections of the output volume: CSF area (because no tumor cells are to be expected here), tumor area (because the algorithm should focus on this area ignoring the significantly larger normal brain), remaining brain area (composed of WM and GM voxels not present under tumor area), background (BG). We experimented with different combinations of this compartments in the loss, Tab. 1. The following loss definition outperforms other ways of defining spatial dependency in the cost function:

$$L_{\mathbf{total}} = \|u_{\mathbf{sim}} - u_{\mathbf{pred}}\|_1^{tumor} + \|u_{\mathbf{sim}} - u_{\mathbf{pred}}\|_1^{CSF} \tag{3}$$

where the error between the predicted by the surrogate ($u_{\mathbf{pred}}$) and simulated by the numerical solver ($u_{\mathbf{sim}}$) cell concentration is computed under L_1 norm separately in the *tumor* area and in the *CSF* area.

3) Bayesian model personalization: To demonstrate the applicability of the neural surrogate, we perform Bayesian tumor growth model personalization substituting the numerical solver with the learnable one. As calibration data we use two types of imaging modalities: (a) T1 contrast-enhanced and FLAIR MRI modalities that allow estimating the morphological characteristic of the visible tumor; (b) FET-PET scans that provide information about the metabolic activity of the tumor.

Analogous to [15], [16], we relate the output of the tumor growth solver $\mathbf{u}(\theta_P)$ to imaging information via a probabilistic model,

$$\begin{aligned} p(\mathcal{D} \mid \mathbf{u}, \theta) = p\left(\mathbf{y}^{T1c} \mid \mathbf{u}, \theta_I\right) \cdot p\left(\mathbf{y}^{FLAIR} \mid \mathbf{u}, \theta_I\right) \cdot \\ \cdot\, p\left(\mathbf{y}^{PET} \mid \mathbf{u}, \theta_I\right) \end{aligned} \tag{4}$$

where image observations $\mathcal{D} = \{\mathbf{y}^{T1c}, \mathbf{y}^{FLAIR}, \mathbf{y}^{PET}\}$ are assumed to be independent, and $\theta = \{\theta_P, \theta_I\}$ constitute parameters of the pathology model θ_P and the probabilistic imaging model θ_I. The latter is defined differently according to the type of modality:

- *MRI modalities* provide information in the form of binary tumor segmentations ($\mathbf{y}^{T1c}, \mathbf{y}^{FLAIR}$). Thus we assign for each voxel a discrete label $y_i^M \in \{0, 1\}$, $M \in \{T1c, FLAIR\}$. We model the probability of observing $\mathbf{y}^{T1c}, \mathbf{y}^{FLAIR}$ for a given concentration $\mathbf{u}(\theta_P)$ with Bernoulli distribution,

$$\begin{aligned} p\left(\mathbf{y}^{T1c,FLAIR} \mid \mathbf{u}, \theta_I^M\right) &= \prod_i p\left(y_i^M \mid u_i, \theta_I^M\right) = \\ &= \prod_i \alpha_{i,M}^{y_i^M} \cdot \left(1 - \alpha_{i,M}\right)^{1-y_i^M}, \end{aligned} \tag{5}$$

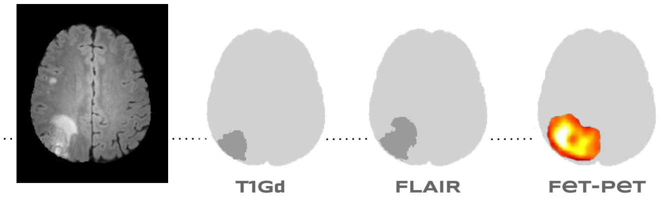

Fig. 3: An illustration of the imaging information used for Bayesian model calibration: binary segmentations obtained from T1Gd and FLAIR modalities, and FET-PET signal which is proportional to the tumor density.

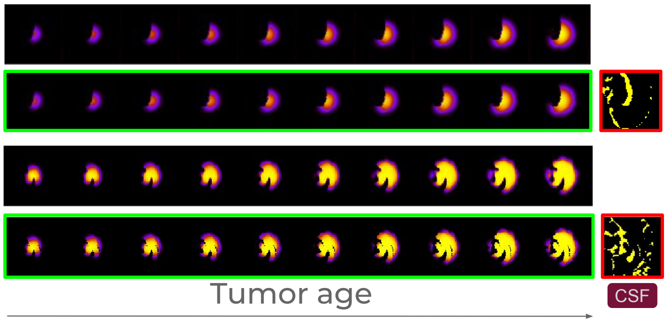

Fig. 4: Comparison between tumor volumes predicted by the surrogate and numerical solver. Each column represent a different value of tumor age T from 100 to 1000 days with intervals of 100 days. The images were obtained by inferring individual 3D volumes for all time points and taking a central 2D slice from each volume. The rows framed in green correspond to the ground truth simulation. The CSF delineation constraining the tumor growth is framed in red.

where $\alpha_{i,M}$ is the probability of observing tumor-induced changes defined as a double logistic sigmoid,

$$\alpha_{i,M}\left(u_i,u_c\right)=0.5+0.5\cdot\text{sign}\left(u_i-u_c^M\right)\left(1-e^{-\frac{\left(u_i-u_c^M\right)^2}{\sigma_\alpha^2}}\right) \tag{6}$$

With this formulation, we postulate that the tumor is not visible on a scan below the threshold level u_c^M. The parameter σ_α is introduced to take uncertainty in the threshold concentration u_c^M into account.

- *FET-PET modality* ($\mathbf{y}^{PET}$) provides continuous information in each voxel and can be assumed to be proportional to the tumor density [3], [41], [42]. In this case we model the likelihood imaging function by a Gaussian distribution with unknown constant of proportionality b:

$$p\left(\mathbf{y}^{PET}\mid\mathbf{u},\theta_I^M\right)=\prod_i p\left(y_i\mid u_i,\theta_I^M\right)=\prod_i\mathcal{N}\left(y_i-bu_i,\sigma^2\right) \tag{7}$$

Here, σ models uncertainty in the PET signal, which is considered to be normalized $y_i^{PET}\in[0,1]$.

In total there are eleven parameters (six pathology model parameters $\theta_P=\{D_w,\rho,x,y,z,T\}$ and five imaging model parameters $\theta_I=\{u_c^{T1c},u_c^{FLAIR},\sigma_\alpha,b,\sigma\}$) which we infer from the triplet of medical scans $\mathcal{D}=\{\mathbf{y}^{T1c},\mathbf{y}^{FLAIR},\mathbf{y}^{PET}\}$ using a Markov Chain Monte Carlo (MCMC) sampling algorithm [43].

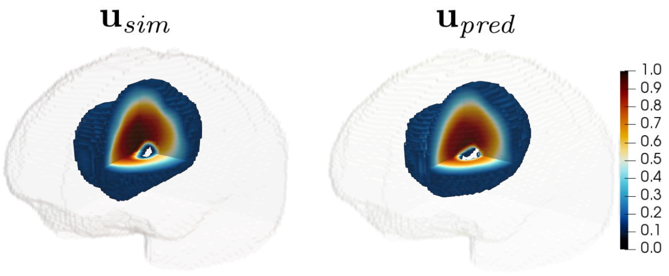

Fig. 5: A qualitative comparison between the simulated and predicted tumors embedded in the 3D brain anatomy.

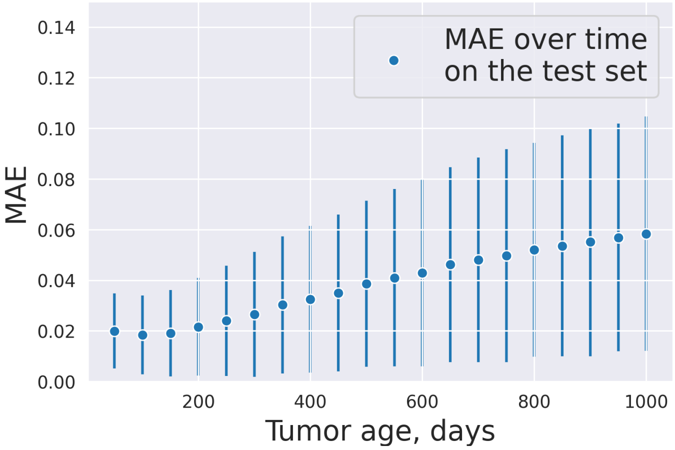

Fig. 6: The mean absolute error analysis over the tumor age (an input parameter) on the in-house test set. The error between the tumors (from numerical and neural solvers) is computed analogous to the first term in the proposed loss - in the area where the tumor concentration simulated using the numerical solver is greater than zero. Dots depict the means of the MAE distribution, error bars - the standard deviations. The analysis shows minor error increase with the increase of the tumor size.

4) Implementation: The numerical tumor solver used for obtaining the simulation dataset is a highly parallelized glioma solver returning a 3D normalized tumor concentration profile on a uniform spatial grid [3].

The surrogate network is trained using the Adam optimizer with decay rates $\beta_1=0.5$ and $\beta_2=0.999$ for 30 epochs, which was observed to be sufficient for convergence. The learning rate is cosine annealed from 10^{-4} to 2.5×10^{-6} over the training and the batch size is 16. The input parameters' ranges are min-max normalized. The training time is about 6 days on an NVIDIA Quadro P8000 using the Tensorflow framework.

For Bayesian MCMC inference we use an implementation of Transitional MCMC from [44] with 2048 samples per iteration. When the MCMC is used with the numerical solver, simulations of $128\times128\times128$ size are calibrated against the imaging information of the corresponding size. In the case of the surrogate, the network predicts $64\times64\times64$ tumor volumes, which are then embedded in the original $128\times128\times128$ domain for the calibration.

III. EXPERIMENTS

Conventional application of tumor modeling such as model personalization (via solving an inverse problem) implies estimating the model parameters by sampling over fixed, physiologically plausible ranges. Thus, we aim to employ the learnable surrogate which i) possesses an interpolation capacity for the parameters $\{D,\rho,T\}$, and ii) is capable to extrapolate for simulations in new geometries. To probe these properties, the validation and test sets were formed to have only the brain

	Validation set			Test set		
	$u_c = 0.2$	$u_c = 0.4$	$u_c = 0.8$	$u_c = 0.2$	$u_c = 0.4$	$u_c = 0.8$
Ablation analysis on the network structure						
(a) With skip-connections (U-net)	0.572	0.575	0.576	0.563	0.565	0.569
(b) Only CSF as input	0.739	0.734	0.734	0.768	0.767	0.763
Ablation analysis on the loss function						
(c) [Tumor+CSF+Brain+BG]	0.615	0.607	0.585	0.588	0.571	0.544
(d) Tumor+[CSF+Brain+BG]	0.755	0.761	0.762	0.787	0.782	0.778
(e) Tumor+CSF+Brain	0.771	0.764	0.762	0.791	0.786	0.780
(f) Tumor+CSF	**0.796**	**0.783**	**0.788**	**0.802**	**0.795**	**0.794**

TABLE I: Ablation analysis on the in-house validation (2k samples) and test (10k samples) tests: a) with skip connections between the encoder and decoder (U-Net like); b) instead of inputting 3 volumes of different tissue types (WM, GM, CSF) only a single volume of CSF tissue serves as an input to the network; c) a single loss is computed for the whole volume; d) the second term in the loss is computed in the spatial complement of the first *tumor* term in Eq. 3; e) in addition to the two terms in Eq. 3, a third term is added to penalize for false predictions in the remaining brain, f) the proposed architecture. The values represent the means of the DICE score histograms. The DICE is computed for the tumor volumes $\mathbf{u}_{pred}$ and $\mathbf{u}_{sim}$ thresholded at 0.2, 0.4, and 0.8 values of tumor cell concentration. The square brackets denote the areas included in a single loss term, e.g. for [CSF+Brain+BG] the term is computed in the volume combining all three compartments.

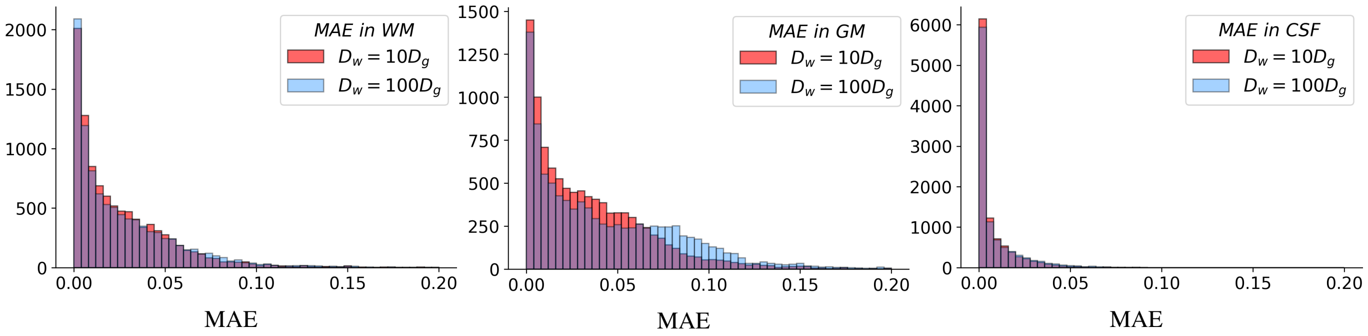

Fig. 7: Histograms of the mean absolute error $\|\mathbf{u}_{pred} - \mathbf{u}_{sim}\|_1$ computed on the in-house test set (10k samples) within each class of the brain tissues (WM, GM, CSF). In red, samples from the tumor model with $D_w = 10D_g$, in blue - from the tumor model with $D_w = 100D_g$. The means of the distributions for the $D_w = 10D_g$ model are: 0.026 (WM), 0.035 (GM), 0.007 (CSF). The means for the $D_w = 100D_g$ model are : 0.028 (WM), 0.044 (GM), 0.008 (CSF). The fact that the network performance on two tumor models with notably different diffusion character stays close suggests that the network is capable of capturing not only pixel-wise CSF dependency, but also highly non-local diffusion dependency on the WM and GM.

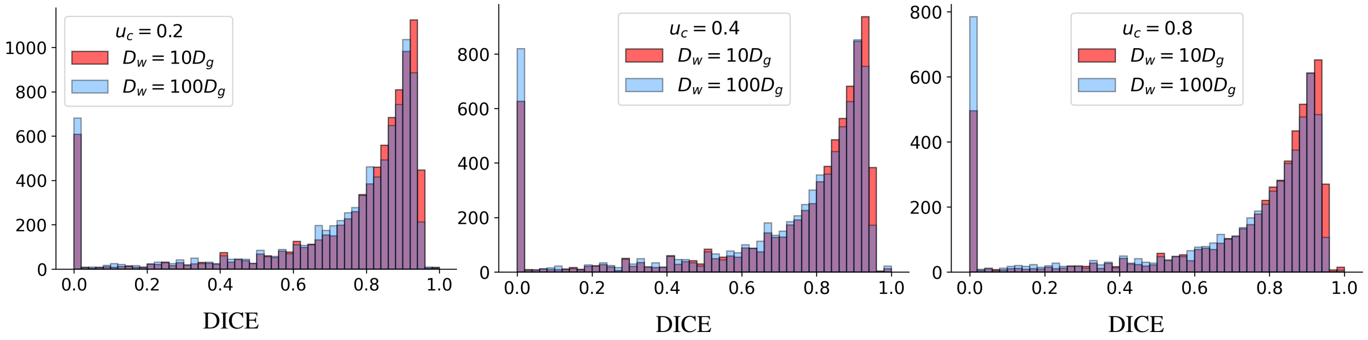

Fig. 8: Histograms of the DICE score computed on the test set (10k samples) for the tumor volumes $\mathbf{u}_{pred}$ and $\mathbf{u}_{sim}$ thresholded at 0.2, 0.4, and 0.8 values of tumor cell concentration. In red, samples from the tumor model with $D_w = 10D_g$, in blue - from the tumor model with $D_w = 100D_g$. The means of the distributions for the $D_w = 10D_g$ model are: 0.802 ($u_c = 0.2$), 0.795 ($u_c = 0.4$), 0.794 ($u_c = 0.8$). The means for the $D_w = 100D_g$ model are: 0.782 ($u_c = 0.2$), 0.775 ($u_c = 0.4$), 0.764 ($u_c = 0.8$).

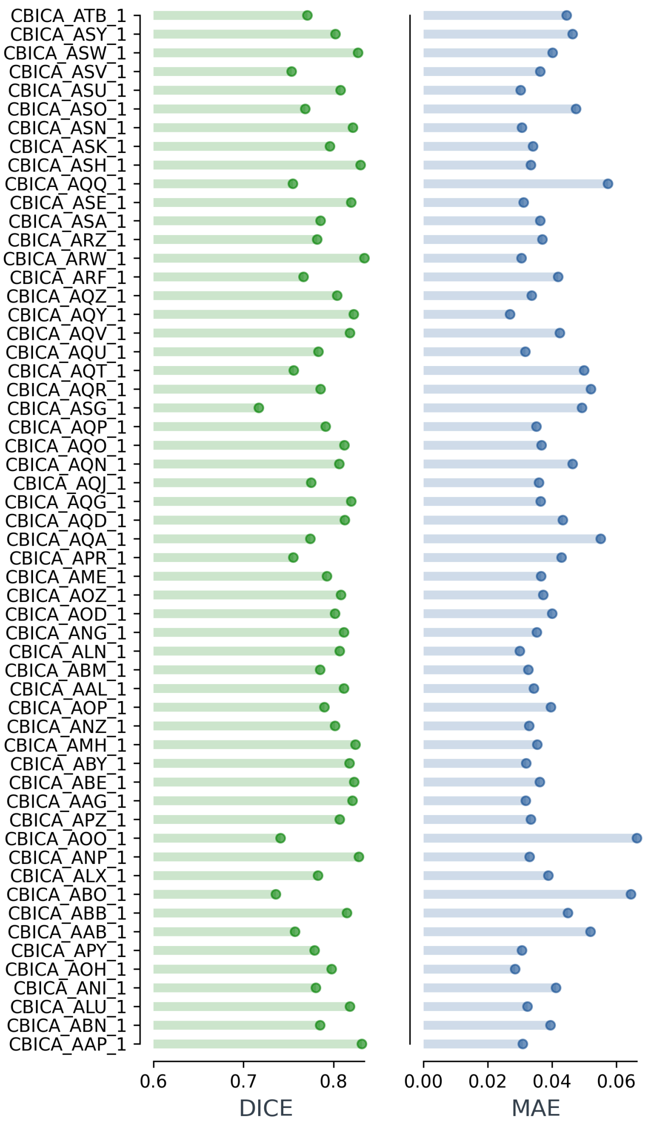

Fig. 9: Distribution of the per-patient mean errors between the tumors predicted by the numerical and neural solvers over 56 patients from the BraTS dataset. The computed errors are DICE (thresholded at $u_c = 0.2$) and MAE. For each patient the mean DICE and MAE are computed on 1k tumor simulations. Such evaluation on an independent test set demonstrates that the proposed neural surrogate possesses a high generalization ability that is stable across a high number of tumor and patient anatomies.

anatomies unseen by the network during training, while the parametric $\{D, \rho, T\}$ triplets were sampled from the same ranges as for the training.

1) Data and parameterization: We simulate the tumors in probabilistic brain tissue (WM, GM, CSF) segmentations that are obtained by registering a patient scan to the healthy brain atlas [3], [45] (available at github.com/JanaLipkova/s3). Different from a plain segmentation of the patient anatomy directly from an MRI scan, such procedure offers an estimate of the anatomy underneath the tumor. Fig.2 shows an example of such an approximation. We use two datasets: a) *in-house set* composed of MRI and PET images from 20 patients (simulations in 10 patients' anatomies are used as training data, simulations in additional 5 patients for validation, and in the remaining 5 patients for test), b) all CBICA 56 patients of high-resolution MRI from *BraTS* [46] (56k simulations are used only for test). Both datasets have resolution of isotropic voxels of 1mm side-length, but for simulation purposes, the calibration is performed in data downsampled to 128x128x128 that corresponds to 2mm per side-length.

For most of experiments, we considered a tumor model in which the diffusion in white matter D_w is greater than in grey matter D_g by 10 times (except Fig. 7 and 8, where two diffusion models are compared - with the ratio of 10 and of 100 between the coefficients). The following ranges are used for random uniform sampling of the model parameters: $D_w \in [0.01, 0.08]$ mm^2/day, $\rho \in [0.0001, 0.03]$ $1/day$, and $T \in [50, 1000]$ day with a step size of 50 days. The tumor location coordinates x, y, and z are sampled uniformly within the brain volumes. It should be noted that this uniform prior can be modified to non-uniform distributions, such as those reported in [1], [47], to increase sampling for locations with higher tumor incidence. The initial locations $\{x, y, z\}$ were samled within the WM and GM only. For samples with initial location closer than 32 pixels to any of the volume borders, we do extra padding when cropping to 64x64x64 size. In total, we have a set of 20k parameters-simulation pairs for training.

2) Identifying the optimal algorithm on in-house dataset: We perform several experiments on the in-house dataset to identify an optimal architectural setup:

a) Experiments on the loss function: As mentioned in the method section the predicted 3D image can be decomposed into several semantic parts (CSF, tumor, WM/GM, background). An ablation analysis demonstrates that the highest performance is achieved using the loss that focuses solely on the tumor and CSF areas, Tab. 1f. Notably, such loss is superior despite absence of penalization in rest of the volume.

b) Experiments on the network structure: These days in the field of medical imaging computing, the natural baseline is a U-Net. We performed experiments comparing the U-net type architecture with the proposed encoder-decoder design. Fig. 10 depicts the problem with the U-Net, namely, worse capturing of the output volume's dependency on the model parameters, and more pronounced false predictions in the non-tumor area.

We also analyzed how much information the proposed network is capable to extract from the WM/GM signal. Results from Tab.1 show that despite the highly non-local character of dependency of the output tumor volume on the WM/GM, the performance is dropped by up to 6% if the WM/GM is not included in the input.

c) Experiments on robustness: We perform experiments comapring two diffusion models with notably different diffusion character ($D_w/D_g = 10$, $D_w/D_g = 100$). Fig. 7 demonstrates the distribution of the mean absolute error within each class of the brain tissues evaluated on the whole in-house test set (10K samples) for two types of diffusion models. Even though we observe samples with the error in the order of 10^{-1}, the majority of the distribution lies within the order of 10^{-2}. In Fig. 8 we depict the histograms of the DICE score

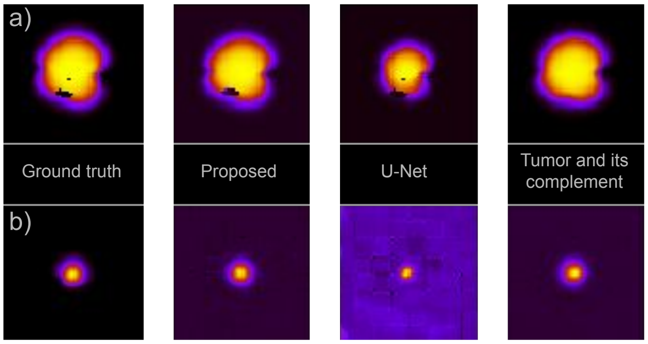

Fig. 10: Visual comparison of the inference obtained by the proposed method, U-Net, and the architecture with two terms in the loss - tumor and tumor complement (as in ”d” in Tab. 1). The U-net exhibits mispredictions of two types: a) better CSF preservation is traded off for worse capturing of the parametric dependency, b) more notable false predictions in the out-of-tumor area. The proposed architecture can also produce false predictions in the non-CSF area (which is expected due to absence of related penalization in the loss). However, using penalization for the tumor complement area does not warrant absence of such noise (”Tumor and complement”).

$(2|X \cap Y|/(|X|+|Y|))$ computed between the simulated and predicted tumor volumes which are thresholded at different levels of the tumor cell concentration (such thresholding is exactly the operation that we perform during the Bayesian inference to relate tumor concentration with MRI signals, Eq. 6). The distributions are centered close to the DICE 1.0. The smaller peak at DICE 0.0 is due to the fact that thresholding of some simulations results in volumes containing a few non-zero voxels, whereas thresholding of the corresponding network predictions outputs volumes of all-zero voxels (or vice versa). In our particular application of the tumor model personalization such volumes are significantly smaller that the binary MRI segmentation volumes to which we calibrate the model and thus do not affect the Bayesian inference outcome. In sum, given that the performance of two different tumor models with notably different diffusion character stays within 2%, we conclude that not only effects imposed by the patient specific CSF anatomy, but also higher order dependency on the WM and GM (through which the diffusion tensor is defined) are captured by the proposed surrogate.

Fig. 6 depicts the dependency of the mean absolute error versus a network's input - tumor age, showcasing somewhat expected error increase with the increase of the tumor size. Fig. 4 qualitatively illustrates the surrogate's predictions over tumor age.

3) Independent testing on the BraTS dataset: As mentioned above, to provide greater anatomical variability to the training set we performed the training on 64x64x64 crops randomly sampled over the brain volumes of the 10 patients (20k samples in total). That means we heavily randomize over the field of view, inducing variability of the simulation domain that is much larger than what one would see from inter-patient variability (even on the large datasets like BraTS). Thus we argue that training on 20k samples of such data is sufficient for generalisability to an arbitrary test set generated by the same solver within the fixed parametric ranges. To examine our point, we performed an independent test of the proposed network's generalization performance on a large subset from BraTS composed of all CBICA patients. This test, Fig. 9, demonstrates generalization capacity that is stable across a high number of tumor and patient anatomies.

4) Bayesian model personalization in patient data: As a final test of the surrogate, we performed the Bayesian brain tumor model calibration. We ran the inference on preoperative scans of 5 patients from the in-house validation set using the proposed neural surrogate and numerical solver. The max-a-posterior (MAP) estimates of the tumor density are provided in Tab.2 and tumor concentrations modeled with the MAP estimates are shown in Fig. 11. We observe an agreement of the glioma profiles obtained by the two methods and the variability of the estimations is within the variability of the Bayesian calibration.

We also observe that the learnable surrogate is trained on a dataset with continuous uniform distribution of the diffusion coefficient D and the proliferation rate ρ, whereas the T parameter has a discrete interval. However, during the model calibration, the time parameter is sampled from a continuous interval. This implicitly suggests that the network's interpolation capacity is sufficient for parametric inference.

5) Performance speed-up: The computing time for a single simulation using the numerical solver on 64x64x64 grid is on average 15 seconds using an Intel Xeon with 8 CPU cores and 64GB of RAM. The time required by using the neural solver for processing a batch of size 8 during inference is equal to 2.4s, i.e. our surrogate is 50 times faster than the numerical solver per single simulation. Such a speed-up proportionally reduces the computing time for the model personalization.

We want to point out that a theoretical comparison between CPU- and GPU-based computations should be taken with care for a few reasons. First, the neural surrogate does not need to do any sequential computations and can batch process multiple frames, whereas the numerical solver is successively integrating the PDE equation in time and for a single sample. Second, the convolutional layers are very efficiently realized in most GPU-tailored deep learning frameworks. In this regard, a fair comparison could imply to contrast the GPU compute with a CPU cluster containing as many processing units as the GPU. However, the latter is rarely available for practical applications, and as our main goal is clinical translation, we compare CPU and GPU platforms that are widely available. Also, it is worth to mention that ideally the comparison should be with a numerical solver optimized to run on GPUs.

IV. DISCUSSION

As depicted in Fig. 8, even though the Dice distributions are peaked close to a perfect score, there is room for improvement. For example, the anatomy encoder well captures the global growth constraints imposed by the CSF anatomy, but tiny CSF components of a few voxels size are often missing in the network predictions, Fig. 4. From our experiments we

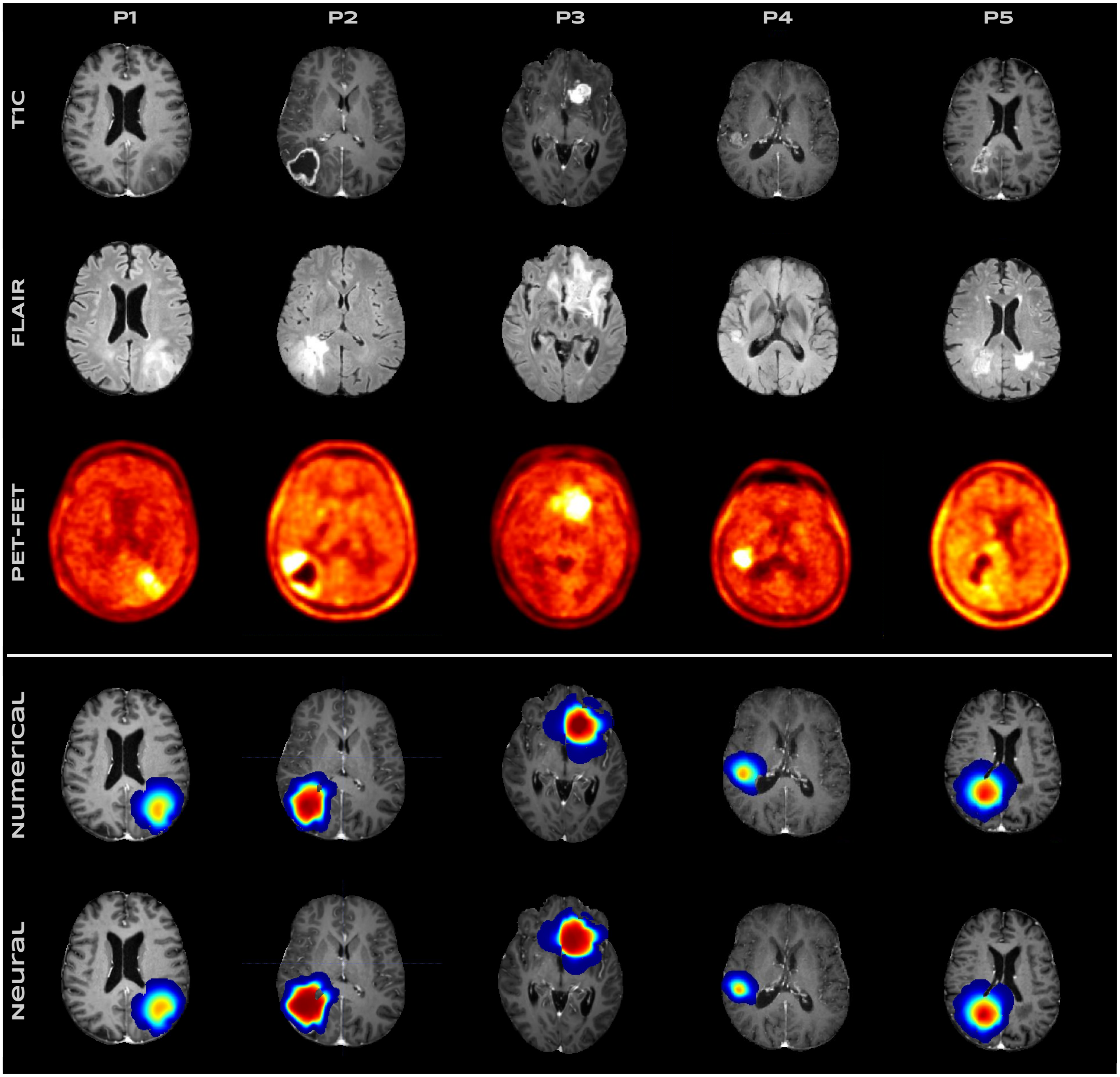

Fig. 11: Results of the Bayesian inference for five patients to which we refer as P1-P5. The three upper rows correspond to the imaging modalities used for tumor model calibration, namely T1c, FLAIR, and FET-PET. The two bottom rows show the simulations of glioma with model parameters inferred via the Bayesian inference using the numerical solver and proposed neural surrogate.

observe, that the U-Net type model better preserves the CSF, Fig. 10. However, in the rest of the volume, the U-Net outputs mispredictions of two types: a) better CSF preservation is traded off for worse capturing of the parametric dependency, b) pronounced false predictions in out-of-tumor area. This leads to a significant performance decrease, Tab. 1. The U-Net is known to work well on datasets which preserve voxel-wise semantics between the network's input and output [48]–[51]. In our case, though, two out of three input anatomies (WM, GM) do not affect the output (simulated 3D tumor) in a voxel-wise fashion, but rather in a highly non-local way the whole 3D tumor volume. Arguably, the mispredictions (at least type "a") may be explained by the higher emphasis on the translation of the anatomy input signal via skip-connections which makes the mapping from the tissues dominate the second input-output mapping - from parameters (D, ρ, T) to the tumor volumes. A more intelligent separation of the input-output mappings or hybrid approaches [37] explicitly enforcing boundary conditions might alleviate such issues.

Another aspect we want to point out is that we do not empirically observe pronounced false predictions in the non-CSF area, even without explicit penalization, as in the proposed

		D	ρ	T	x	y	z	σ	b	u_c^{T1c}	u_c^{FLAIR}	σ_α	**MAE**
P1	Numerical solver	0.08994	0.01550	572.582	0.3098	0.6748	0.2819	0.2177	0.6384	0.7976	0.3718	0.0514	0.09
	Neural surrogate	0.08976	0.01657	553.019	0.3071	0.6657	0.2818	0.1702	0.7198	0.7344	0.3364	0.0500	
P2	Numerical solver	0.08920	0.01993	503.146	0.6121	0.6706	0.3411	0.2358	0.7309	0.6502	0.4376	0.0688	0.18
	Neural surrogate	0.08945	0.02882	418.631	0.6217	0.6445	0.3316	0.2434	0.6324	0.6316	0.4337	0.0799	
P3	Numerical solver	0.08998	0.01555	572.128	0.4129	0.3544	0.2900	0.2326	0.7023	0.6088	0.3608	0.0735	0.15
	Neural surrogate	0.08990	0.02897	418.946	0.4230	0.3678	0.2901	0.2498	0.6561	0.6294	0.3662	0.0772	
P4	Numerical solver	0.07967	0.00813	733.13	0.6421	0.5502	0.3151	0.1832	1.0146	0.6007	0.5458	0.0757	0.06
	Neural surrogate	0.04362	0.00817	776.41	0.6707	0.5613	0.3224	0.1966	1.0136	0.6237	0.5192	0.0579	
P5	Numerical solver	0.08938	0.01080	686.923	0.5530	0.6203	0.3147	0.2484	0.6000	0.6023	0.3775	0.0796	0.05
	Neural surrogate	0.08360	0.01066	687.841	0.5701	0.6259	0.3276	0.2411	0.7044	0.6098	0.3790	0.0759	

TABLE II: MAP estimates of the marginal distribution from the Bayesian calibration on the patient data P1-P5 (from Fig. 11) using the numerical solver and proposed neural surrogate. The prior ranges are chosen as follows: $D_w \in [0.01, 0.08]\ mm^2/day$, $\rho \in [0.0001, 0.03]\ 1/day$, $T \in [30, 1000]\ days$, $\sigma \in [0.01, 0.25]$, $b \in [0.6, 1.02]$, $u_c^{T1c} \in [0.6, 0.8]$, $u_c^{FLAIR} \in [0.05, 0.6]$, and $\sigma_\alpha \in [0.05, 0.08]$. The MAE between the predicted and simulated tumors is computed analogous the 1st term of the network loss - in the area where the tumor from the numerical solver is greater than zero. We observe an agreement of the model parameters obtained by the two methods and the variability of the estimations is within the variability of the Bayesian calibration.

loss definition. Moreover, penalization for false predictions in the loss does not improve performance (Tab. 1). We attribute the superiority of the loss with only two compartments to the fact that adding extra compartments to the loss introduces a "conflict" of gradients - the optimization focus distorts from the tumor and CSF areas, which are the only areas that affect the numerical simulations. We do sometimes observe low-level noise in the non-penalized area, but at the same time, as Fig. 10b also shows, using penalization for the tumor complement area does not warrant absence of such noise.

Next, further research would benefit from considering different scenarios for the diffusion process with respect to its homogeneity and isotropy to evaluate the proposed network's learning capacity. In our work, we consider inhomogeneity by using spatially varying diffusion tensor (Sec. II-1). However, the question of how to consider anisotropy and how to introduce and weight the off-diagonal elements of the diffusion tensor into the simulation we believe is an open one. It is also closely related to how to model the Diffusion Tensor Imaging signal itself (the common ellipse, only the first eigencomponent to suppress the tumor edema effect, a stick and ball model). We felt this questions have not been answered well yet elsewhere, and thus we focused on isotropy only.

Lastly, we do not have a definite answer to whether the mentioned limitations and resulting approximation errors are acceptable for clinical translation. Neither do we know whether the Bayesian calibration itself under the simplistic Fisher-Kolmogorov formalism is suitable for the translation. Both require a study on a large cohort of patients, post-surgery analysis, etc. [1]–[4]. However, we find the proposed method to be a solid baseline in the search for optimal tumor model surrogate, which in turn can significantly speed up our search for a biophysical model descriptive enough for clinical trials.

V. CONCLUSION

We present the first learnable surrogate with anatomy encoder for tumor growth modeling that is capable of mapping the model parameters to the corresponding simulations while respecting patient-specific anatomy. Our method achieves real-time simulation 50× faster than numerical solver. Since the learnable surrogate was tested in general PDE settings, we believe the technique can be adopted to more complicated tumor growth models and similar 4D inverse modeling problems.

REFERENCES

[1] J. Unkelbach *et al.*, "Radiotherapy planning for glioblastoma based on a tumor growth model: improving target volume delineation," *Physics in Medicine and Biology*, vol. 59, no. 3, pp. 747–770, 2014.

[2] R. Rockne *et al.*, "A patient-specific computational model of hypoxia-modulated radiation resistance in glioblastoma using 18f-fmiso-pet," *JRS, Interface*, vol. 12, 2015.

[3] J. Lipkova *et al.*, "Personalized radiotherapy design for glioblastoma: Integrating mathematical tumor models, multimodal scans and bayesian inference," *IEEE Transactions on Medical Imaging*, pp. 1–1, 2019.

[4] E. Konukoglu, O. Clatz, H. Delingette, and N. Ayache, "Personalization of reaction-diffusion tumor growth models in mr images: Application to brain gliomas characterization and radiotherapy planning," *Multiscale Cancer Modeling*, 12 2010.

[5] M. Le *et al.*, "Personalized radiotherapy planning based on a computational tumor growth model," *IEEE Transactions on Medical Imaging*, vol. 36, no. 3, pp. 815–825, mar 2017.

[6] S. Subramanian, K. Scheufele, N. Himthani, and G. Biros, "Multiatlas calibration of biophysical brain tumor growth models with mass effect," *arXiv preprint arXiv:2006.09932*, 2020.

[7] K. Scheufele, S. Subramanian, and G. Biros, "Automatic mri-driven model calibration for advanced brain tumor progression analysis," *arXiv: Medical Physics*, 2020.

[8] K. Scheufele, S. Subramanian, A. Mang, G. Biros, and M. Mehl, "Image-driven biophysical tumor growth model calibration," *arXiv: Quantitative Methods*, 2019.

[9] L. Zhang, L. Lu, R. M. Summers, E. Kebebew, and J. Yao, "Convolutional invasion and expansion networks for tumor growth prediction," *IEEE transactions on medical imaging*, vol. 37, no. 2, pp. 638–648, 2017.

[10] J. Petersen *et al.*, "Deep probabilistic modeling of glioma growth," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2019, pp. 806–814.

[11] E. Ahmed *et al.*, "Macroscopic cerebral tumor growth modeling from medical images: A review," *IEEE Access*, vol. 6, pp. 30 663–30 679, 2018.

[12] A. Mang, S. Bakas, S. Subramanian, C. Davatzikos, and G. Biros, "Integrated biophysical modeling and image analysis: Application to neuro-oncology," *Annual Review of Biomedical Engineering*, vol. 22, pp. 309–341, 2020.

[13] O. Clatz *et al.*, "Realistic simulation of the 3-d growth of brain tumors in mr images coupling diffusion with biomechanical deformation," *IEEE transactions on medical imaging*, vol. 24, no. 10, pp. 1334–1346, 2005.

[14] P. R. Jackson, J. Juliano, A. Hawkins-Daarud, R. C. Rockne, and K. R. Swanson, "Patient-specific mathematical neuro-oncology: Using a simple proliferation and invasion tumor model to inform clinical practice," *BMB*, vol. 77, no. 5, pp. 846–856, mar 2015.

[15] B. Menze *et al.*, "A generative approach for image-based modeling of tumor growth," in *IPMI*, 2011, pp. 735–747.

[16] M. Le *et al.*, "Bayesian personalization of brain tumor growth model," in *MICCAI*, 2015, pp. 424–432.

[17] I. Ezhov *et al.*, "Neural parameters estimation for brain tumor growth modeling," in *MICCAI*, 2019, pp. 787–795.

[18] V. Cristini and J. Lowengrub, *Multiscale modeling of cancer: an integrated experimental and mathematical modeling approach*. Cambridge University Press, 2010.

[19] J. Kadupitiya, F. Sun, G. Fox, and V. Jadhao, "Machine learning surrogates for molecular dynamics simulations of soft materials," *Journal of Computational Science*, p. 101107, 2020.

[20] Z. Li *et al.*, "Fourier neural operator for parametric partial differential equations," *arXiv preprint arXiv:2010.08895*, 2020.

[21] B. Moseley, A. Markham, and T. Nissen-Meyer, "Solving the wave equation with physics-informed deep learning," *arXiv preprint arXiv:2006.11894*, 2020.

[22] C. Rao, H. Sun, and Y. Liu, "Physics-informed deep learning for incompressible laminar flows," *arXiv preprint arXiv:2002.10558*, 2020.

[23] B. Lütjens *et al.*, "Physics-informed gans for coastal flood visualization," *arXiv preprint arXiv:2010.08103*, 2020.

[24] X. Jin, S. Cai, H. Li, and G. E. Karniadakis, "Nsfnets (navier-stokes flow nets): Physics-informed neural networks for the incompressible navier-stokes equations," *arXiv preprint arXiv:2003.06496*, 2020.

[25] Z. Fang and J. Zhan, "Deep physical informed neural networks for metamaterial design," *IEEE Access*, vol. 8, pp. 24 506–24 513, 2019.

[26] J. Han, L. Zhang, and E. Weinan, "Solving many-electron schrödinger equation using deep neural networks," *Journal of Computational Physics*, vol. 399, p. 108929, 2019.

[27] C. Beck, S. Becker, P. Cheridito, A. Jentzen, and A. Neufeld, "Deep learning based numerical approximation algorithms for stochastic partial differential equations and high-dimensional nonlinear filtering problems," *arXiv preprint arXiv:2012.01194*, 2020.

[28] E. Geremia, B. H. Menze, M. Prastawa, M.-A. Weber, A. Criminisi, and N. Ayache, "Brain tumor cell density estimation from multi-modal mr images based on a synthetic tumor growth model," in *International MICCAI Workshop on Medical Computer Vision*. Springer, 2012, pp. 273–282.

[29] K. Dmitrii *et al.*, "Machine learning accelerated computational fluid dynamics," *arXiv preprint arXiv:2102.01010*, 2021.

[30] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *Journal of Computational Physics*, vol. 378, pp. 686–707, 2019.

[31] V. Sitzmann *et al.*, "Implicit neural representations with periodic activation functions," *arXiv preprint arXiv:2006.09661*, 2020.

[32] B. Stevens and T. Colonius, "Finitenet: A fully convolutional lstm network architecture for time-dependent partial differential equations," *arXiv preprint arXiv:2002.03014*, 2020.

[33] N. Thuerey, K. Weißenow, L. Prantl, and X. Hu, "Deep learning methods for reynolds-averaged navier–stokes simulations of airfoil flows," *AIAA Journal*, vol. 58, no. 1, pp. 15–26, 2020.

[34] M. Kasim *et al.*, "Up to two billion times acceleration of scientific simulations with deep neural architecture search," *arXiv preprint arXiv:2001.08055*, 2020.

[35] B. Kim *et al.*, "Deep fluids: A generative network for parameterized fluid simulations," *Comput. Graph. Forum*, vol. 38, pp. 59–70, 2018.

[36] J. Hsieh, S. Zhao, S. Eismann, L. Mirabella, and S. Ermon, "Learning neural pde solvers with convergence guarantees," *ArXiv*, vol. abs/1906.01200, 2019.

[37] S. Shit *et al.*, "Implicit neural solver for time-dependent linear pdes with convergence guarantee," *arXiv preprint arXiv:1910.03452*, 2019.

[38] A. Sanchez-Gonzalez *et al.*, "Learning to simulate complex physics with graph networks," *arXiv preprint arXiv:2002.09405*, 2020.

[39] A. Odena, V. Dumoulin, and C. Olah, "Deconvolution and checkerboard artifacts," *Distill*, 2016. [Online]. Available: http://distill.pub/2016/deconv-checkerboard

[40] D. Berthelot, T. Schumm, and L. Metz, "Began: Boundary equilibrium generative adversarial networks," *arXiv preprint arXiv:1703.10717*, 2017.

[41] F. Stockhammer *et al.*, "Correlation of f-18-fluoro-ethyl-tyrosin uptake with vascular and cell density in non-contrast-enhancing gliomas," *Journal of neuro-oncology*, vol. 88, no. 2, pp. 205–210, 2008.

[42] M. Hutterer *et al.*, "[18f]-fluoro-ethyl-l-tyrosine pet: a valuable diagnostic tool in neuro-oncology, but not all that glitters is glioma," *Neuro-oncology*, vol. 15, no. 3, pp. 341–351, 2013.

[43] J. Ching and Y.-C. Chen, "Transitional markov chain monte carlo method for bayesian model updating, model class selection, and model averaging," *Journal of engineering mechanics*, vol. 133, no. 7, pp. 816–832, 2007.

[44] P. Hadjidoukas *et al.*, "π4u: A high performance computing framework for bayesian uncertainty quantification of complex models," *Journal of Computational Physics*, vol. 284, pp. 1–21, 2015.

[45] F. Kofler *et al.*, "Brats toolkit: Translating brats brain tumor segmentation algorithms into clinical and scientific practice," *Frontiers in Neuroscience*, vol. 14, 2020.

[46] B. H. Menze *et al.*, "The multimodal brain tumor image segmentation benchmark (brats)," *IEEE transactions on medical imaging*, vol. 34, no. 10, pp. 1993–2024, 2014.

[47] M. Esmaeili, A. L. Stensjøen, E. M. Berntsen, O. Solheim, and I. Reinertsen, "The direction of tumour growth in glioblastoma patients," *Scientific reports*, vol. 8, no. 1, pp. 1–6, 2018.

[48] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.

[49] S. Shit *et al.*, "cldice–a topology-preserving loss function for tubular structure segmentation," *arXiv preprint arXiv:2003.07311*, 2020.

[50] K. Payette *et al.*, "A comparison of automatic multi-tissue segmentation methods of the human fetal brain using the feta dataset," *arXiv preprint arXiv:2010.15526*, 2020.

[51] X. Li, H. Chen, X. Qi, Q. Dou, C.-W. Fu, and P.-A. Heng, "H-denseunet: hybrid densely connected unet for liver and tumor segmentation from ct volumes," *IEEE transactions on medical imaging*, vol. 37, no. 12, pp. 2663–2674, 2018.

Chapter 6: Neural parameters estimation for brain tumor growth modeling

Ivan Ezhov, Jana Lipkova, Suprosanna Shit, Florian Kofler, Nore Collomb, Benjamin Lemasson, Emmanuel Barbier, Bjoern Menze

Conference: International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 787–795, 2019.

To avoid violation of copyrights, the final manuscript submitted to Springer for typesetting is reproduced below.

Synopsis: Understanding the dynamics of brain tumor progression is essential for optimal treatment planning. Cast in a mathematical formulation, it is typically viewed as an evaluation of a system of partial differential equations, wherein the physiological processes that govern the growth of the tumor are considered. To personalize the model, i.e. find a relevant set of parameters, with respect to the tumor dynamics of a particular patient, the model is informed from empirical data, e.g., medical images obtained from diagnostic modalities, such as magnetic-resonance imaging. Existing model-observation coupling schemes require a large number of forward integrations of the biophysical model and rely on simplifying assumptions on the functional form, linking the output of the model with the image information. In this work, we propose a learning-based technique for the estimation of tumor growth model parameters from medical scans. The technique allows for explicit evaluation of the posterior distribution of the parameters by sequentially training a mixture-density network, relaxing the constraint on the functional form, and reducing the number of samples necessary to propagate through the forward model for the estimation. We test the method on synthetic and real scans of rats injected with brain tumors to calibrate the model and predict tumor progression. .

6. CHAPTER 6: NEURAL PARAMETERS ESTIMATION FOR BRAIN TUMOR GROWTH MODELING

Contributions of thesis author: data processing, designing and implementing the algorithmic solution, manuscript writing.

Neural parameters estimation for brain tumor growth modeling

Ivan Ezhov*[1], Jana Lipkova*[1], Suprosanna Shit[1], Florian Kofler[1], Nore Collomb[2], Benjamin Lemasson[2], Emmanuel Barbier[2], and Bjoern Menze[1]

[1] Department of Informatics, Technical University of Munich
[2] Grenoble Institut des Neurosciences, University Grenoble Alpes
`ivan.ezhov@tum.de`

Abstract. Understanding the dynamics of brain tumor progression is essential for optimal treatment planning. Cast in a mathematical formulation, it is typically viewed as evaluation of a system of partial differential equations, wherein the physiological processes that govern the growth of the tumor are considered. To personalize the model, i.e. find a relevant set of parameters, with respect to the tumor dynamics of a particular patient, the model is informed from empirical data, e.g., medical images obtained from diagnostic modalities, such as magnetic-resonance imaging. Existing model-observation coupling schemes require a large number of forward integrations of the biophysical model and rely on simplifying assumption on the functional form, linking output of the model with the image information. In this work, we propose a learning-based technique for the estimation of tumor growth model parameters from medical scans. The technique allows for explicit evaluation of the posterior distribution of the parameters by sequentially training a mixture-density network, relaxing the constraint on the functional form and reducing the number of samples necessary to propagate through the forward model for the estimation. We test the method on synthetic and real scans of rats injected with brain tumors to calibrate the model and to predict tumor progression.

1 Introduction

Modeling brain tumor progression holds a promise of optimizing clinical treatment planning. An appropriate tumor model, personalised with respect to the patient-specific growth dynamics, could quantify clinically relevant information - the tumor's morphology and its character of evolution [1, 2]. Existing mathematical description of the pathophysiological system spans from the intracellular level of gene expression to the macroscopic level of bio-mechanical tumor-tissue interaction. The latter is the scale at which the medical imaging analysis is typically carried out as this is the scale where medical scans are most interpretative.

Among the family of macroscopic models, the reaction-diffusion class of equations [3] is most widely adopted to characterize information visible on medical scans. Under such equations the evolution of tumor cell density is tracked by considering tumor-relevant physiological processes, such as proliferation of cancerous cells, i.e. increase of the cells number due to its division, and the cells' migration into surrounding tissue. Various approaches have been developed to link the output of the model, the distribution of the cell density, with the tumor visible on images [1, 2, 4–9]. Methods as in [5] make a certain assumption on the cell density along visible tumor outlines and fit the model output to image observation that includes lesion growth and tissue displacement. The model adjustment, realized by means of a PDE-constrained optimisation scheme, allows to obtain a point estimate of free model parameters. Bayesian methods [1, 7–9] cast the problem in a probabilistic formulation and provide estimation of the parameters along with confidence intervals via Markov Chain Monte Carlo (MCMC) sampling. The authors of [7] rely on the travelling wave formulation of [4] together with a Bayesian parameter estimation. In [1, 9], authors construct a probabilistic graphical model wherein the probability of imaging signal is defined to be dependent on the biophysical model's output. For magnetic-resonance images (MRI), the probability of observing abnormality is defined as a logistic sigmoid function of the tumor cell density. Phenomenological introduction of the functional form leaves the question whether it possesses a capacity to approximate the mapping between cell density and the imaging information. Also, generating samples from the posterior distribution as with the MCMC methods requires large number of evaluations of the forward model, which can be of the order 10-100 thousand evaluations [1, 9]. This results in an expensive computational cost, impeding clinical validation of more complex models and eventually the approach's adoptability to a routine daily use within clinical settings.

---

* The authors contributed equally to the work.

In this paper, we adopt methodological advances in the estimation of forward model parameters, relying on learning-based strategy [10, 11]. The technique allows for explicit evaluation of the distribution over the parameters by training a mixture-density network (MDN) [12]. The MDN, modeled as a feedforward fully-connected network, maps the output of the model to parameters of the distribution in a non-linear fashion. As theoretical works [13] prove, such a network can serve as a universal function approximation, thus relaxing the necessity of introducing an explicit form for the likelihood, relating the model output and imaging data. In summary, the contributions of this paper are threefold: (1) We make the technique applicable to PDE-based tumor growth models, (2) We validate our method on synthetic and real data of rats implanted with cancer cell lines, using two time points for the model initialization and calibration, (3) We demonstrate that the technique provides more accurate parametric estimations and requires less forward model's samples as compared to explicit Bayesian formulation even with highly efficient MCMC sampling method.

2 Method

Tumor growth model. We base our forward model on the reaction-diffusion equation, describing the tumor progression via spatial and temporal evolution of the cancerous cell density. Particularly, a special type of the reaction-diffusion formalism, the Fisher-Kolmogorov equation, is used:

$$\frac{\partial u}{\partial t} = \nabla(\mathbf{D}\nabla u) + \rho u(1-u), \quad in \quad \Omega \tag{1}$$

$$\mathbf{n} \cdot \nabla u = 0, \quad in \quad \Gamma_{\Omega}. \tag{2}$$

Eq. (1) considers two pathophysiological processes: the logistic proliferation of the cells and its diffusion into neighbouring tissue. u denotes the tumor cell density in the volume of the brain Ω, $\mathbf{D}$ is the diffusion tensor and ρ denotes tumor proliferation rate. The diffusion is assumed to be heterogeneous: with different degree of infiltration in the white and the grey matters, and restricted in the ventricles area. We impose no-flux boundary condition Eq. (2), $\mathbf{n}$ denotes the unit vector orthogonal to the boundary of the simulation domain Γ_{Ω} and ∇ is the gradient operator. We performed experiments with two variants of the model initialization: as a seed point at a fixed location $\mathbf{r}^*$ ($u(\boldsymbol{r}, 0) = u_0$ if $\boldsymbol{r} = \boldsymbol{r*}$, $u(\boldsymbol{r}, 0) = 0$ elsewhere), and as an approximation of the cell density distribution, obtained from an image observation at the first monitoring time point ($u(\boldsymbol{r}, 0) = u_0(\boldsymbol{r})$).

Linking tumor model and image observation. We calibrate the model parameters from image observations in the form of 3D binary tumor segmentations, obtained from MRI modalities. Two MRI modalities, T1-gadolinium hyper-intensities (featuring active tumor core) and T2-hyper-intensities (featuring whole tumor), were used in order to better describe the right tumor morphology. To make the output of the model consistent with the segmentations we make a physiologically plausible assumption that regions of abnormalities visible on the images correspond to regions of high cell infiltration. Respecting the assumption, we introduce two additional parameters u^{T1}, u^{T2} for thresholding the simulated cell density profile, leading to isolines of the tumor cell density that we assume to match outlines of the tumor visible in a given modality. The thresholded binary volumes are combined by element-wise summation to form a 3D label map.

Neural parameters inference. We can view the forward model's output X – the 3D label map – as a sample from a likelihood distribution $p(X|\theta)$ conditioned on a set of parameters $\theta = \{D, \rho, u^{T1}, u^{T2}\}$. The distribution $p(X|\theta)$ cannot be in general evaluated, but its samples are readily available from the tumor model. Given an observation X_{obs} – segmentations of the tumor in the MRI modalities (summed element-wise), our goal is to infer the posterior distribution of the tumor model parameters, using the Bayes rule: $p(\theta|X_{obs}) \propto p(X_{obs}|\theta)p(\theta)$.

In [1, 9], the likelihood is approximated by Bernoulli distribution with the parameter of the distribution defined as logistic sigmoid function. In our work, for inference of the forward model's parameters, we adopt a methodology that allows to learn a nonlinear mapping from the output of the model directly to posterior distribution over its parameters [10]. The inference is based on the neural posterior estimation (NPE), wherein an approximated posterior $q_\phi(\theta|X)$, modeled as a mixture density network, converges to the true posterior $p(\theta|X)$ (via the Kullback-Leibler divergence minimization) by iteratively performing the following steps illustrated in Fig. 1:

1) [*Blue box*] Pairs $\{\theta_i, X_i\}_{i=1}^N$ are generated to form a training dataset. First, the tumor simulator parameters $\{D_i, \rho_i\}$ are sampled from a prior $p(\theta)$ distribution (which is uniform at the first iteration step $s = 1$) and corresponding simulation is propagated until the fixed time point t^* to obtain the 3D cell density profile u_i. Then, u_i is transformed to obtain binary segmentation masks, using the other two sampled parameters $\{u_i^{T1}, u_i^{T2}\}$. Together, the segmentation masks form X_i.

2) [*Yellow box*] The MDN is trained by taking X_i as input and outputting parameters $\alpha_k^s, \boldsymbol{\mu}_k^s, \boldsymbol{\Sigma}_k^s$ of a mixture of Gaussians $q_{\phi^s}(\theta|X) = \sum_k \alpha_k^s N(\theta|\boldsymbol{\mu}_k^s, \boldsymbol{\Sigma}_k^s)$ of K components. The objective of the approximated posterior $q_{\phi^s}(\theta|X)$ training is to maximize the total log-loss, $L(\phi^s) = \sum_i log(q_{\phi^s}(\theta_i|X_i))$.

3) [*Orange box*] The trained MDN is used to infer observation specific parameters of the Gaussian mixture $\alpha_{obs}^s, \boldsymbol{\mu}_{obs}^s, \boldsymbol{\Sigma}_{obs}^s$ by evaluating $q_{\phi^s}(\theta|X = X_{obs})$ at the observation X_{obs} - the label map, obtained from MRI segmentations at t^*.

4) [*Red box*] Finally, the observation specific parameters are used to update estimation of the posterior $p^s(\theta|X_{obs}) = q_{\phi^s}(\theta|X = X_{obs}; \alpha_{obs}^s, \boldsymbol{\mu}_{obs}^s, \boldsymbol{\Sigma}_{obs}^s)$, which then used as a proposal distribution for sampling model parameters during the next iteration.

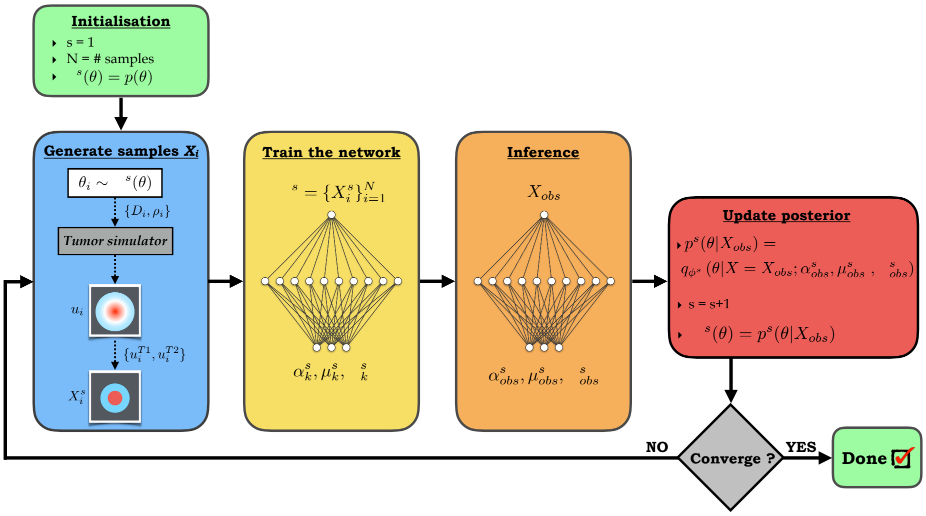
Initialisation
s = 1
N = # samples
$^s(\theta) = p(\theta)$
Generate samples X_i
$\theta_i \sim \; ^s(\theta)$
$\{D_i, \rho_i\}$
Tumor simulator
u_i
$\{u_i^{T1}, u_i^{T2}\}$
X_i^s
Train the network
$^s = \{X_i^s\}_{i=1}^N$
$\alpha_k^s, \mu_k^s, \; _k^s$
Inference
X_{obs}
$\alpha_{obs}^s, \mu_{obs}^s, \; _{obs}^s$
Update posterior
$p^s(\theta|X_{obs}) =$
$q_{\phi^s}(\theta|X = X_{obs}; \alpha_{obs}^s, \mu_{obs}^s, \; _{obs}^s)$
s = s+1
$^s(\theta) = p^s(\theta|X_{obs})$
NO
Converge ?
YES
Done ☑

Fig. 1. The neural posterior estimator. At the heart of it is a mixture density network that maps input data to closed form estimates of the model parameters. It guides the patient-specific simulation of the tumor growth in an efficient iterative fashion.

During successive iterations all the four steps are identical except the step 2, that requires modification of the training objective. To compensate for the fact that we sample from the proposal distribution, the objective function is weighted by a ratio between the prior and proposal distributions $p(\theta_i)/p^s(\theta_i|X_{obs})$ [10]:

$$L(\phi^s) = \sum_i \frac{p(\theta_i)}{p^s(\theta_i|X_{obs})} log(q_{\phi^s}(\theta_i|X_i)) \tag{3}$$

Implementation. We implement the tumor simulator using 3D extension of the multi-resolution adaptive grid solver [14], allowing for high-parallelization. The typical execution time is 20-40 seconds with 8 CPU cores. The architecture of the neural estimator represents a feedforward fully-connected network with a single hidden layer of 100 units with *tanh* as an activation function. We initialize the weights of the network with He-normal [15] at the first iteration and use the weights trained at the iteration step s for initialization at the s+1 step. The latter allows to implicitly reuse the samples from previous iterations: for memory efficiency, it is desirable to avoid having to store and reuse directly old samples since we are dealing with 3D volumes. The network was trained using the Adam optimizer [16] for 100 epochs at each iteration. For each subject *a separate network is employed.* We run the experiments on NVIDIA Quadro P6000 GPU.

3 Experiments

Data. In our experiments, we use synthetic and real data of human glioma cells injected in the rat brain. For producing the synthetic data, we simulate a 3D tumor in the anatomy, obtained from rats brain atlas. We initialize the tumor as a seed point at a fixed location with the diffusion coefficient in the white matter $D_w = 0.02$ $[mm^2/day]$ greater than in the grey matter $D_w = 10D_g$, and proliferation rate $\rho = 0.6$ $[1/day]$. To generate tumor segmentations masks, we threshold the simulated normalized cell density profile at $u^{T1} = 0.7$ and $u^{T2} = 0.25$ for T1 and T2 modalities, respectively, at a single calibration time point ($t^* =$ day 9). The day 11 is used for the model validation. The real data were obtained by injecting F98 tumor cell lines in rats brain. The tumor progression was monitored at several time points from day 9 to day 16, using T1w, T2w, and DWI imaging modalities. The images were expert-annotated. Since the initial condition of tumor location and shape is unknown in the real rats due to the injection, we made use of the DWI modality at the first monitoring time point (day 9) for model initialization. The apparent diffusion coefficient (ADC), calculated from the DWI, can be considered to be inversely proportional to the tumor cell density [17]. We used the ADC, confined within the T2w segmentation volume, as initial condition (in the late time states of tumor progression, the complex tumor microenviroment, hypoxia and necrosis complicate the simple inversely proportional relation). The binary segmentations from the T1w and T2w at the next time point ($t^* =$ day 11) are used for inference of the model parameters, and at the following days (14 and 16) we validate the model predictions, Fig. 3.

Results on synthetic data. For a sensitivity analysis of the inference, experiments on the synthetic data were first performed. In Fig. 2 we show a pairwise correlation of the forward model parameters $\{D, \rho, u^{T1}, u^{T2}\}$ obtained with the neural posterior estimator and the explicit Bayesian inference with MCMC sampling from [1]. For both methods, 1000 samples were used for the inference. Depicted by red stars and orange vertical lines are ground truth (gt) data. The proposed method provides the maximum a posteriori estimation (MAP)

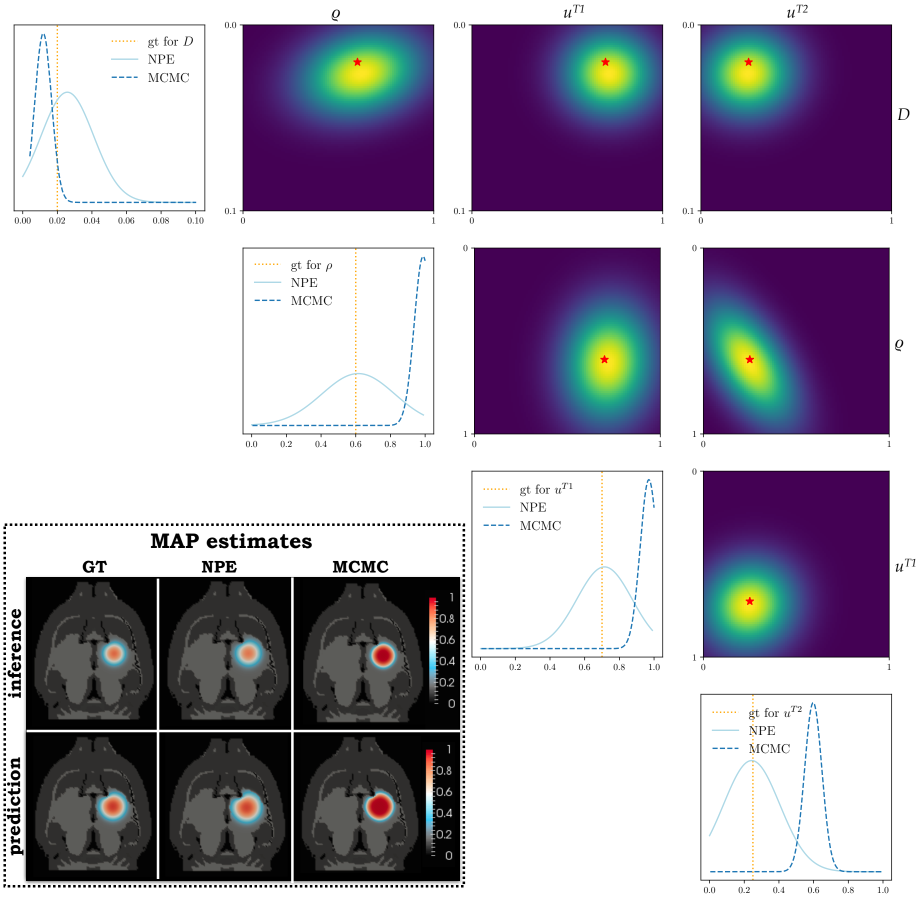

Fig. 2. Posterior distribution of the tumor growth model's parameters inferred for the synthetic rats data: 1D distributions along the diagonal (for the NPE and MCMC methods) and 2D marginals (for the NPE) elsewhere. Depicted by red stars and orange vertical lines are ground truth data. Tumors simulated in the rats brain atlas using the ground truth parameters, and MAP parametric estimates obtained by the NPE and MCMC-based methods are shown on the inset. The top row depicts 2D slices of the cell density profile at the inference time point (day 9), and the bottom row - at the prediction time point (day 11).

for all the parameters in a close agreement with the gt data. In consistency with [1], for the MCMC, based on the likelihood formulation as a logistic sigmoid, the information in the form of binary segmentations is not sufficient to recover the gt parameters. At the same time, the NPE is more computationally efficient, since we observe accurate estimates after running 4 iterations of the posterior update, whereas the MCMC-based method requires about 20 sampling generations for convergence. This is attributed to a) sampling from a range of the parametric space more relevant to the observation after each iteration, b) efficient use of the samples as the technique does not imply any rejection thereof. The inset on the Fig. 2 shows the tumor cell density computed with the MAP parametric estimations from each method, in comparison with the ground truth data.

Results on real data. We validated the method on two real rat cases. Fig. 3 shows the tumor cell distributions for one of the rats,

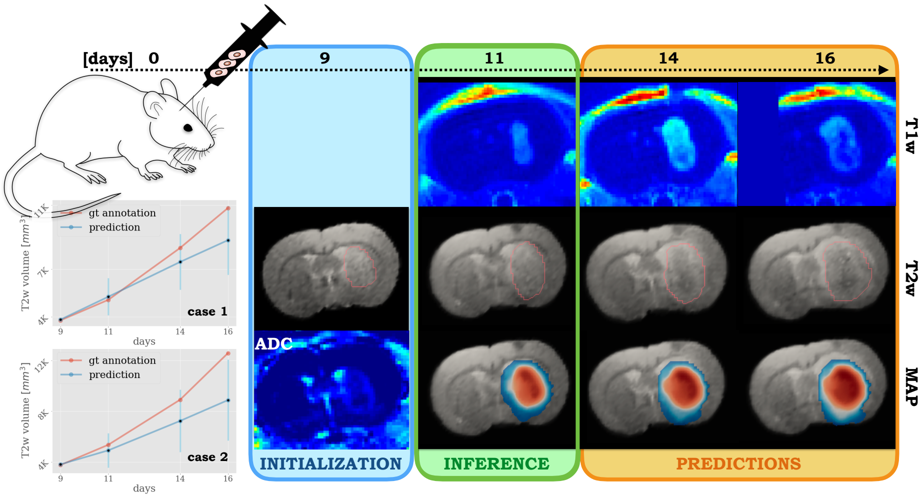

Fig. 3. Data preparation and results of the inference for the real rat cases. After the injection of cancer lines, the rats are monitored at four time points by means of the T1w, T2w and DWI modalities. The inverse ADC from the DWI at day 9 serves as an initial condition for the tumor model. The model parameters are inferred from the binary segmentations, obtained from the T1w and T2w scans at day 11. The model predictions, simulated using MAP estimations of the D, ρ parameters, are validated at days 14 and 16. In the bottom row, the predicted tumor cell density profiles, overlayed on the T2w, are shown. In the middle row, the pink outlines are boundaries of the T2w segmentations. The inset shows volume dynamics of the T2w binary segmentations, obtained from the annotation and predicted by the model, for two rats (the rat case 1 corresponds to the scans shown above). The inferred parametric uncertainties were propagated through the model to obtain mean and standard deviation (blue bars) of the dynamics.

simulated using the inferred MAP estimates, at the calibration time point (day 11) and at the validation points (days 14 and 16).The volume dynamics of T2 tumor segmentations for both rats, calculated from the annotation and predicted by the model, is shown on the inset. The predicted segmentations were obtained by thresholding the simulated cell density profile at the inferred u^{T2} level. While the volumes are in good agreement at the calibration time (an indication of plausible parametric inference), the model prediction underestimates the real tumor dynamics at the validation points. This can be attributed to simplifying assumptions of the Fisher-Kolmogorov model, such as constant proliferation and diffusion for all time points, which limit model's ability to describe the nonlinear character of the real tumor progression. As the proposed inference scheme opens an avenue for efficient parametric estimation, we will test more complicated tumor growth formalisms, e.g., accounting for tissue displacement and microenviromental influence, in future work.

4 Conclusion

We present an approach for inferring parameters of a tumor model from information available on medical scans relying on a learning-based strategy. The approach allows for more efficient parametric estimation, as compared to the conventional Bayes description with MCMC type of sampling, and exempts from the necessity to introduce an explicit form, linking the biophysical model with image observation. Despite we demonstrate the applicability of the method to tumor modeling, the method can be adopted to other physical modeling problems that require calibration from imaging modalities.

Acknowledgment: Ivan Ezhov and Suprosanna Shit have received funding from the European Unions Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement TRABIT No 765148.

References

1. Lipkova, J., et al.: Personalized radiotherapy design for glioblastoma: Integrating mathematical tumor models, multimodal scans and bayesian inference. IEEE Transactions on Medical Imaging (2019) 1–1
2. Unkelbach, J., et al.: Radiotherapy planning for glioblastoma based on a tumor growth model: improving target volume delineation. Physics in Medicine and Biology **59**(3) (2014) 747–770
3. Chaplain, M., Stuart, A.M.: A Mathematical Model for the Diffusion of Tumour Angiogenesis Factor into the Surrounding Host Tissue. Mathematical Medicine and Biology: A Journal of the IMA **8**(3) (1991) 191–220
4. Konukoglu, E., et al.: Image guided personalization of reaction-diffusion type tumor growth models using modified anisotropic eikonal equations. IEEE Transactions on Medical Imaging **29**(1) (2010) 77–95
5. Hogea, C., et al.: An image-driven parameter estimation problem for a reaction–diffusion glioma growth model with mass effects. J Math Biol **56**(6) (2008) 793–825

6. Rockne, R., et al.: A patient-specific computational model of hypoxia-modulated radiation resistance in glioblastoma using 18f-fmiso-pet. JRS, Interface **12** (2015)
7. Le, M., et al.: Bayesian personalization of brain tumor growth model. In: MICCAI. (2015) 424–432
8. Hawkins-Daarud, A., et al.: Quantifying uncertainty and robustness in a biomathematical modelbased patient-specific response metric for glioblastoma. JCO Clinical Cancer Informatics (3) (2019) 1–8
9. Menze, B., et al.: A generative approach for image-based modeling of tumor growth. In: IPMI. (2011) 735–747
10. Lueckmann, J.M., et al.: Flexible statistical inference for mechanistic models of neural dynamics. In: NeurIPS. (2017) 1289–1299
11. Papamakarios, G., Murray, I.: Fast ϵ-free inference of simulation models with bayesian conditional density estimation. In: NeurIPS. (2016) 1028–1036
12. Bishop, C.M.: Mixture density networks. Technical report (1994)
13. Hornik, K.: Approximation capabilities of multilayer feedforward networks. Neural Networks **4**(2) (1991) 251 – 257
14. Rossinelli, D., et al.: MRAG-i2d: Multi-resolution adapted grids for remeshed vortex methods on multicore architectures. Journal of Computational Physics **288** (may 2015) 1–18
15. He, K., et al.: Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In: ICCV. (2015) 1026–1034
16. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. In: ICLR 2015. (2015)
17. Surov, A., et al.: Correlation between apparent diffusion coefficient and cellularity is different in several tumors: a meta-analysis. Oncotarget **8 35** (2017) 59492–59499

Chapter 7: Learn-Morph-Infer: A new way of solving the inverse problem for brain tumor modeling

Ivan Ezhov, Kevin Scibilia, Katharina Franitza, Felix Steinbauer, Suprosanna Shit, Lucas Zimmer, Jana Lipkova, Florian Kofler, Johannes C. Paetzold, Luca Canalini, Diana Waldmannstetter, Martin J. Menten, Marie Metz, Benedikt Wiestler, Bjoern Menze

Journal: Medical Image Analysis, Volume 83, January 2023, 102672.

To avoid violation of copyrights, the final manuscript submitted to Elsevier for typesetting is reproduced below.

Synopsis: Current treatment planning of patients diagnosed with a brain tumor, such as glioma, could significantly benefit by accessing the spatial distribution of tumor cell concentration. Existing diagnostic modalities, e.g. magnetic resonance imaging (MRI), contrast sufficiently well areas of high cell density. In gliomas, however, they do not portray areas of low cell concentration, which can often serve as a source for the secondary appearance of the tumor after treatment. To estimate tumor cell densities beyond the visible boundaries of the lesion, numerical simulations of tumor growth could complement imaging information by providing estimates of full spatial distributions of tumor cells. Over recent years a corpus of literature on medical image-based tumor modeling was published. It includes different mathematical formalisms describing the forward tumor growth model. Alongside, various parametric inference schemes were developed to perform an efficient tumor model personalization, i.e. solving the inverse problem. However, the unifying drawback of all existing approaches is the time complexity of the model personalization which prohibits a potential integration of the modeling into clinical settings. In this work, we introduce a deep learning based methodology for inferring the patient-specific spatial distribution of brain tumors from T1Gd and FLAIR MRI medical scans. Coined as Learn-Morph-

Infer, the method achieves real-time performance in the order of minutes on widely available hardware and the compute time is stable across tumor models of different complexity, such as reaction-diffusion and reaction– advection–diffusion models. We believe the proposed inverse solution approach not only bridges the way for clinical translation of brain tumor personalization but can also be adopted to other scientific and engineering domains..

Contributions of thesis author: data processing, designing and implementing the algorithmic solution, manuscript writing.

Learn-Morph-Infer: a new way of solving the inverse problem for brain tumor modeling

Ivan Ezhov[a,b,1], Kevin Scibilia[a], Katharina Franitza[c], Felix Steinbauer[a], Suprosanna Shit[a,b], Lucas Zimmer[d,b], Jana Lipkova[e,f,g], Florian Kofler[a,b,i], Johannes Paetzold[a,b], Luca Canalini[h], Diana Waldmannstetter[a], Martin J. Menten[a,b], Marie Metz[i,b], Benedikt Wiestler[b,i,2], Bjoern Menze[d,2]

[a]*Department of Informatics, TUM, Munich, Germany*
[b]*TranslaTUM - Central Institute for Translational Cancer Research, TUM, Munich, Germany*
[c]*Department of Mechanical Engineering, TUM, Munich, Germany*
[d]*Department of Quantitative Biomedicine, UZH, Zurich, Switzerland*
[e]*Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA*
[f]*Broad Institute of Harvard and MIT, Cambridge, MA*
[g]*Data Science Program, Dana-Farber Cancer Institute, Boston, MA*
[h]*Fraunhofer MEVIS, Bremen, Germany*
[i]*Neuroradiology Department of Klinikum Rechts der Isar, TUM, Munich, Germany*

Abstract

Current treatment planning of patients diagnosed with a brain tumor, such as glioma, could significantly benefit by accessing the spatial distribution of tumor cell concentration. Existing diagnostic modalities, e.g. magnetic resonance imaging (MRI), contrast sufficiently well areas of high cell density. In gliomas, however, they do not portray areas of low cell concentration, which can often serve as a source for the secondary appearance of the tumor after treatment. To estimate tumor cell densities beyond the visible boundaries of the lesion, numerical simulations of tumor growth could complement imaging information by providing estimates of full spatial distributions of tumor cells. Over recent years a corpus of literature on medical image-based tumor modeling was published. It includes different mathematical formalisms describing the forward tumor growth model. Alongside, various parametric inference schemes were de-

---

Email address: `ivan.ezhov@tum.de` (Ivan Ezhov)

[1] Corresponding author.

[2] Contributed equally as senior authors.

Preprint submitted to Journal of LaTeX Templates *January 11, 2023*

veloped to perform an efficient tumor model personalization, i.e. solving the inverse problem. However, the unifying drawback of all existing approaches is the time complexity of the model personalization which prohibits a potential integration of the modeling into clinical settings. In this work, we introduce a deep learning based methodology for inferring the patient-specific spatial distribution of brain tumors from T1Gd and FLAIR MRI medical scans. Coined as *Learn-Morph-Infer*, the method achieves real-time performance in the order of minutes on widely available hardware and the compute time is stable across tumor models of different complexity, such as reaction-diffusion and reaction-advection-diffusion models. We believe the proposed inverse solution approach not only bridges the way for clinical translation of brain tumor personalization but can also be adopted to other scientific and engineering domains.

Keywords: Inverse modeling, physics-based deep learning, glioma, model calibration, tumor modeling, MRI

---

1. Introduction

Glioblastoma (GBM) is the most aggressive brain tumor, characterized by varying and unknown infiltration into the surrounding tissue. After resection of the tumor mass visible in MRI scans, current treatment includes radiotherapy targeting tissue around the visible lesion to account for residual tumor cells. Tumor recurrence is however present in most cases, possibly due to patient-specific and non-uniform distribution of residual tumor cells. Personalization of the clinical (irradiation) target volume could spare more healthy tissue and increase progression-free survival by potentially avoiding recurrence (Stupp et al., 2014, Harpold et al., 2007, Jackson et al., 2015, Lipkova et al., 2019).

Current computational approaches for personalizing radiotherapy planning often rely on solving an inverse problem for GBM growth models (Hogea et al., 2008, Konukoglu et al., 2010b, Geremia et al., 2012, Menze et al., 2011, Le et al., 2017, Lipkova et al., 2019, Scheufele et al., 2020, Subramanian et al., 2020a, Hormuth et al., 2021, Lorenzo and et al, 2021). In this context, the growth (forward) models are based on partial differential equations (PDEs) that describe the evolution of tumor cell density in the brain anatomy. The inverse model aims to identify free parameters of the forward model that best match the observation, e.g. tumor outlines from medical imaging modalities. To identify such parameters, the inverse problem can be cast as constrained optimization (Hogea et al., 2008, Mang et al., 2012, Scheufele et al., 2019) or Bayesian inference formulations (Menze et al., 2011, Lipkova et al., 2019, Ezhov et al., 2020, 2019).

Predominantly, existing forward GBM models view tumor progression at the macroscopic level by describing gross biomechanical phenomena. These include diffusive motion and proliferation of tumor cells (under simplistic reaction-diffusion PDEs) (Menze et al., 2011), interaction between the tumor and surrounding tissue (i.e. mass-effect) (Subramanian et al., 2020a), necrotic core formation (Patel and Hathout, 2017), etc. Despite methodological advances in computing the inverse model, the total time for model personalization is still large amounting to multiple hours using such simplistic forward models (Subramanian et al., 2020a, Hormuth et al., 2018, Scheufele et al., 2019). For example, in (Scheufele et al., 2019) the authors exploit a highly efficient quasi-Newton optimization scheme to infer parameters of the reaction-diffusion model. The inference converges after $\sim$5 hours of compute on 11 dual-x86 CPU nodes for 256^3 resolution. In (Subramanian et al., 2020a), the mass-effect model is solved using an analogous optimization scheme but implemented on a GPU leading to the same order of compute time for the 256^3 grid (and up to 1 hour for 128^3 resolution). Bayesian methods providing uncertainty estimate of the parametric inference (Lipkova et al., 2019) can take an even longer time (up to days) of computing on specialized CPU clusters.

Recently, machine learning solutions entered the field of PDEs. Learnable solutions for both forward (Raissi et al., 2019, Sitzmann et al., 2020, Stevens and Colonius, 2020, Thuerey et al., 2020, Kasim et al., 2020, Kim et al., 2018) and inverse (Papamakarios and Murray, 2016, Lueckmann et al., 2017, Dax et al., 2021) models were developed. In (Papamakarios and Murray, 2016, Lueckmann et al., 2017), the authors proposed a Bayesian framework that allows bypassing a compute of intractable likelihoods necessary for the Bayesian inference. This notably speeds up the inference, however, it still requires thousands of simulations to be generated by the forward model, which in turn can still result in many hours of compute for the inverse problem. In concurrent to our work (Dax et al., 2021), a solution to the inverse problem providing point estimates was proposed. This is achieved via learning a mapping from the physical model simulations directly to the model parameters by leveraging access to large simulated data. Unfortunately, this method is not capable to deal with the varying geometry of the simulation domain. Moreover, as our paper shows, direct mapping to physical model parameters can result in a deterioration of prediction accuracy since the inverse mapping between simulations and parameters is not bijective.

Few machine learning approaches (Ezhov et al., 2020, Pati et al., 2020) also appeared in the brain tumor modeling context. However, (Ezhov et al., 2020) requires a vast amount of forward model evaluations for convergence of parametric estimation under Bayesian settings for each new patient. In turn, (Pati et al., 2020) requires access to a dataset of inferred model parameters that can become prohibitively expensive to collect with the growing complexity of the tumor model.

Potential integration of brain tumor modeling into clinical practice would require access to a large cohort of longitudinal clinical data, allowing to estimate the clinical value of a patient outcome's forecast by the tumor models (Yankeelov et al., 2013). Integration of current macroscopic models would also require a thorough analysis of forecast consistency between the macroscopic description and higher complexity microscopic models encompassing subcellular biophysics. For this, it is paramount to bring the computational cost of the inverse modeling to a reasonable time. Here, we propose a neural network based methodology for predicting a patient-specific spatial distribution of GBM (from single time-point medical scans, namely T1Gd and FLAIR MRI) that requires neither sampling nor optimization. As such, it is potentially well suited for rapid model personalization in a clinical workflow, but also for scaling up preceding feasibility studies to large patient cohorts.

Our contribution is a) a pipeline allowing to perform tumor model personalization in a fixed for all patients space, and b) a special learning setup for the network (the main part of the pipeline) performing inverse model inference. The method achieves lightning-fast performance in the order of minutes on widely available hardware and the compute time is stable across tumor models of different complexity. This in turn opens the possibility of rapid testing various biophysical models on a large dataset of patients and hence bridging the way for clinical translation.

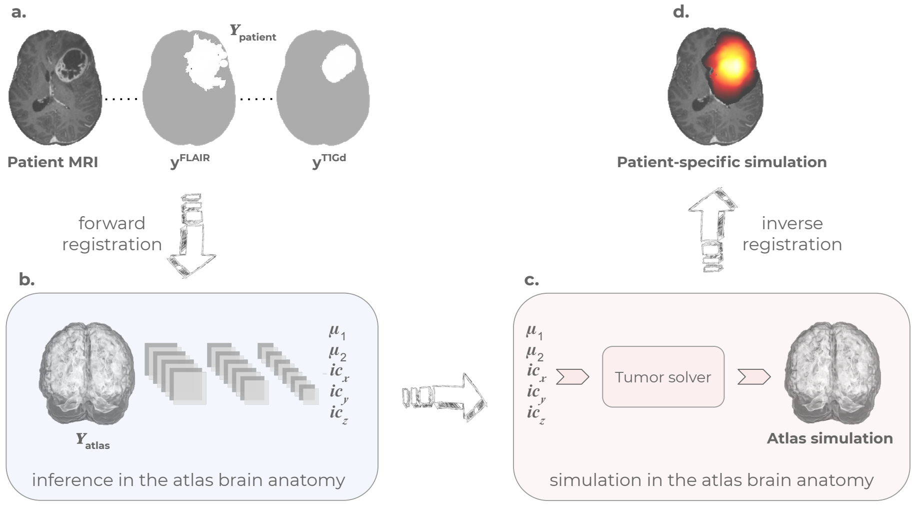
a.
Y_{patient}
Patient MRI
y^{FLAIR}
y^{T1Gd}
d.
Patient-specific simulation
forward registration
inverse registration
b.
c.
μ_1
μ_2
ic_x
ic_y
ic_z
Y_{atlas}
inference in the atlas brain anatomy
Tumor solver
Atlas simulation
simulation in the atlas brain anatomy

Figure 1: A sketch of the inference procedure of the *Learn-Morph-Infer* pipeline. First, the patient's brain tumor segmentations $\{\mathbf{y}^{T1Gd}, \mathbf{y}^{FLAIR}\}$ are morphed to the brain atlas space (a). A network trained on synthetic data inputs the morphed segmentations and outputs corresponding tumor model parameters $\{\mu_1, \mu_2, ic_x, ic_y, ic_z\}$ (b). The inferred parameters are used to simulate a tumor in the atlas space (c). Finally, the simulated tumor is morphed back to the patient space (d).

2. Method

2.0.1. Learn-Morph-Infer Pipeline

In contrast to a few works like (Hormuth et al., 2019, Ezhov et al., 2019, Tunç et al., 2021), we solve the inverse problem by relying only on single time-point MRI observations of the brain tumor $Y = \{\mathbf{y}^{T1Gd}, \mathbf{y}^{FLAIR}\}$. This is the most realistic clinical scenario as normally treatment follows immediately after the first MRI scanning. Given a tumor observation Y, our goal is to calibrate a set of personalized parameters θ_c of the forward tumor growth model that infers the underlying patient-specific tumor cell density $c_{patient}$ in the patient anatomy[3]. For a given patient, this is achieved via the proposed *Learn-Morph-Infer* pipeline:

- We register a patient MRI image to the brain atlas (Rohlfing et al., 2010) and obtain a transformation matrix.
- The transformation matrix is used to morph scans based on the patient anatomy $Y_{patient}$ to scans in the atlas anatomy Y_{atlas} as illustrated in Fig. 1a-b.
- A neural network, that has learned to solve the inverse problem $Y_{atlas} \rightarrow \theta_c$ through prior training on simulated data Y_{sim} (Fig. 1), predicts θ_c during inference, Fig. 1b.

---

[3] Let us explain how we see the potential clinical utility of glioma modeling. In clinical practice, for glioma patients normally only a single MRI scan is taken before treatment. It is not feasible to capture the dynamics of the pathology by calibrating a tumor model from a single scan - multiple sets of model parameters can result in the same simulated tumor profile. Thus, one cannot reliably predict tumor evolution over time as different sets of parameters will result in tumor trajectories diverging over time. Instead, we are only interested in identifying a simulated tumor whose morphology (and, hence, tumor cell distribution) best matches the morphology of the tumor we see on a single MRI scan. The best matching simulated profile will then serve to predict tumor infiltration beyond contrast-enhancing tumor areas and thereby define the target volume for radiotherapy (as opposed to the currently used simple EORTC criteria - 2cm border surrounding the cavity (Young et al., 1999)).

- These parameters θ_c are used as input in the tumor growth model (forward solver) to infer a tumor cell density c_{atlas} in the atlas space, Fig. 1c.
- The tumor volume c_{atlas} is transformed back to the patient space with the inverse transformation matrix, yielding $c_{patient}$ as displayed in Fig. 1d.

2.0.2. Forward tumor model

We independently probe two types of non-linear PDEs describing tumor growth:

Reaction-diffusion equation. First, we consider the Fisher-Kolmogorov PDE describing the evolution of the tumor cell concentration c by considering cell diffusion and proliferation,

$$\frac{\partial c}{\partial t} = \nabla \cdot (\mathbf{D}\nabla c) + \rho c(1-c) \tag{1}$$

$$\nabla c \cdot \mathbf{n} = 0 \quad \textit{boundary condition} \tag{2}$$

Here, ρ denotes the tumor proliferation rate while the infiltrative behaviour of the tumor is modelled by the diffusion tensor $\mathbf{D} = D \cdot \mathbb{I}$. The equation is solved in a three dimensional atlas brain anatomy segmented into white matter (WM), grey matter (GM) and cerebrospinal fluid (CSF). The diffusion coefficient D is computed for each voxel $\mathbf{i}$ with location (i_x, i_y, i_z) as $D_i = p_{w_i} D_w + p_{g_i} D_g$, where p_{w_i}, p_{g_i} describe percentages at voxel $\mathbf{i}$ and D_w, D_g diffusion coefficients of WM and GM respectively, and a relation $D_w = 10 \cdot D_g$ is assumed (Lipkova et al., 2019). No cell diffusion into CSF is feasible according to the model. The solver based on this growth model takes $\theta_c = \{D_w, \rho, T, ic_x, ic_y, ic_z\}$ as input and returns a tumor cell density c. The parameters $\mathbf{x} = (ic_x, ic_y, ic_z)$ define the initial condition where the tumor is initialized at time t=0 as a point seed. The tumor is simulated until the time of detection T.

Reaction-diffusion-advection equation. The second type is a non-linear reaction-diffusion-advection PDE analogous to (Subramanian et al., 2020a). In the following, the brain tissue will be represented by $\mathbf{m} = (m_{WM}(\mathbf{i}, t),\ m_{GM}(\mathbf{i}, t),\ m_{CSF}(\mathbf{i}, t))$ for each voxel $\mathbf{i}$ and time t. The normalized tumor cell density $c = c(\mathbf{i}, t)$ can be modelled by the following equations:

$$\frac{\partial c}{\partial t} = \nabla \cdot (\mathbf{D} \nabla c) - \nabla(c\mathbf{v}) + \rho c(1 - c) \tag{3}$$

$$\frac{\partial \mathbf{m}}{\partial t} + \nabla \cdot (\mathbf{m} \otimes \mathbf{v}) = 0 \tag{4}$$

$$\nabla \cdot (\lambda \nabla \mathbf{u} + \mu(\nabla \mathbf{u} + \nabla \mathbf{u}^{\top})) = \gamma \nabla c \tag{5}$$

$$\frac{\partial \mathbf{u}}{\partial t} = \mathbf{v} \tag{6}$$

$$\nabla c \cdot \mathbf{n} = 0 \quad \textit{boundary condition} \tag{7}$$

$$\mathbf{m} = 0 \quad \textit{boundary condition} \tag{8}$$

$$\mathbf{u} = 0 \quad \textit{boundary condition} \tag{9}$$

$$\mathbf{v} = 0 \quad \textit{boundary condition} \tag{10}$$

Coupling Eqn. 3 to a linear elasticity model, Eqn. 5, allows considering deformation in the anatomy due to a mass effect induced by tumor growth (Subramanian et al., 2019). The linear elasticity model is defined by the Lamè coefficients λ and μ as specified in Eqn. 5. The displacement $\mathbf{u}$ is represented in the advection term of Eqn. 3. The degree of the mass effect depends on the selection of the mass effect parameter γ.

Parameter symbol	Parameter meaning
c	Tumor cell density
$\mathbf{D}$	Diffusion tensor
D	Diffusion coefficient
$\mathbb{I}$	Identity tensor
ρ	Proliferation rate
$\mathbf{n}$	Unit vector normal to boundary
$\mathbf{u}$	Advection displacement
$\mathbf{v}$	Advection velocity
$\mathbf{m}$	Brain tissue maps
λ, μ	Lamè coefficients
γ	Mass effect parameter

Table 1: Description of parameters driving the reaction-diffusion and reaction-advection-diffusion PDEs.

2.0.3. Linking cell density with MRI signal

MRI modalities capture structural information about the brain tumor. T1Gd contrasts the tumor core, whereas FLAIR informs about the area of the edema in addition to the tumor core. It is established practice (Lê et al., 2016, Lipkova et al., 2019, Subramanian et al., 2020a, Tunc et al., 2021, Konukoglu et al., 2010a, Menze et al., 2011) to consider binary segmentations corresponding to the MRI scans to inform biophysical models. The binary masks contain zeros in the area of healthy tissues and are non-zero in the tumor-related area. In order to relate the segmentations $Y = \{\mathbf{y}^{T1Gd}, \mathbf{y}^{FLAIR}\}$ to a simulated tumor cell density c, we threshold the density at randomly sampled levels c_t^{T1Gd} and c_t^{FLAIR} ($c_t^{T1Gd} > c_t^{FLAIR}$) to obtain $Y_{sim} = \{\mathbf{y}_{sim}^{T1Gd}, \mathbf{y}_{sim}^{FLAIR}\}$ reproducing the real segmentations [4].

---

[4] We want to clarify here the thresholding step in more detail. The output of our tumor solver is a continuous tumor cell density distribution. The experimental observation to which we want to fit our tumor model simulations comes in a form of binary tumor segmentation. In order to identify a simulation best fitting the binary segmentation, we threshold our simulated continuous tumor cell distribution to obtain a binary volume, which in turn can already be nicely compared with the real patient segmentation. Oftentimes, in literature people resort to fixed values of cell density for the threshold levels. However, there is no understanding

2.0.4. Learning the inverse model in atlas space

As discussed in the previous section, the key step of the *Learn-Morph-Infer* pipeline is to learn the inverse tumor model using a neural network that can infer a set of personalized parameters θ_c from corresponding tumor observations Y_{atlas}. In order to create a dataset for the network training, we generate 100,000 tumors in the atlas space by randomly sampling tumor model parameters $\{D_w, \rho, T, ic_x, ic_y, ic_z, c_t^{T1Gd}, c_t^{FLAIR}\}$ within physiologically plausible ranges (Swanson et al., 2000). To form the neural network input, the simulated MRI segmentations are combined into one volume $\mathbf{y}_{sim}^{MRI} = 0.666 \cdot \mathbf{y}_{sim}^{T1Gd} + 0.333 \cdot \mathbf{y}_{sim}^{FLAIR}$.

Reformulation of the inverse problem. Now, the question is what should be used as a network prediction? It is tempting to try to predict the tumor model parameters directly. However, it is well known that the inverse problem is highly ill-posed, i.e. numerous sets of dynamic parameters $\{D_w, \rho, T\}$ correspond to the same simulated cell density profile. Thus, we have two sources of prediction error: a) coming from the fact that we learn a mapping from one to many. Imagine we have two sets of dynamic parameters $\{D_w, \rho, T\}$ and $\{D_w^*, \rho^*, T^*\}$ that result in the same tumor profile $\mathbf{y}_{sim}^{MRI}$. If we train a network in a supervised fashion, every time the network predicts $\{D_w^*, \rho^*, T^*\}$ for a $\{D_w, \rho, T\}$-$\mathbf{y}_{sim}^{MRI}$ pair (and vice-versa), it will be falsely penalized; b) the actual error that comes from limited network capacity to accurately learn the mapping.

(neither agreement) in the community of what levels would best correspond to a real scenario. Different works used different numbers in their studies (Le et al., 2017). To consider the most general case, we randomly sampled the values within the ranges reported in the literature.

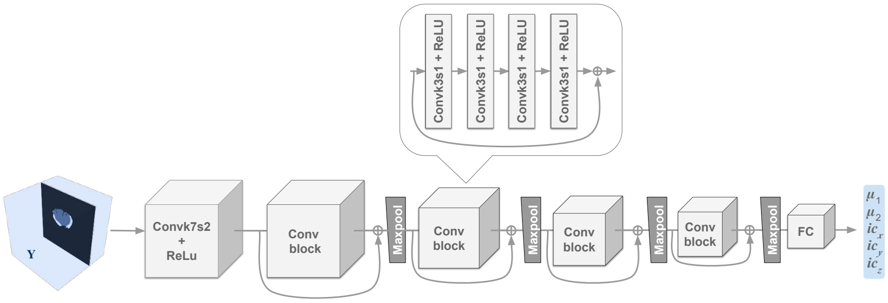

Figure 2: A design of the inverse model network - one of the key elements of the *Learn-Morph-Infer* pipeline. The network design represents a ResNet type architecture. It takes as an input the binary brain tumor segmentations $\{\mathbf{y}^{T1Gd}, \mathbf{y}^{FLAIR}\}$ and outputs $\{\mu_1, \mu_2, ic_x, ic_y, ic_z\}$. Crucially, the network predicts $\{\mu_1, \mu_2\}$ which are not the exact forward model parameters but time-independent combinations thereof.

Clearly, the error of type (a) should negatively affect the learning performance, as the network may be penalized for making a sensible prediction. To circumvent this, we do not predict the parameters directly. As evident from Eqn. 1, normalization of the time parameter T in Eqn. 1 is equivalent to re-scaling of the proliferation ρ and diffusion D coefficients (Subramanian et al., 2020b). This means that for sets $\{D_w, \rho, T\}$ and $\{D_w^*, \rho^*, T^*\}$ corresponding to the same simulated tumor, combinations of time-independent parameters $\mu_1 = \sqrt{D_w T}$, $\mu_2 = \sqrt{T\rho}$ stay constant ($\sqrt{D_w T}$=$\sqrt{D_w^* T^*}$, $\sqrt{T\rho}$=$\sqrt{T^*\rho^*}$) (Konukoglu et al., 2010b). Hence, predicting these combinations of time-independent parameters μ_1 and μ_2 relaxes the error type (a). In order to calculate back the $\{D_w, \rho, T\}$, we introduce a third combination as $v = 2\sqrt{D_w \rho}$. As it is not possible to infer the velocity v from a single time-point observation, we set the velocity equal to the mean velocity of the used sampling range, 200 $mm/year$ (note also that for our purpose the choice of v is irrelevant as any tumor simulation can be obtained with arbitrary v (Menze et al., 2011)). Given $\{\mu_1, \mu_2, v\}$, we can calculate $\{D_w = \frac{\mu_1 v}{2\mu_2}, \rho = \frac{\mu_2 v}{2\mu_1}, T = \frac{2\mu_2\mu_1}{v}\}$.

Following this reasoning, we make the network to predict five parameters: $\{\mu_1, \mu_2\}$ and $\{ic_x, ic_y, ic_z\}$. Note that we do not predict $\{c_t^{T1Gd}, c_t^{FLAIR}\}$, as we do not need the threshold parameters at further steps of the *Learn-Morph-Infer* pipeline.

The inverse network design. The network we chose to learn the inverse model represents a convolutional architecture, depicted in Fig. 2. Every convolution in the network is followed by a rectified linear unit (ReLU) nonlinearity. The input is passed through an initial convolution of kernel size 7, stride 2 and 64 filters (k7s2f64), downsampling the volume to 64^3 and increasing the number of channels to 64. This volume of size 64^3 x 64 is input through four convolutional blocks, where every convolutional block contains four convolutions of kernel size 3, stride 1, and 64 filters (i.e. number of filters is kept constant throughout the network). A convolutional block uses a skip connection to learn a residual mapping (He et al., 2016), with the input being added element-wise to the output of the four convolutions. The first three convolutional blocks are followed by a MaxPool3D layer (with parameters k2s2) to downsample the 3D volumes by two. The last convolutional block is followed by a global average pooling layer, shrinking the 64 3D volumes to 64 neurons that can be linked through a fully connected (FC) layer to the output. These outputs are linearly interpolated into the [-1, 1] range for training.

3. Results

3.1. Data and implementation details

Synthetic data. The simulated tumors used for training the network have resolution of 128^3. The simulations were generated by randomly sampling patient-specific parameters from the following ranges: $D_w \in [0.0002\,\frac{\mathrm{cm}^2}{\mathrm{d}}, 0.015\,\frac{\mathrm{cm}^2}{\mathrm{d}}]$, $\rho \in [0.002\,\frac{1}{\mathrm{d}}, 0.2\,\frac{1}{\mathrm{d}}]$, $T \in [50\mathrm{d}, 1500\mathrm{d}]$, $x \in [0.15, 0.7]$, $y \in [0.2, 0.8]$, $z \in [0.15, 0.7]$, $c_t^{T1Gd} \in [0.5, 0.85]$, $c_t^{FLAIR} \in [0.05, 0.5]$. The elasticity model parameters λ, μ, γ were taken from (Subramanian et al., 2019). Tumors that are unrealistically small or large have been discarded (based on minimum and maximum tumor sizes of real tumors from BraTS dataset (Menze et al., 2014)). The simulations dataset is then divided into a training set (80000 tumors), validation set (8000 tumors) and test set (12000 tumors).

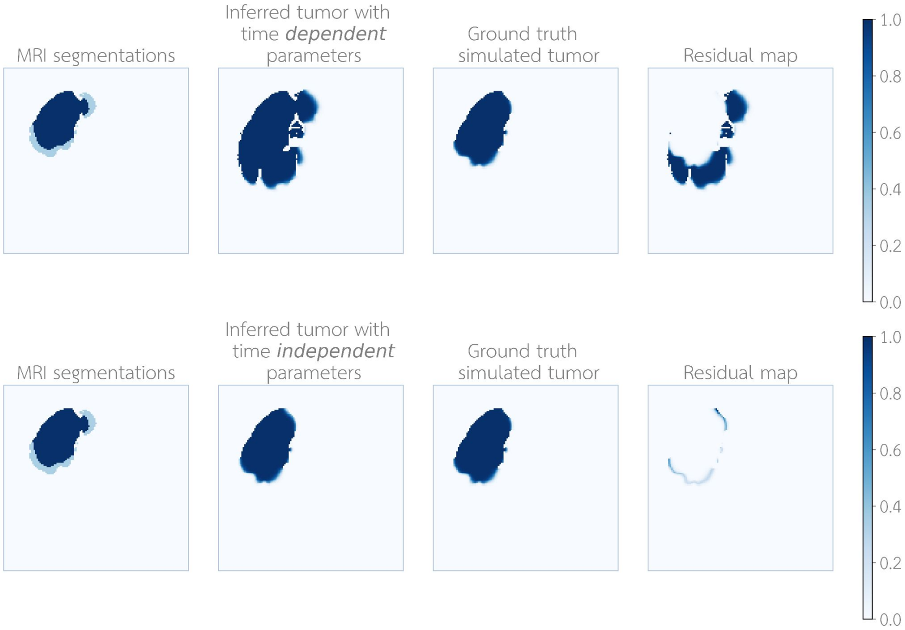
MRI segmentations
Inferred tumor with time *dependent* parameters
Ground truth simulated tumor
Residual map
1.0
0.8
0.6
0.4
0.2
0.0
MRI segmentations
Inferred tumor with time *independent* parameters
Ground truth simulated tumor
Residual map
1.0
0.8
0.6
0.4
0.2
0.0

Figure 3: Qualitative comparison between the tumors inferred using the *Learn-Morph-Infer* pipeline with two different network training strategies: with time-dependent $(D_w, \rho, T, \mathbf{x})$ parameters as network output, and time-independent $(\mu_1, \mu_2, \mathbf{x})$ parameters as output. The examples correspond to the Fisher-Kolmogorov tumor model with MAE equal to 0.495 (within WM) and 0.496 (within GM) for time-dependent inference, and 0.048 (within WM) and 0.048 (within GM) for time-independent inference. The first column ”MRI segmentations” corresponds to the FLAIR+T1Gd segmentation, the second one ”Inferred tumor” corresponds to the simulation inferred by the proposed *Learn-Morph-Infer* pipeline, the third column ”Ground truth simulated tumor” corresponds to the simulated tumor which we used for the sensitivity analysis (see the ”Synthetic data” section for details how these test data were formed), and the fourth ”Residual map” column is the difference between the 2nd and 3rd columns’ images.

Analogously to how we created the training data, we form synthetic test data by thresholding a simulated tumor c_{gt} at two levels. Then we pass the obtained thresholded volumes through the pipeline to infer the tumor cell distribution c_{pred}. Finally, we quantitatively compare the difference between c_{gt} and c_{pred}.

Real data. The data set of real MRI images consists of an in-house cohort of 80 patients with a newly diagnosed glioblastoma, IDH wild type as per the 2016 WHO classification of brain tumors. All patients gave written informed consent to be part of an observational cohort. Preoperative MR images were converted to NIfTI format. For image preprocessing (co-registration, skull stripping) and automated tumor segmentation, we used BraTS Toolkit (Kofler et al., 2020), a tool we developed locally and which is freely available. The average age of the patient cohort is 60 years with a minimum of 26 and a maximum of 79 years. The cohort is equally represented by gender.

Implementation details. For the registration between the patient and atlas brain MRI scans, we use the Advanced Normalization Tools (ANTs) (Avants et al., 2009). We choose a deformable SyN registration that ensures providing both forward and inverse transformation with step-size 0.25, weight 1, and region radius for cross-correlation computation r=4. Cross-correlation is used as a similarity metric. The optimization is performed over two resolutions with a maximum of 50 iterations at the coarsest level, and 20 at the final level. The tumor area on the patient scan was masked for the registration. We use a Gaussian regularizer with a sigma of 3 operating on the similarity gradient. These settings provided high morphing quality at relatively fast compute ($\sim$2 minutes).

The network is initialized with He initialization as in the original ResNet architecture (He et al., 2016), and is trained with the AdamW (Loshchilov and Hutter, 2019) optimizer, which is a variant of the Adam optimizer (Kingma and Ba, 2015) using decoupled weight decay. We use an initial learning rate of 6 x 10^{-5} which is decayed exponentially after every batch by a factor of 0.999997. Weight decay with a factor of 0.05 is used as a regularization technique. Furthermore, we train the network with a batch size of 32 and the Mean Squared Error (MSE) loss function. All training and testing runs were executed on an NVIDIA Quadro RTX 6000 with the PyTorch framework.

3.2. Experiments

We perform two sets of experiments: a) on synthetic data to estimate the accuracy of the learnable inverse model, and b) on patient MRI scans to qualitatively probe the transferability of the method to real data.

Experiment			MAE			DICE		
ID	Tumor Model	Predicted Parameters	GM	WM	CSF	c_t=0.01	c_t=0.1	c_t=0.8
1		$\{D_w, \rho, T\}, \{\mathbf{x}\}$	0.461	0.463	0.0	0.607	0.558	0.423
2	FK	$\{\mu_1, \mu_2\}, \{\mathbf{x}\}$	0.059	0.059	0.0	0.940	0.928	0.855
3		$\{\mu_1, \mu_2, \mathbf{x}\}$	**0.057**	**0.057**	**0.0**	**0.943**	**0.932**	**0.861**
4		$\{D_w, \rho, T\}, \{\mathbf{x}\}$	0.089	0.087	0.059	0.846	0.807	0.737
5	ME	$\{\mu_1, \mu_2\}, \{\mathbf{x}\}$	0.059	0.057	0.054	0.877	0.841	0.772
6		$\{\mu_1, \mu_2, \mathbf{x}\}$	**0.055**	**0.054**	**0.054**	**0.886**	**0.850**	**0.783**

Table 2: Ablation analysis on the test set (12k samples). In total, we perform 6 experiments. First three experiments are performed for the Fisher-Kolmogorov (FK) tumor model: 1) Two separate neural networks predicting $\{D_w, \rho, T\}$ and $\{\mathbf{x} = (ic_x, ic_y, ic_z)\}$, 2) Two separate neural networks for prediction of the growth $\{\mu_1, \mu_2\}$ and location $\{\mathbf{x}\}$ parameters, 3) Single neural network predicting $\{\mu_1, \mu_2, \mathbf{x}\}$. The last experiments 4-6 are analogous but performed for the mass-effect (ME) model. As error measure for all experiments, we use the mean absolute error (MAE) $||c_{sim} - c_{gt}||$ in WM, GM, and CSF, as well as DICE score between ground truth and simulated tumor cell density thresholded at different c_t. The non-zero error in the CSF area for the ME model comes from the fact that the simulated tumor is allowed to displace the healthy tissue including the CSF. For both FK and ME models, the usage of a single network predicting time-independent parameters results in a notable increase in accuracy compared to the time-dependent counterpart.

3.3. Synthetic test

In Tab. 2, for two different tumor models, we show results of the ablation analysis, wherein we perform multiple experiments varying the neural network input and output configurations. First, we provide empirical proof that a network predicting $\{D_w, \rho, T\}$ instead of time-independent parameter combinations

$\{\mu_1, \mu_2\}$ cannot be trained reliably. Mean absolute error, as well as DICE score, improve significantly when the network predicts $\{\mu_1, \mu_2\}$ (for the ME model the improvement is less pronounced that we attribute to a higher numerical error in our forward solver implementation for this model). Second, we tested whether the performance can be improved by learning two separate networks predicting growth $\{\mu_1, \mu_2\}$ and initial location parameters $\{\mathbf{x} = (ic_x, ic_y, ic_z)\}$, respectively. This test did not reveal an improvement compared to a single network predicting all parameters.

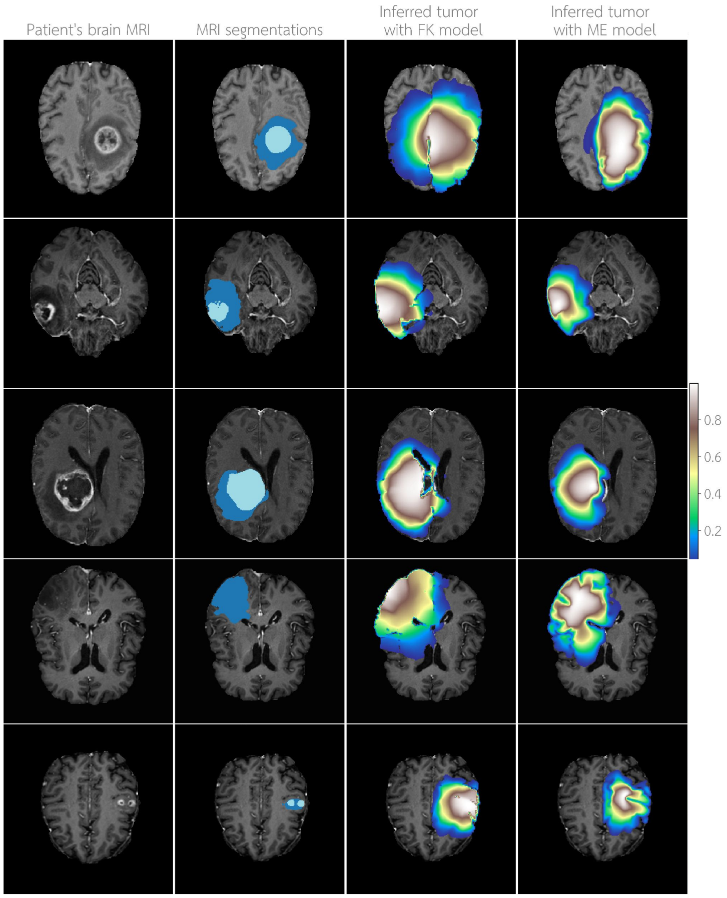
Patient's brain MRI
MRI segmentations
Inferred tumor with FK model
Inferred tumor with ME model
0.8
0.6
0.4
0.2

Figure 4: Examples of patient-specific simulations produced using the *Learn-Morph-Infer* method. The inverse model network was trained on samples from the Fisher-Kolmogorov (3rd column) and mass effect (4th column) forward models.

Fig. 3 qualitatively showcases the accuracy of inference using the *Learn-Morph-Infer* pipeline. As discussed before, depending on the network training strategy (either predicting time-dependent $\{D_w, \rho, T, \mathbf{x}\}$ parameters, or time-independent $\{\mu_1, \mu_2, \mathbf{x}\}$ parameters as network output), the accuracy of the final simulated tumor notably differs. If the proposed learning with time-independent combinations of parameters provides close to the ground truth tumor profile, then the learning with time-dependent combinations makes the inference hardly useful.

Finally, we also analyzed the robustness of the proposed inverse network against the wide range of model parameters. Fig. 5 demonstrates the distribution of the mean absolute error over the range of values for the parameters μ_1, μ_2 pertaining to the test set of the FK model. From these scatter plots, we conclude that the network performance is stable across the ranges used for the test set.

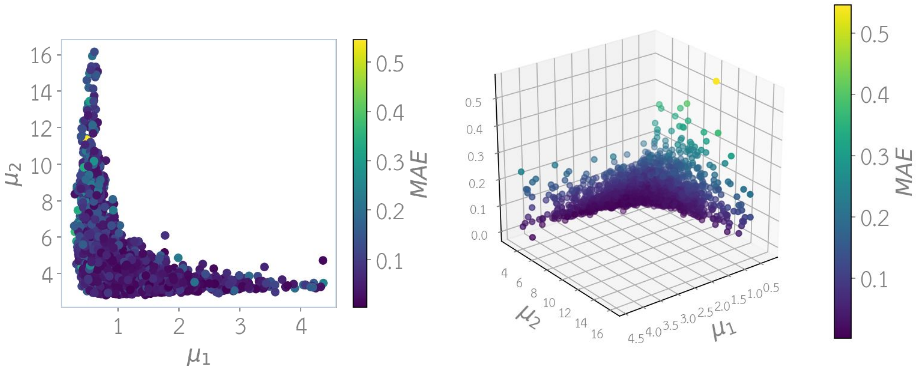

Figure 5: Distribution of the mean absolute error over the range of values for the parameters μ_1, μ_2, in 2D (left) and 3D (right) views.

3.4. Real MRI patient data

We performed qualitative validation of the method on a large cohort of brain tumor patients who underwent MRI testing. Binary segmentation corresponding to the T1Gd and FLAIR modalities were used as input to the *Learn-Morph-*

Infer method. Fig. 4 showcases examples of inferred tumor simulations for various tumor grades and locations in patients' brains. Out of 80 cases, there were less than 10% cases in which the pipeline outputted wrong results: tumor occupying the whole brain or absence of tumor.

To quantitatively estimate the accuracy of the proposed method, we compared it with the publicly available glioma solver from (Lipkova et al., 2019). Due to the computationally costly inference of the latter, we evaluated the performance on 10 randomly chosen glioma patients. The results of the comparison

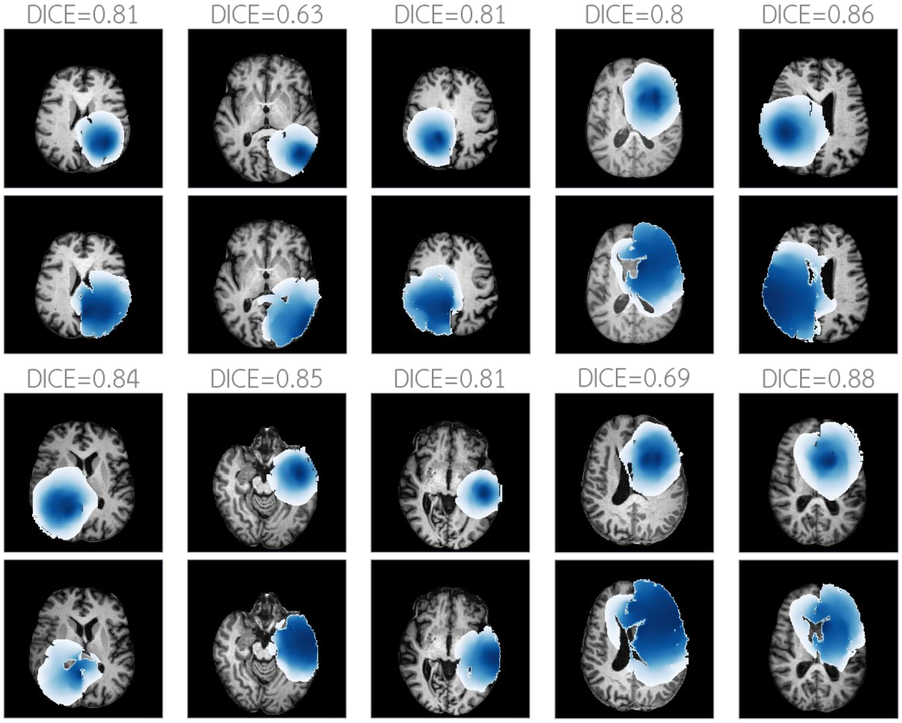

Figure 6: Examples of patient-specific simulations produced using the *Learn-Morph-Infer* method (above) and the MCMC-based glioma solver from (Lipkova et al., 2019) (below) for the FK model. Despite of the different nature of the methods and inevitable errors coming from the network misprediction, the DICE score for most of the cases is around 0.8 (the DICE was computed after thresholding the tumor cell concentration at $c_t = 0.01$).

are provided in Fig. 6. For most of the cases, the overlap in terms of DICE is around 0.8. Note that this error includes not only the contribution from network misprediction and registration back and forth to the brain atlas space but likely also from just different nature of the methods (our proposed method provides point estimates, while the glioma inverse solver by (Lipkova et al., 2019) is a probabilistic method). A natural extension of our work to reduce the error would be to run the sampling-based (or optimization) method in patient space after performing inference with the proposed *Learn-Morph-Infer* method but within narrow parametric ranges centered around estimates predicted by our learnable method. Or even a simpler extension - instead of the steps "c" and "d" in Fig.1, one can run the tumor solver directly in patient space with model parameters inferred at step "b" (such procedure would mitigate the error coming from the mismatch between brain atlas and patient anatomy).

3.5. Computing time

The total time including registration, morphing to atlas space, inference, tumor simulation, and morphing back to patient space is 4-7 minutes. The time for the inverse model network's inference is around 2 seconds for 128^3 resolution. Crucially, the inference time for the more complicated model with mass effect stays the same as for the Fisher-Kolmogorov tumor model. This emphasizes the key practical contribution of the proposed method in that it allows constant time model personalization for an arbitrary tumor model complexity.

4. Conclusion

We present a learnable brain tumor model personalization methodology. We demonstrate that it is feasible to learn an inverse model in a supervised fashion from a data set of numerical simulations. We show that the choice of output can crucially affect network's performance - predicting time-independent combinations of parameters notably outperforms time-dependent counterparts. Such time-independent parametrization is not limited to the PDEs considered here, and thus the proposed *Learn-Morph-Infer* pipeline can be adapted to other inverse problems in natural science and engineering disciplines. For the brain tumor growth model, the *Learn-Morph-Infer* pipeline provides real-time performance of the parametric inference. Most importantly, the personalization time is stable across tumor models of different numerical complexity. These performance benefits pave the way for clinical testing of various mathematical tumor descriptions on a large cohort of patients.

5. Acknowledgment

I. Ezhov, S. Shit and L. Canalini are supported by the Translational Brain Imaging Training Network under the EU Marie Sklodowska-Curie program (Grant ID: 765148). B. Menze, B. Wiestler, and F. Kofler are supported through the SFB 824, subproject B12, by DFG through TUM International Graduate School of Science and Engineering, GSC 81, and by the Institute for Advanced Study, funded by the German Excellence Initiative. The authors acknowledge funding from DCoMEX (Grant agreement ID: 956201). B. Menze is supported by Helmut Horten Foundation.

References

B. B. Avants, N. Tustison, and G. Song. Advanced normalization tools (ants). *Insight j*, 2(365):1–35, 2009.

M. Dax, S. R. Green, J. Gair, J. H. Macke, A. Buonanno, and B. Schölkopf. Real-time gravitational wave science with neural posterior estimation. *Physical review letters*, 127(24):241103, 2021.

I. Ezhov, T. Mot, S. Shit, J. Lipková, J. C. Paetzold, F. Kofler, F. Navarro, C. Pellegrini, M. Kollovieh, M. Metz, B. Wiestler, and B. Menze. Geometry-aware neural solver for fast bayesian calibration of brain tumor models. 2020.

I. Ezhov et al. Neural parameters estimation for brain tumor growth modeling. In *MICCAI*, pages 787–795. 2019.

E. Geremia, B. H. Menze, M. Prastawa, M.-A. Weber, A. Criminisi, and N. Ayache. Brain tumor cell density estimation from multi-modal mr images based on a synthetic tumor growth model. In *International MICCAI Workshop on Medical Computer Vision*, pages 273–282. Springer, 2012.

H. L. Harpold, E. C. Alvord Jr, and K. R. Swanson. The evolution of mathematical modeling of glioma proliferation and invasion. *Journal of Neuropathology & Experimental Neurology*, 66(1):1–9, 2007.

K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778, 2016. doi: 10.1109/CVPR.2016.90.

C. Hogea et al. An image-driven parameter estimation problem for a reaction–diffusion glioma growth model with mass effects. *J Math Biol*, 56(6):793–825, 2008.

D. A. Hormuth, S. L. Eldridge, J. A. Weis, M. I. Miga, and T. E. Yankeelov. Mechanically coupled reaction-diffusion model to predict glioma growth: methodological details. In *Cancer Systems Biology*, pages 225–241. Springer, 2018.

D. A. Hormuth, A. M. Jarrett, X. Feng, and T. E. Yankeelov. Calibrating a predictive model of tumor growth and angiogenesis with quantitative mri. *Annals of biomedical engineering*, 47(7):1539–1551, 2019.

D. A. Hormuth et al. Image-based personalization of computational models for predicting response of high-grade glioma to chemoradiation. *Scientific Reports*, 11(1):1–14, 2021.

P. R. Jackson, J. Juliano, A. Hawkins-Daarud, R. C. Rockne, and K. R. Swanson. Patient-specific mathematical neuro-oncology: Using a simple proliferation and invasion tumor model to inform clinical practice. *BMB*, 77(5): 846–856, mar 2015.

M. Kasim et al. Up to two billion times acceleration of scientific simulations with deep neural architecture search. *arXiv preprint arXiv:2001.08055*, 2020.

B. Kim et al. Deep fluids: A generative network for parameterized fluid simulations. *Comput. Graph. Forum*, 38:59–70, 2018.

D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. In *ICLR 2015*, 2015.

F. Kofler et al. Brats toolkit: Translating brats brain tumor segmentation algorithms into clinical and scientific practice. *Frontiers in Neuroscience*, 14, 2020.

E. Konukoglu et al. Extrapolating glioma invasion margin in brain magnetic resonance images: Suggesting new irradiation margins. *MedIA*, 14(2):111 – 125, 2010a.

E. Konukoglu et al. Image guided personalization of reaction-diffusion type tumor growth models using modified anisotropic eikonal equations. *IEEE Transactions on Medical Imaging*, 29(1):77–95, 2010b.

M. Le et al. Personalized radiotherapy planning based on a computational tumor growth model. *IEEE Transactions on Medical Imaging*, 36(3):815–825, mar 2017.

J. Lipkova et al. Personalized radiotherapy design for glioblastoma: Integrating mathematical tumor models, multimodal scans and bayesian inference. *IEEE Transactions on Medical Imaging*, pages 1–1, 2019.

G. Lorenzo and et al. Quantitative in vivo imaging to enable tumor forecasting and treatment optimization. *arXiv preprint arXiv:2102.12602*, 2021.

I. Loshchilov and F. Hutter. Decoupled weight decay regularization. In *International Conference on Learning Representations*, 2019. URL `https://openreview.net/forum?id=Bkg6RiCqY7`.

J.-M. Lueckmann et al. Flexible statistical inference for mechanistic models of neural dynamics. In *NeurIPS*, pages 1289–1299, 2017.

M. Lê, H. Delingette, J. Kalpathy-Cramer, E. R. Gerstner, T. Batchelor, J. Unkelbach, and N. Ayache. Mri based bayesian personalization of a tumor growth model. *IEEE Transactions on Medical Imaging*, 35(10):2329–2339, 2016. doi: 10.1109/TMI.2016.2561098.

A. Mang, A. Toma, T. A. Schuetz, S. Becker, T. Eckey, C. Mohr, D. Petersen, and T. M. Buzug. Biophysical modeling of brain tumor progression: From unconditionally stable explicit time integration to an inverse problem with parabolic pde constraints for model calibration. *Medical Physics*, 39(7Part1): 4444–4459, 2012.

B. Menze et al. A generative approach for image-based modeling of tumor growth. In *IPMI*, pages 735–747, 2011. ISBN 978-3-642-22092-0.

B. H. Menze et al. The multimodal brain tumor image segmentation benchmark (brats). *IEEE transactions on medical imaging*, 34(10):1993–2024, 2014.

G. Papamakarios and I. Murray. Fast ϵ-free inference of simulation models with bayesian conditional density estimation. In *NeurIPS*, pages 1028–1036. 2016.

V. Patel and L. Hathout. Image-driven modeling of the proliferation and necrosis of glioblastoma multiforme. *Theoretical Biology and Medical Modelling*, 14(1): 1–9, 2017.

S. Pati, V. Sharma, H. Aslam, S. Thakur, H. Akbari, A. Mang, S. Subramanian, G. Biros, C. Davatzikos, and S. Bakas. Tmod-09. glioblastoma biophysical growth estimation using deep learning-based regression. *Neuro-Oncology*, 22 (Supplement_2):ii229–ii229, 2020.

M. Raissi, P. Perdikaris, and G. E. Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, 2019.

T. Rohlfing, N. M. Zahr, E. V. Sullivan, and A. Pfefferbaum. The sri24 multichannel atlas of normal adult human brain structure. *Human brain mapping*, 31(5):798–819, 2010.

K. Scheufele, A. Mang, A. Gholami, C. Davatzikos, G. Biros, and M. Mehl. Coupling brain-tumor biophysical models and diffeomorphic image registration. *Computer methods in applied mechanics and engineering*, 347:533–567, 2019.

K. Scheufele, S. Subramanian, and G. Biros. Automatic mri-driven model calibration for advanced brain tumor progression analysis. *arXiv: Medical Physics*, 2020.

V. Sitzmann et al. Implicit neural representations with periodic activation functions. *arXiv preprint arXiv:2006.09661*, 2020.

B. Stevens and T. Colonius. Finitenet: A fully convolutional lstm network architecture for time-dependent partial differential equations. *arXiv preprint arXiv:2002.03014*, 2020.

R. Stupp, M. Brada, M. Van Den Bent, J.-C. Tonn, and G. Pentheroudakis. High-grade glioma: Esmo clinical practice guidelines for diagnosis, treatment and follow-up. *Annals of oncology*, 25:iii93–iii101, 2014.

S. Subramanian, A. Gholami, and G. Biros. Simulation of glioblastoma growth using a 3d multispecies tumor model with mass effect. *Journal of mathematical biology*, 79(3):941–967, 2019.

S. Subramanian, K. Scheufele, N. Himthani, and G. Biros. Multiatlas calibration of biophysical brain tumor growth models with mass effect. *arXiv preprint arXiv:2006.09932*, 2020a.

S. Subramanian, K. Scheufele, M. Mehl, and G. Biros. Where did the tumor start? an inverse solver with sparse localization for tumor growth models. *Inverse Problems*, 36(4):045006, 2020b.

K. R. Swanson, E. C. Alvord Jr, and J. Murray. A quantitative model for differential motility of gliomas in grey and white matter. *Cell proliferation*, 33(5):317–329, 2000.

N. Thuerey, K. Weißenow, L. Prantl, and X. Hu. Deep learning methods for reynolds-averaged navier–stokes simulations of airfoil flows. *AIAA Journal*, 58(1):15–26, 2020.

B. Tunc, D. Hormuth, G. Biros, and T. E. Yankeelov. Modeling of glioma growth with mass effect by longitudinal magnetic resonance imaging. *IEEE Transactions on Biomedical Engineering*, 2021.

B. Tunç, D. A. Hormuth II, G. Biros, and T. E. Yankeelov. Modeling of glioma growth with mass effect by longitudinal magnetic resonance imaging. *IEEE Transactions on Biomedical Engineering*, 68(12):3713–3724, 2021.

T. E. Yankeelov, N. Atuegwu, D. Hormuth, J. A. Weis, S. L. Barnes, M. I. Miga, E. C. Rericha, and V. Quaranta. Clinically relevant modeling of tumor growth and treatment response. *Science translational medicine*, 5(187): 187ps9–187ps9, 2013.

H. Young, R. Baum, U. Cremerius, K. Herholz, O. Hoekstra, A. Lammertsma, J. Pruim, P. Price, et al. Measurement of clinical and subclinical tumour response using [18f]-fluorodeoxyglucose and positron emission tomography: review and 1999 eortc recommendations. *European journal of cancer*, 35(13): 1773–1782, 1999.

Concluding Remarks

Although there has been progress in treating gliomas in recent years, they remain a significant public health problem. Ongoing research is needed to improve outcomes for patients with such tumors. This dissertation focused on computational glioma modeling and one of its main bottlenecks in translating the modeling into clinical practice - the time it takes to personalize the biophysical model. Paper-by-paper, we were addressing methodological constraints resulting in the computational bottleneck:

- Our first work resulted in the development of a surrogate based on a neural network that can substitute a numerical forward glioma solver while achieving 50 times speed up, Chapter 5.
- In the next part of the thesis, Chapter 6, we addressed the issue of sampling scale, i.e., the number of samples required to reach parametric convergence. We proposed a learnable method that allows reducing by one order the number of samples for glioma model personalization.
- Finally, in the last Chapter 7, we demonstrated a new scheme for model personalization that does not require neither sampling nor optimization procedure while bringing the compute times for every new patient to the clinically acceptable range.

The main implications of the results achieved within the thesis are twofold: the possibility of prompt testing tumor models of varying complexity and consequently validating these models on a large patient cohort. Both implications are crucial for a clinical translation.

First, modeling brain tumor growth is a complex task, and there is still much to be understood about the biological processes involved. As such, the specific form of the reaction, diffusion, advection, and other potential terms in a brain tumor growth model may vary and need to be modified as a deeper understanding of the pathology will be reached. This new information would likely necessitate constructing glioma models of higher computational complexity. Traditional model personalization methods would not allow performing this within practically adequate time. Second, our methods allow now validation on a statistically significant sample of patients. This is necessary for potential clinical trials examining radiotherapy strategies defined through image-based modeling.

Reflecting on the limitations of the developed techniques, we have to admit that, despite the gains in computational time during inference brought by our proposed methods, the training phase requires a collection of simulations in the order of thousands. The latter is achieved by employing numerical solvers as they provide error guarantees on the PDE solution. Using the numerical solvers, the time for collecting the training dataset scales with the model complexity. Thus, future research probing more advanced glioma models would benefit by developing similar learnable forward solvers that can provide error guarantees without undermining the computational speed up. Another limitation worth mentioning is the approximation error of the simulation domain. Our proposed methods rely on patient-specific brain tissues, obtained by atlas registration, as the simulation domain. More elaborate tumor models can be more dependent on brain tissue morphology and thus be more sensitive to errors in inferred tissue maps. Therefore, the proposed works might need to be adjusted for more complicated tumor models: either using more accurate registration or avoiding the registration step by learning to reconstruct the brain tissues in the patient geometry [55].

This doctoral dissertation set a goal of developing learnable personalization of brain tumor growth models by marrying well-established biophysical numerics with emerging data-driven paradigms. Despite the mentioned limitations, from the computational time point of view, the marriage outcome allows us to consider the model personalization for radiotherapy planning. From the point of view of the plausibility of the used tumor growth descriptions, there is clearly room for improvement. However, given that the standard irradiation volume resorts merely to a rule-of-thumb of a few-centimeter uniform extension around the tumor area, we hope that even simplistic tumor models used in this thesis might be advantageous compared to the standard planning.

Appendices

Bibliography

[1] A. Ramesh, M. Pavlov, G. Goh, S. Gray, C. Voss, A. Radford, M. Chen, and I. Sutskever. “Zero-shot text-to-image generation.” In: *International Conference on Machine Learning*. PMLR. 2021, pp. 8821–8831.

[2] R. Stupp, M. Brada, M. Van Den Bent, J.-C. Tonn, and G. Pentheroudakis. “High-grade glioma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up.” In: *Annals of oncology* 25 (2014), pp. iii93–iii101.

[3] M. E. Davis. “Glioblastoma: overview of disease and treatment.” In: *Clinical journal of oncology nursing* 20.5 (2016), S2.

[4] E. C. Halperin, G. Bentel, E. R. Heinz, and P. C. Burger. “Radiation therapy treatment planning in supratentorial glioblastoma multiforme: an analysis based on post mortem topographic anatomy with CT correlations.” In: *International Journal of Radiation Oncology* Biology* Physics* 17.6 (1989), pp. 1347–1350.

[5] R. Stupp, W. P. Mason, M. J. Van Den Bent, M. Weller, B. Fisher, M. J. Taphoorn, K. Belanger, A. A. Brandes, C. Marosi, U. Bogdahn, et al. “Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma.” In: *New England journal of medicine* 352.10 (2005), pp. 987–996.

[6] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer. “High-Resolution Image Synthesis With Latent Diffusion Models.” In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 2022, pp. 10684–10695.

[7] A. Mang, S. Bakas, S. Subramanian, C. Davatzikos, and G. Biros. “Integrated biophysical modeling and image analysis: application to neuro-oncology.” In: *Annual review of biomedical engineering* 22 (2020), p. 309.

[8] N. V. Krakhmal, M. Zavyalova, E. Denisov, S. Vtorushin, and V. Perelmuter. “Cancer invasion: patterns and mechanisms.” In: *Acta Naturae* 7.2 (25) (2015), pp. 17–28.

[9] G. Oliver. “Lymphatic vasculature development.” In: *Nature Reviews Immunology* 4.1 (2004), pp. 35–45.

[10] V. Leinonen, R. Vanninen, and T. Rauramaa. “Raised intracranial pressure and brain edema.” In: *Handbook of Clinical Neurology* 145 (2018), pp. 25–37.

[11] I. I. Rabi, J. R. Zacharias, S. Millman, and P. Kusch. “A new method of measuring nuclear magnetic moment.” In: *Physical review* 53.4 (1938), p. 318.

[12] E. M. Purcell, H. C. Torrey, and R. V. Pound. “Resonance absorption by nuclear magnetic moments in a solid.” In: *Physical review* 69.1-2 (1946), p. 37.

[13] F. Bloch. “Nuclear induction.” In: *Physical review* 70.7-8 (1946), p. 460.

[14] E. F. Jackson, L. E. Ginsberg, D. F. Schomer, and N. E. Leeds. “A review of MRI pulse sequences and techniques in neuroimaging.” In: *Surgical Neurology* 47.2 (1997), pp. 185–199.

[15] D. L. Bailey, M. N. Maisey, D. W. Townsend, and P. E. Valk. *Positron emission tomography*. Vol. 2. Springer, 2005.

[16] C. L. Vaughan and B. M. Mayosi. “Origins of computed tomography.” In: *The Lancet* 369.9568 (2007), p. 1168.

[17] M. Bausart, V. Préat, and A. Malfanti. “Immunotherapy for glioblastoma: the promise of combination strategies.” In: *Journal of Experimental & Clinical Cancer Research* 41.1 (2022), pp. 1–22.

[18] S. B. Tatter, E. G. Shaw, M. L. Rosenblum, K. C. Karvelis, L. Kleinberg, J. Weingart, J. J. Olson, I. R. Crocker, S. Brem, J. L. Pearlman, et al. “An inflatable balloon catheter and liquid 125I radiation source (GliaSite Radiation Therapy System) for treatment of recurrent malignant glioma: multicenter safety and feasibility trial.” In: *Journal of neurosurgery* 99.2 (2003), pp. 297–303.

[19] L. Rogers, E. Shaw, J. Rock, A. Sills, M. Vogelbaum, and M. Ewend. “Interim results of a Phase II study of resection and GliaSite brachytherapy for a single brain metastasis.” In: *International Journal of Radiation Oncology, Biology, Physics* 60.1 (2004), S259–S260.

[20] R. Airley. *Cancer chemotherapy: basic science to the clinic*. John Wiley & Sons, 2009.

[21] S. Benzekry, C. Lamont, A. Beheshti, A. Tracz, J. M. Ebos, L. Hlatky, and P. Hahnfeldt. “Classical mathematical models for description and prediction of experimental tumor growth.” In: *PLoS computational biology* 10.8 (2014), e1003800.

[22] C. Vaghi, A. Rodallec, R. Fanciullino, J. Ciccolini, J. P. Mochel, M. Mastri, C. Poignard, J. M. Ebos, and S. Benzekry. “Population modeling of tumor growth curves and the reduced Gompertz model improve prediction of the age of experimental tumors.” In: *PLoS computational biology* 16.2 (2020), e1007178.

[23] L. Buchauer, M. A. Khan, Y. Zhuo, C. Shao, P. Zou, W. Feng, M. Qian, G. Bekki, C. Bunne, A. Neuerburg, et al. “Exponential Growth of Glioblastoma In Vivo Driven by Rapidly Dividing and Outwardly Migrating Cancer Stem Cells.” In: *BioRxiv* (2019), p. 723601.

[24] A. L. Stensjøen, O. Solheim, K. A. Kvistad, A. K. Håberg, Ø. Salvesen, and E. M. Berntsen. “Growth dynamics of untreated glioblastomas in vivo.” In: *Neuro-oncology* 17.10 (2015), pp. 1402–1411.

[25] J. A. Koziol, T. J. Falls, and J. E. Schnitzer. “Different ODE models of tumor growth can deliver similar results.” In: *BMC cancer* 20.1 (2020), pp. 1–10.

[26] B. Ribba, G. Kaloshi, M. Peyre, D. Ricard, V. Calvez, M. Tod, B. Čajavec-Bernard, A. Idbaih, D. Psimaras, L. Dainese, et al. “A Tumor Growth Inhibition Model for Low-Grade Glioma Treated with Chemotherapy or Radiotherapy.” In: *Clinical Cancer Research* 18.18 (2012), pp. 5071–5080.

[27] R. Rockne, E. Alvord, J. Rockhill, and K. Swanson. “A mathematical model for brain tumor response to radiation therapy.” In: *Journal of mathematical biology* 58.4 (2009), pp. 561–578.

[28] L. Grinberg, D. A. Fedosov, and G. E. Karniadakis. “Parallel multiscale simulations of a brain aneurysm.” In: *Journal of computational physics* 244 (2013), pp. 131–147.

[29] F. G. Ashby and S. Helie. “The neurodynamics of cognition: A tutorial on computational cognitive neuroscience.” In: *Journal of mathematical psychology* 55.4 (2011), p. 273.

[30] E. Crampin and P. Maini. “Reaction-diffusion models for biological pattern formation.” In: (2001).

[31] C. Cosner. “Reaction–diffusion equations and ecological modeling.” In: *Tutorials in mathematical biosciences IV*. Springer, 2008, pp. 77–115.

[32] S. Kondo and T. Miura. “Reaction-diffusion model as a framework for understanding biological pattern formation.” In: *science* 329.5999 (2010), pp. 1616–1620.

[33] O. Saut, J.-B. Lagaert, T. Colin, and H. M. Fathallah-Shaykh. “A multilayer grow-or-go model for GBM: effects of invasive cells and anti-angiogenesis on growth.” In: *Bulletin of mathematical biology* 76.9 (2014), pp. 2306–2333.

[34] J. Liu, D. A. Hormuth, J. Yang, and T. E. Yankeelov. “A Multi-Compartment Model of Glioma Response to Fractionated Radiation Therapy Parameterized via Time-Resolved Microscopy Data.” In: *Frontiers in oncology* (2022), p. 257.

[35] G. Vilanova, I. Colominas, and H. Gomez. “A mathematical model of tumour angiogenesis: growth, regression and regrowth.” In: *Journal of The Royal Society Interface* 14.126 (2017), p. 20160918.

[36] V. Cristini and J. Lowengrub. *Multiscale modeling of cancer: an integrated experimental and mathematical modeling approach.* Cambridge University Press, 2010.

[37] J. D. Murray. *Mathematical biology: I. An introduction.* Springer, 2002.

[38] S. Subramanian, A. Gholami, and G. Biros. “Simulation of glioblastoma growth using a 3D multispecies tumor model with mass effect.” In: *Journal of mathematical biology* 79.3 (2019), pp. 941–967.

[39] L. T. Biegler, O. Ghattas, M. Heinkenschloss, D. Keyes, and B. van Bloemen Waanders. *Real-time PDE-constrained Optimization.* SIAM, 2007.

[40] A. N. Tikhonov. “On the stability of inverse problems.” In: *Dokl. Akad. Nauk SSSR.* Vol. 39. 1943, pp. 195–198.

[41] H. Antil, D. P. Kouri, M.-D. Lacasse, and D. Ridzal. *Frontiers in PDE-constrained Optimization.* Vol. 163. Springer, 2018.

[42] A. Mang, A. Gholami, C. Davatzikos, and G. Biros. “PDE-constrained optimization in medical image analysis.” In: *Optimization and Engineering* 19.3 (2018), pp. 765–812.

[43] K. Scheufele, S. Subramanian, A. Mang, G. Biros, and M. Mehl. “Image-driven biophysical tumor growth model calibration.” In: *SIAM journal on scientific computing: a publication of the Society for Industrial and Applied Mathematics* 42.3 (2020), B549.

[44] C. Hogea, C. Davatzikos, and G. Biros. “An image-driven parameter estimation problem for a reaction–diffusion glioma growth model with mass effects.” In: *Journal of mathematical biology* 56.6 (2008), pp. 793–825.

[45] Y. Liu, S. M. Sadowski, A. B. Weisbrod, E. Kebebew, R. M. Summers, and J. Yao. “Patient specific tumor growth prediction using multimodal images.” In: *Medical image analysis* 18.3 (2014), pp. 555–566.

[46] D. Knopoff, D. Fernández, G. Torres, and C. Turner. “A mathematical method for parameter estimation in a tumor growth model.” In: *Computational and Applied Mathematics* 36.1 (2017), pp. 733–748.

[47] B. Liang, L. Lozenski, U. Villa, and D. Faghihi. “PDE-constrained Optimization Formulation for Tumor Growth Model Calibration.” In: *arXiv preprint arXiv:2302.06445* (2023).

[48] C. M. Bishop and N. M. Nasrabadi. *Pattern recognition and machine learning.* Vol. 4. 4. Springer, 2006.

[49] T. Kloek and H. K. Van Dijk. “Bayesian estimates of equation system parameters: an application of integration by Monte Carlo.” In: *Econometrica: Journal of the Econometric Society* (1978), pp. 1–19.

[50] A. A. Markov. “Extension of the limit theorems of probability theory to a sum of variables connected in a chain.” In: *Dynamic probabilistic systems* 1 (1971), pp. 552–577.

[51] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. “Equation of state calculations by fast computing machines.” In: *The journal of chemical physics* 21.6 (1953), pp. 1087–1092.

[52] M. Lê, H. Delingette, J. Kalpathy-Cramer, E. R. Gerstner, T. Batchelor, J. Unkelbach, and N. Ayache. “Bayesian personalization of brain tumor growth model.” In: *International Conference on Medical Image Computing and Computer-Assisted Intervention.* Springer. 2015, pp. 424–432.

[53] 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: *Biennial International Conference on Information Processing in Medical Imaging.* Springer. 2011, pp. 735–747.

[54] J. Lipková, P. Angelikopoulos, S. Wu, E. Alberts, B. Wiestler, C. Diehl, C. Preibisch, T. Pyka, S. E. Combs, P. Hadjidoukas, et al. “Personalized radiotherapy design for glioblastoma: Integrating mathematical tumor models, multimodal scans, and bayesian inference.” In: *IEEE transactions on medical imaging* 38.8 (2019), pp. 1875–1884.

[55] F. Kofler, F. Meissen, F. Steinbauer, R. Graf, E. Oswald, E. de da Rosa, H. B. Li, U. Baid, F. Hoelzl, O. Turgut, et al. "The Brain Tumor Segmentation (BraTS) Challenge 2023: Local Synthesis of Healthy Brain Tissue via Inpainting." In: *arXiv preprint arXiv:2305.08992* (2023).