Data driven Computational Mechanics at EXascale

Data driven Computational Mechanics at EXascale

Work program topic: EuroHPC-01-2019
Type of action: Research and Innovation Action (RIA)

Data processing software final evaluation and documentation

DELIVERABLE D5.4

Version No 1

http://www.dcomex.eu/
This project has received funding from the European High-Performance Computing Joint Undertaking Joint Undertaking ('the JU'), under Grant Agreement No 956201

DOCUMENT SUMMARY INFORMATION

Project Title	**Data driven Computational Mechanics at EXascale**
Project Acronym	DCoMEX
Project No:	956201
Call Identifier:	EuroHPC-01-2019
Project Start Date	01/04/2021
Related work package	WP 5
Related task(s)	Task 5.1, Task 5.2, Task 5.3
Lead Organisation	TUM
Submission date	11/7/2024
Re-submission date	
Dissemination Level	PU

Quality Control:

	Who	Affiliation	Date
Checked by internal reviewer	I. Ezhov, B. Menze	TUM	1.7.2024
Checked by WP Leader	I. Ezhov	TUM	11.7.2024
Checked by Project Coordinator			

Document Change History:

Version	Date	Author (s)	Affiliation	Comment
1.0	11.7.2024	B. Menze, I. Ezhov	TUM	

Content

1. Description

This report outlines the capabilities, design, and implementation of the UQ-aware image processing module, including its functional requirements, existing technology leverage, modular framework, uncertainty quantification, and visualization tools, concluding with information on software availability. These software modules facilitate geometry reconstruction from 3D images and incorporate algorithms to handle uncertainties in image segmentation and boundary conditions, ultimately aiding downstream modeling in the DCoMEX framework.

This report also serves the purpose of documenting and demonstrating the use of the DCoMEX UQ-aware image segmentation software initially described in deliverables D5.1, D5.2, and D5.3. We will do so by using data, processing tasks, and specialized algorithms and mathematical models originating from the DCoMEX-BIO use case.

1.1 Functional and Technical Requirements for UQ-Aware Image Processing in DCoMEX

The UQ-aware image processing module aims to translate 3D image scans into well-defined spatial domains suitable for downstream modeling. During the course of the project, and driven by the DCoMEX usecase – and, in particular, the DCoMEX-Bio usecase on tumor growth modeling – we identified a number of functionalities the DCoMEX image processing need to enable:

1. **Segmentation Algorithms:** These algorithms define and delineate the modeling domain visible from images, creating subvolumes with task-specific boundary conditions. Nodes of these meshes may contain local properties.

2. **Local Model Parameters (e.g., Material Properties) and Boundary Conditions:** Nodes on the mesh or grid can be identified with local material properties relevant to the modeling, derived from the spatial context in the image domain. Similarly, image segments can be identified with relevant boundaries and boundary conditions for the modeling domain.

3. **Representing Uncertainties for Uncertainty Quantification:** The nodes/grids representing the modeling domain should express spatial uncertainties from image delineations, including uncertainties in local material properties.

4. **Enabling expert-guided Interaction with Image Data:** For applications requiring manual processing, tools should enable highly interactive image segmentations adaptable to new 3D image data. In the same way, a visualization of intermediate and final examples should be enabled by the DCoMEX tools.

5. **Automated Processing:** For applications with well-defined domains and pre-developed segmentation tools, these tools should integrate seamlessly, scaling efficiently to handle large datasets and to prepare them for large-scale computing tasks in MSolve.

2. Implementation of the DCoMEX UQ-Aware Image Processing Software

2.1. Design, Implementation, and Availability

The DCoMEX image processing tool employs a modular framework with well-defined interfaces to integrate the above tools. This modularity allows for customizable and scalable image processing workflows, including segmentation, meshing, and uncertainty quantification. The tool enables both pre- and post-processing capabilities, as listed in 1.2.

2.1.1. Modular data processing software architecture

A modular framework enables a mapping of functionalities to (image processing) functions, enabling a straightforward of individual tools, as well as, a pipeline set of processing tools. Modules and functions of the framework include:

1. **Workspace Manager**: Organizes data into workspaces, allowing easy storage and retrieval.
2. **Viewers**: Provides tools for 2D/3D data visualization and manipulation, including an in_built viewer and interfaces to other community visualization tools.
3. **Processing**: Supports plugins for various pre- and post-processing tasks. It is enabling integration with external image pre-processing tools, like ITK-SNAP and Ilastik, but also the DCoMEX-Bio “brats” and “tumor_growth” algorithms. It is also enabling the generation of summary statistics for post-processing tasks.
4. **Pipelines**: Users can create pipelines combining multiple plugins to automate processing tasks.

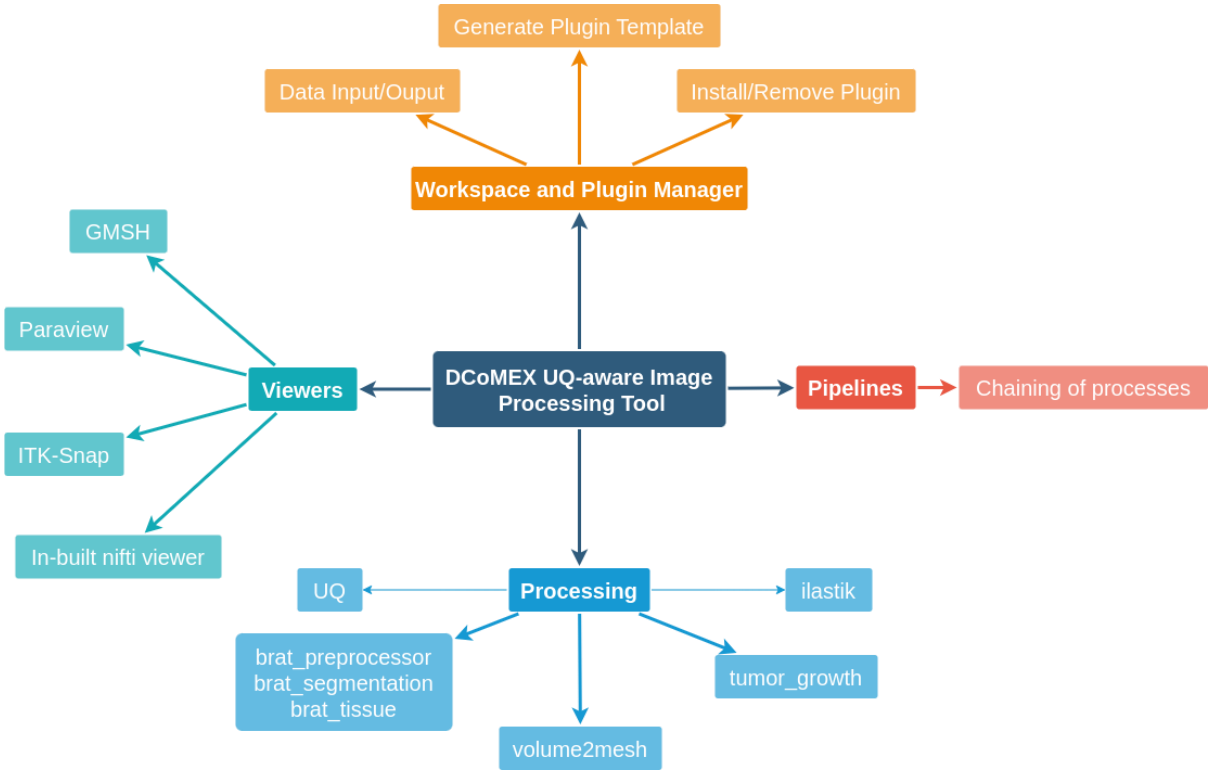

Figure 1: Modular architecture of the DCOMEX UQ-aware Image Processing Tool, offering interfaces to image import, visualization, interactive processing, and MSolve-specific meshing. Serving as data processing pipeline.

2.2.2 Integration with other community tools

The modular design of the DCoMEX image processing module leverages several established open-source software projects. This enriches its functionality and guarantees the further availability – and development - of relevant software functions of the toolbox via dedicated open-source software initiations.

1. **ITK and VTK**: These toolkits provide extensive functionality for 2D, 3D, and 4D data handling, visualization, and processing in the biomedical domain.

2. **ITK-snap**: A tool for multimodal 3D image annotation, offering semi-automatic segmentation and image navigation.
3. **Ilastik**: A user-friendly tool for ML-guided adaptive image segmentation, that is popular in bioimaging.
4. **BRATS-Toolkit**: Specifically for DCoMEX-Bio and brain tumor image segmentation, this tool offers comprehensive preprocessing, segmentation, and fusion capabilities.

2.2.3 Visualization and interaction

The DCoMEX tool also includes modules to visualize simulation outputs, particularly those affected by uncertainties in boundary conditions and model parameters.

The visualization module provides summary statistics like averages and standard deviations for simulation results. It integrates with other visualization tools in the DCoMEX pipeline and supports standard file formats for compatibility with specialized software like 3D Slicer and Napari.

2.2.4 Availability and licensing of the code

Key elements of the DCoMEX UQ aware image processing framework are written in Python to enable full integration with other computational tools and, at the same time, to enable users to modify and adjust the given to their specific needs. In addition, all tools are available as command line functions (to be used in Linux or Windows WLS), together with appropriate installation scripts to broaden the user base.

Module and its components are available as open-source software on GitHub via https://github.com/DComEX.

3. Dedicated Functions for UQ-Aware Data Handling in DCoMEX

3.1 UQ-Aware Preparation of Simulation Input

Using real world image data requires the user to abstract the representation of the problem under study, i.e., defining about boundaries or the localization of local material properties that impact on the simulation model. In many applications that aim at propagating uncertainties of the simulation and modeling task, however, a probabilistic representation is preferred over the abstracted one. The DCoMEX UQ-aware image processing tool offers functionalities and related representations to deal with these types of uncertainties:

1. **Probabilistic Segmentations**: Plugins like Ilastik and ITK-SNAP can generate probabilistic outputs, assigning class probabilities to pixels/voxels, which are then formatted for downstream MSolve simulation by using the ensemble approach proposed by [Krygier].
2. **Node Probabilities and Ensembles**: Uncertainties about local material properties can be represented as probabilities at mesh nodes or as ensembles of meshes sampling from probabilistic domains. For example, in the DCoMEX-Bio use case, tissue properties are associated with probabilities at the nodes of a structured grid. In this example, the downstream simulation models take continuous probabilistic values as input to the patient-specific prediction of tumor growth.

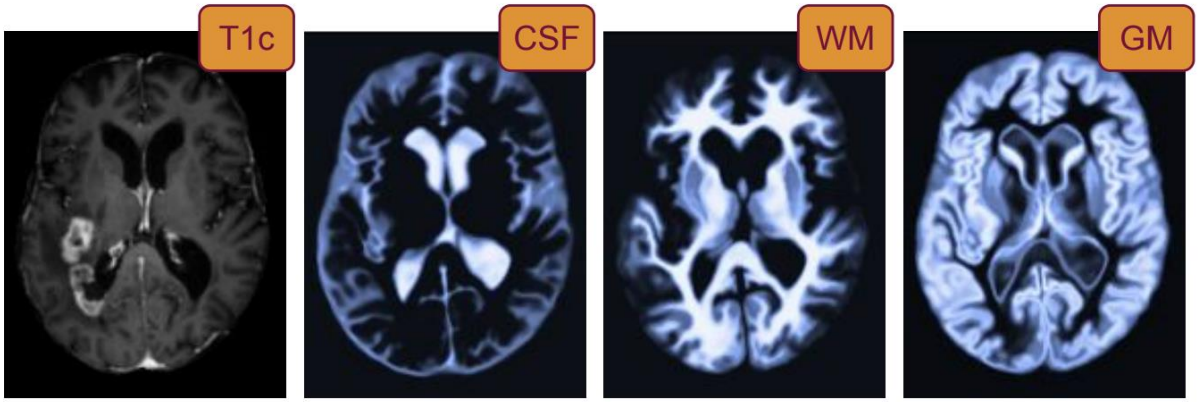

Figure 2. An example of probabilistic tissue segmentation (Cerebrospinal fluid (CSF), White matter (WM), Grey matter (GM)) and corresponding MRI T1c scan. The probabilistic segmentations offer information on the expected tissues underneath and nearby the tumor (from [Ezhov]).

3.2 Dealing with Simulation Output from UQ-Aware Simulations

In addition to uncertainties associated with boundary conditions, there may be uncertainties related to parameters of the simulation algorithm. For example, an exchange rate in a dynamical process may only be know up to a certain range or distribution. As a result, different simulation results are equally likely. In another example, parameters of a differential equation may be estimated from a set of (image) observations via Bayesian inference. Same as before, they may only be known only up to a distribution.

In both cases, the output of the forward simulations is likely to be represented in an ensemble of simulation output, for example, via Markov chain Monte Carlo and other importance sampling approaches. Other inference algorithms may directly generate summary statistics like means or variances. If they do so at the level of the node in the simulation domain, the resulting uncertainties need to be visualized and inspected in the original domains of the input data. This

functionality for visualizing local output of UQ-aware simulations is supported in the data processing tool.

More specifically, the tool is capable of dealing with both local and global uncertainties:

1. **Local Uncertainties**: Uncertainties in boundary conditions translate into local uncertainties in simulation results, requiring visualization of probabilistic segmentations and resulting simulation outputs. Dealing with ensembles of simulations is a direct consequence of the functionalities of the DCoMEX UQ-aware processing tools described above in Section 3.1.
2. **Global Uncertainties**: Uncertainties in global simulation parameters lead to voxel-wise distributions, necessitating visualization tools to summarize and display these uncertainties within the original image domain. These ensembles of (weighted) simulations results, for example, from Bayesian sampling algorithms like Korali (Fig. 3). Both summaries of discrete simulation output can be visualized, as well as summaries generated by the algorithm itself, like maximum a posteriori values, or standard deviations of state variables (e.g., see Fig. 3).

An example of summary statistics arising from global parametric uncertainty is given in Fig. 3.

4. The DCoMEX-Bio Usecase

The DCoMEX-Bio Usecase on tumor image analysis and patient-specific simulation of tumor growth has been a prominent driver project for conceiving

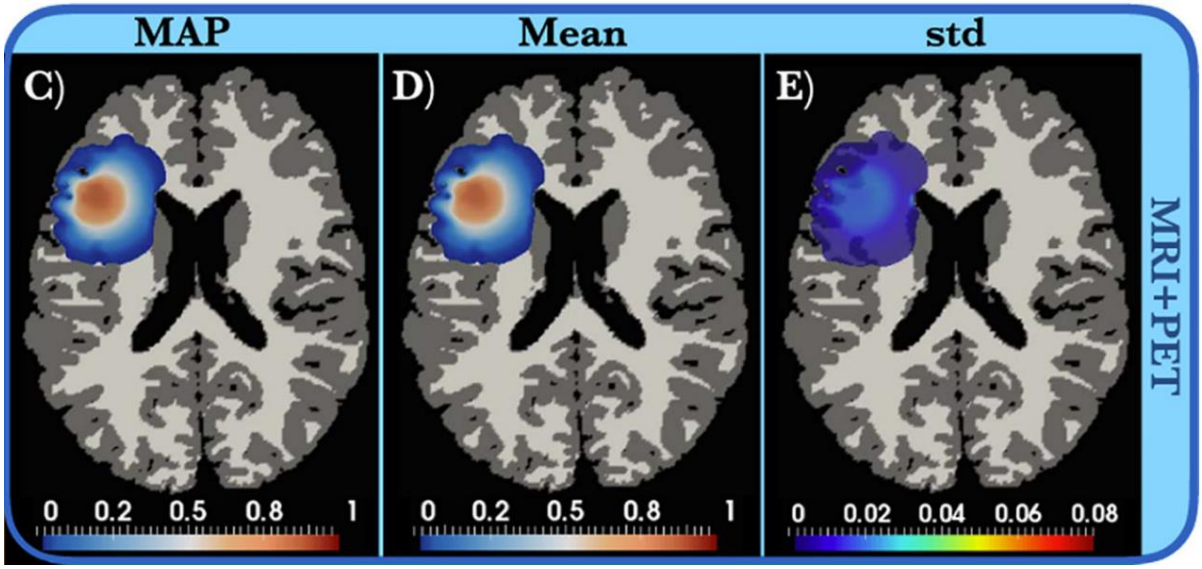

Figure 3. Visualization of simulated tumor in a patient anatomy. The modeling uncertainty results in probabilistic estimated of the tumor cell distribution: MAP, Mean, and std represent the maximum-a-posterior, mean and standar deviation of the distribution, respectively (from [Lipkova])

the design and implementing the DCoMEX UQ-aware image processing software. To enable the usecase, dedicated routines for preprocessing brain images have been implemented, several tumor growth simulation models in MSolve, as well as a Corali-based estimation of personalized model parameters, and a visualization of final simulations within the brain image geometry.

4.1 Pre-processing of Tumor Image Data

Preprocessing brain tumor image data for modeling tumor growth requires two interlinked image segmentation tasks: First, different magnetic resonance images – and, possibly, other image modalities – need to be standardize and aligned with a spatial template domain. Second, anatomical structures of interest need to be localized and segmented, as well as tumor structures and -substructures.

Following the commonly used Fisher-Kolmogorov model with anisotropic growth conditions [Lipkova, Ezhov, Balcerak], the anatomy defines patient-specific spatial boundary conditions in the subsequent tumor growth simulation, and outlines of the various tumor substructures represent targets in the subsequent model personalization, i.e., need to be matched by the simulation output.

4.1.1 Interactive ITK-snap and Ilastik segmentation.

Images can be aligned with tissue templates, also referred to as 3D anatomical "atlases", and resampled to standard resolution across all image modalities in various commercial and non-commercial medical image processing tools (such as 3D Slicer[1]). Upon co-alignment, they can be further processed with generic tools in the DCoMEX UQ-aware image processing software: Starting with the co-aligned atlases, the patient anatomy can be inspected and refined, if necessary, in an interactive fashion using, e.g., level set based label propagation algorithms from ITK-snap[2], or a ML-based segmentation using Ilastik[3].

Subsequently, the tumor can be segmented: ITK-snap, for example, has been the standard tool for interactive tumor image segmentation and an interactive segmentation by an expert can follow the annotation procedure established for the "Multimodal Brain Tumor Image Segmentation Challenge (BRATS)" [Menze]. Alternatively, a segmentation algorithm for delineating the 3D structures can be trained using Ilastik. Early on in the BRATS challenge, an Ilastik based segmentation using random forests and 3D local image features has won the BRATS benchmark [Menze]. It's advantage over the interactive ITK-

[1] www.slicer.ort
[2] www.itksnap.org
[3] www.ilastik.org

snap based segmentation is its ability to train on few examples (few annotated exemplary voxels, or few exemplary cases), and then the trained machine learning model can be applied to the remaining data, for example, by implementing an automated pipeline in the DCoMEX UQ-aware image processing tool.

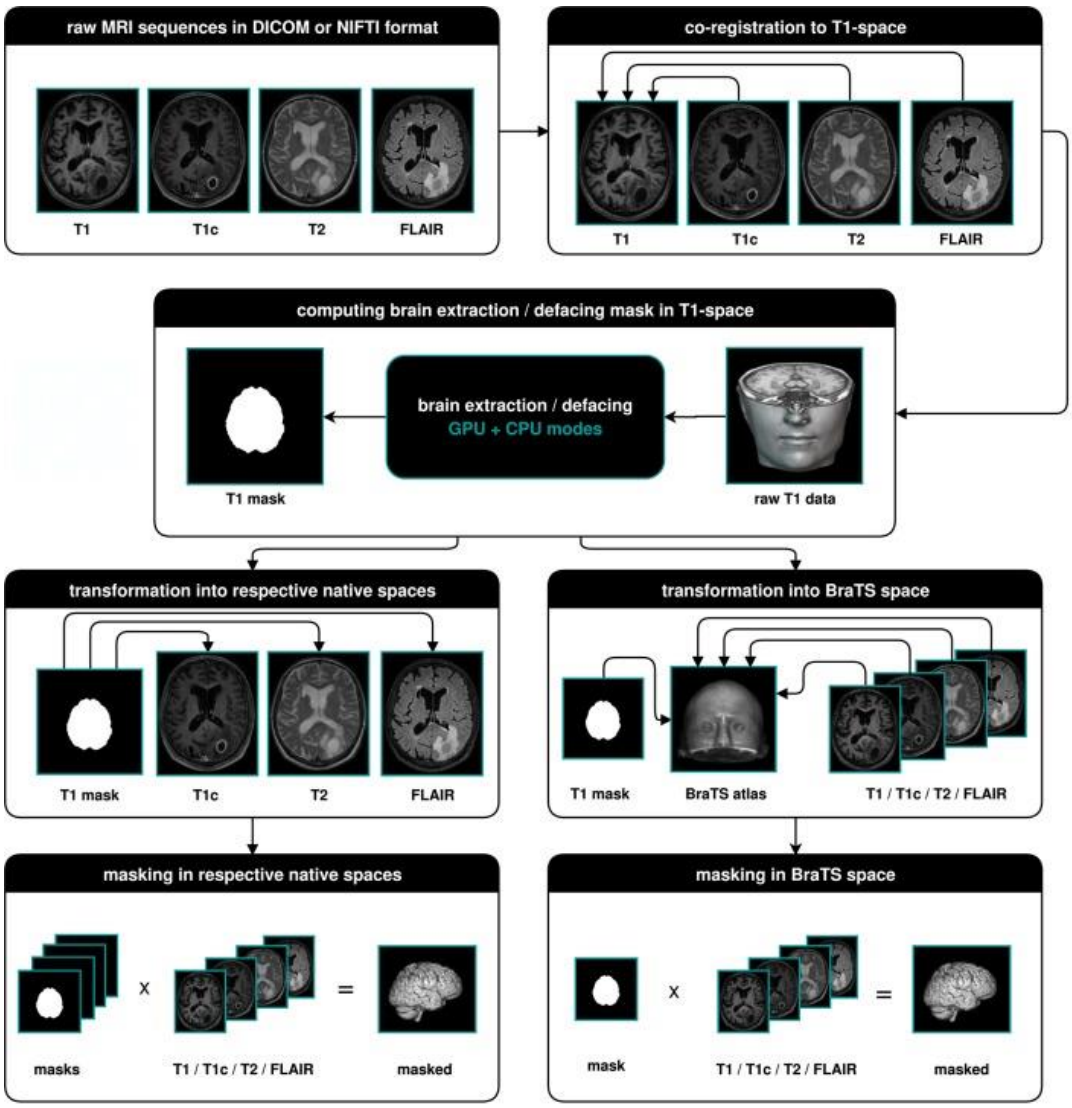

Figure 4: The tumor image data processing as implemented in the "brats" processing modules dealing with image registration (top row), masking (second row), and brain segmentation (third and fourth row). Functions are implemented in [Kofler]. Also see Fig 1 for tissue segmentation output.

4.1.2 Dedicated BRATS processing pipeline

Both ITK-snap and Ilastik based brain tumor image processing require expert knowledge, time, and prior data preparation. To this end, a dedicated module has been implemented that makes use of (a) established protocol for image processing for the BRATS benchmark including co-alignment and tissue segmentation via atlas alignment (Fig x, top), and (b) that enables the deployment of state-of-the-art 3rd party tumor image segmentation algorithms as made available in the BRATS benchmark (Fig x, bottom). Different algorithms can be chosen, including winning deep learning algorithms of the past years.

This set of algorithms is implemented as processing modules (Fig. 1) named "brats_preprocessing" for image alignment and resampling, "brats_tissue" for an atlas-based tissue segmentations, and "brats_segmentation" for the tumor segmentation.

4.2 Simulating Tumor Growth and Estimating Patient-Specific Model Parameters

The aforementioned Fisher-Kolmogorov partial differential equation is the classical tumor simulation model. It describes tumor growth as a reactive flow in porous medical, with the tumor proliferation being described as logistic growth and migration of tumor cells being modeled as a diffusive random walk. Its two key parameters are the diffusion constant and the growth constant. As mentioned, the consideration of anatomical boundary conditions – modifying the local tumor cell diffusion constant - is crucial for modeling it in a patient-specific personalization. We implemented this model in MSolve, also serving as a baseline for the more involved mathematical models for response to tumor

immunotherapy. We also implemented a model that is conditioning cell proliferation on local nutrient availability that, itself, is modeled as a diffusive process. In this model, the absence of nutrients leads to the rise of a second tumor mixture, i.e., necrotic cells. Fig. 5 shows simulations of this model adjusted to anatomy and tumor structure and image markers of an exemplary patient.

The DCoMEX preprocessing and modeling pipeline has been tested on a set of about 200 patients from TUM University Hospital [Balcerak]. The baseline models implemented in MSolve have been used, and different personalization strategies have been compared.

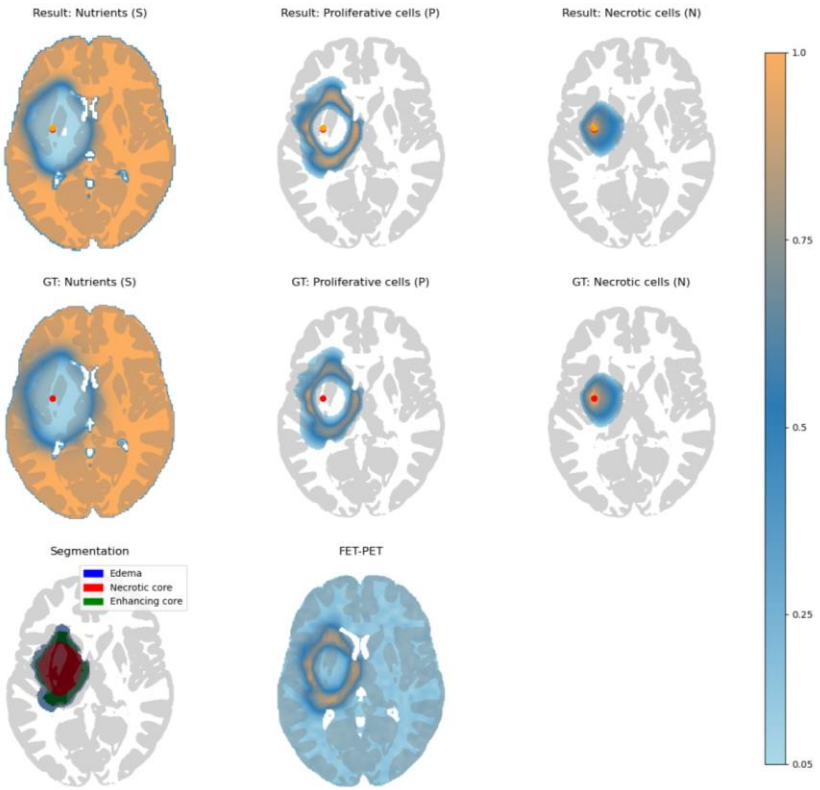

Figure 5: Tumor growth model with migrating, proliferating and necrotic cells making use of the "brats" image processing modules and implemented in MSolve. Shown are estimate tumor cell distributions.

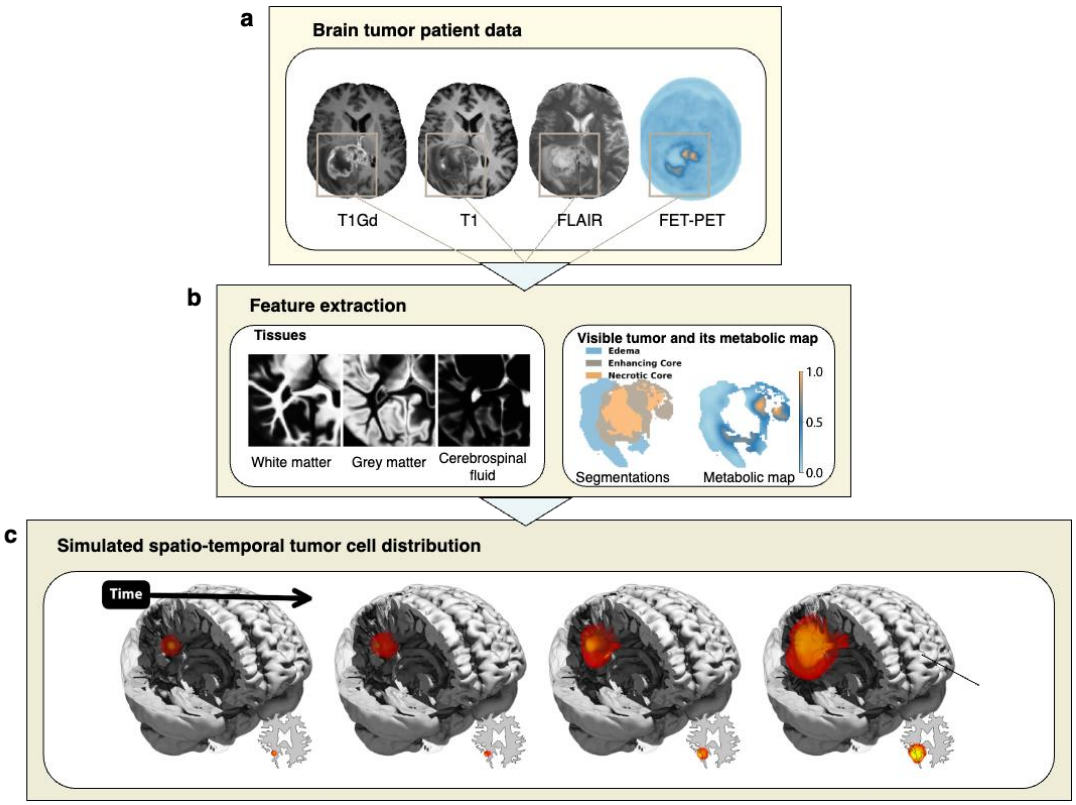

Figure 6: A pipeline of the DCoMEX-BIo usecase: a). The medical imaging scans (such as MRI and PET) are processed using the DCoMEX brat_preprocessor tool, b) The obtained tissue and tumor segmentation by means of the brat_tissue and brat_segmentation serve as input to the tumor growth models, c) The simulated tumor cell distribution can then be visualized using DCoMEX viewer plugin.

4.3 Post-Processing of Simulation Data and Visualization

Results returned from the MSolve simulation can be further processed and visualized in the original domain of the patient anatomy. This includes the joint visualization with the original image data (Fig 6a), as well as with the segmentations (Figs 2, 6b). The visualization of the estimated tumor cell concentrations (Fig. 6c) can be displayed as averages or other via other summary statistics.

5. References

- Kroeger et al. *ilastik: interactive machine learning for (bio)image analysis*. Nature Methods 2019

- Kofler et al. BraTS Toolkit: Translating BraTS Brain Tumor Segmentation Algorithms Into Clinical and Scientific Practice. Frontiers in Neuroscience 2020

- Krygier et al. Quantifying the unknown impact of segmentation uncertainty on image-based simulations. Nature Communications 2022.

- Yushkevich et al. User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. Neuroimage 2006

- Menze et al. The Multimodal Brain Tumor Image Segmentation Benchamark (BRATS). IEEE Transactions on Medical Imaging, 2014.

- Balcerak et al. Individualizing Glioma Radiotherapy Planning by Optimization of Data and Physics-Informed Discrete Loss. arXiv preprint arXiv:2312.05063. (2024)

- Ezhov et al., Geometry-aware neural solver for fast Bayesian calibration of brain tumor models, IEEE Transactions on Medical Imaging, 2021

- Lipková et al., Personalized Radiotherapy Design for Glioblastoma: Integrating Mathematical Tumor Models, Multimodal Scans, and Bayesian Inference, IEEE Transactions on Medical Imaging, 2019