



## **Deliverable No. 9.3**

# **A Multimodal and Longitudinal Brain Tumor Image Analysis Tool**

Grant Agreement No.:	600841
Deliverable No.:	D9.3
Deliverable Name:	<b>A multimodal and longitudinal brain tumor image analysis tool</b>
Contractual Submission Date:	31/01/2017
Actual Submission Date:	31/01/2017



Dissemination Level		
<b>PU</b>	Public	<b>X</b>
<b>PP</b>	Restricted to other programme participants (including the Commission Services)	
<b>RE</b>	Restricted to a group specified by the consortium (including the Commission Services)	
<b>CO</b>	Confidential, only for members of the consortium (including the Commission Services)	

COVER AND CONTROL PAGE OF DOCUMENT	
Project Acronym:	CHIC
Project Full Name:	Computational Horizons In Cancer (CHIC): Developing Meta- and Hyper-Multiscale Models and Repositories for In Silico Oncology
Deliverable No.:	D9.3
Document name:	A multimodal and longitudinal brain tumor image analysis tool
Nature (R, P, D, O) <sup>1</sup>	R
Dissemination Level (PU, PP, RE, CO) <sup>2</sup>	RE
Version:	1
Actual Submission Date:	31/01/2017
Editor: Institution: E-Mail:	Mauricio Reyes UBERN mauricio.reyes@istb.unibe.ch

#### ABSTRACT:

This deliverable 9.3. - **A multimodal and longitudinal brain tumor image analysis tool**, describes the development, testing and validation of algorithms, methods and associated technology for multimodal and longitudinal brain tumor image analysis. Specifically, novel methodologies based on human-machine intelligence (i.e. Machine Learning + Clinical Expert Domain Knowledge) were developed to perform fully automatic segmentation of glioblastomas from multi-sequence MRI imaging.

Results indicate the feasibility of utilizing the proposed technology in the clinical setup, ensuring an appropriate clinical workflow (i.e. execution time, data workflow and human-machine interactions), accuracy and performance, reproducibility and reliability of brain tumor analyses for temporally-resolved imaging information.

#### KEYWORD LIST:

Glioblastoma, automated image analysis, multi-compartmental, multimodality, longitudinal, tumor progression

<sup>1</sup> R=Report, P=Prototype, D=Demonstrator, O=Other

<sup>2</sup> PU=Public, PP=Restricted to other programme participants (including the Commission Services), RE=Restricted to a group specified by the consortium (including the Commission Services), CO=Confidential, only for members of the consortium (including the Commission Services)

*The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 600841.*

*The author is solely responsible for its content, it does not represent the opinion of the European Community and the Community is not responsible for any use that might be made of data appearing therein.*

<b>MODIFICATION CONTROL</b>			
<b>Version</b>	<b>Date</b>	<b>Status</b>	<b>Author</b>
1.0	24/01/2017	Draft	Mauricio Reyes
1.1	25/01/2017	Pre-final	Mauricio Reyes
1.2	31/01/2017	Final	Mauricio Reyes

#### List of contributors

- Raphael Meier, UBERN

## Contents

1	EXECUTIVE SUMMARY.....	9
2	INTRODUCTION .....	10
2.1	CURRENT STATUS IN THE TREATMENT OF GLIOBLASTOMA.....	10
2.2	STATE OF THE ART IN BRAIN TUMOR IMAGE SEGMENTATION .....	10
3	AUTOMATED BRAIN TUMOR SEGMENTATION .....	12
4	CLINICAL EVALUATION OF AUTOMATED BRAIN TUMOR IMAGE SEGMENTATION.....	13
5	COMPARISON TO OTHER STATE-OF-THE ART METHODS.....	14
5.1	EVALUATION OF BRATUMIA AND BRAINLAB'S FDA-APPROVED SMARTBRUSH SEGMENTATION .....	14
5.2	SUMMARY OF BRATUMIA FEATURES .....	15
6	AUTOMATED SEGMENTATION OF POST-OPERATIVE MRI PATIENT IMAGES.....	15
7	LONGITUDINAL BRAIN TUMOR SEGMENTATION .....	15
7.1	LONGITUDINAL BRAIN TUMOR IMAGE ANALYSIS OF HIGH GRADE GLIOMAS .....	15
7.2	CLINICAL EVALUATION OF A FULLY-AUTOMATIC SEGMENTATION METHOD FOR LONGITUDINAL BRAIN TUMOR VOLUMETRY	16
8	UNCERTAINTY ESTIMATION IN BRAIN TUMOR IMAGE SEGMENTATION.....	18
9	APPLICATIONS OF THE DEVELOPED TECHNOLOGY .....	20
9.1	ASSESSMENT OF EXTENT-OF-RESECTION (EOR) AND RESIDUAL-TUMOR (RT) THROUGH AUTOMATED TUMOR VOLUMETRY	20
9.2	AUTOMATED VOLUMETRY FOR RADIOMICS.....	21
10	CONCLUSIONS AND OUTLOOK .....	23
11	REFERENCES .....	23

## Figures

- Figure 1 Main interface of BraTumIA (Brain Tumor Image Analysis). Four views enable visualization of the four MR sequences. After segmentation, the segmentation can be overlaid on the input images where different colors are used to differentiate the tumor layers of necrotic, enhancing core, non-enhancing core, and edema. Volumes per tumor layer can also be retrieved and exported as a text file for further analysis. The segmentation results are also available on the user's local folder for further analysis. .... 13
- Figure 2 Dispersion of Dice scores for the individual algorithms (color coded), and various fused algorithmic segmentations (gray), shown together with the expert results (also shown in gray). Boxplots show quartile ranges of the scores on the test datasets; whiskers and dots indicate outliers. Black squares indicate the mean score. Also shown are results from four "Fused" algorithmic segmentations, and the performance of the "Best Combination" as the upper limit of individual algorithmic performance. The light-grayed rectangle highlights results from BraTumIA. In comparison to other approaches it is highlighted the low variance (higher robustness) of BraTumIA, in the range of the inter-expert variability (also highlighted with a light-grayed box). .... 14
- Figure 3 Exemplary case, T1c -weighted images (green = enhancing tumor), from left to right: Preoperative image, Postoperative image (relevant area is magnified), result for the proposed semi-supervised post-operative segmentation approach, result for baseline approach, ground truth. In this specific case, part of the choroid plexus (yellow box) has been infiltrated by the tumor, which was correctly detected by our approach. .... 15
- Figure 2 . Longitudinal evolution of tumor and healthy tissue compartments shown on an axial slice of one patient. .... 16
- Figure 5. Absolute volumes and volume differences between consecutive time points as measured by BraTumIA and plotted against the estimates of Rater-1 and Rater-2. Strong significant correlations (r-values ranging from 0.83 to 0.96,  $p < 0.001$ ) were observed between all estimates of BraTumIA and of each of the raters (Meier et al. 2016). .... 16
- Figure 4 Preliminary results for uncertainty estimation in fully-connected CRFs. From left to right: Contrast-enhanced T1-weighted Magnetic Resonance Sequence. Segmentation result of the fully-connected CRF (D-CRF). Uncertainty map overlaid on the segmentation result. Manual ground truth data. Note: Best seen in color. .... 18
- Figure 7. Global interpretability of a machine learning model segmenting different tissue types. Top: each row on the top figure corresponds to a MRI sequence (T1, T1c, T2, FLAIR); each bar corresponds to the given image feature that is mostly expressed to model the information on that given MRI sequence. Features are sorted from left to right in decreasing order of importance. Bottom: examples of expressed feature maps for each MRI sequence, highlighting the targeted brain tumor area the feature information describes. .... 19
- Figure 8. Spatial local interpretability of the Challenge subject 0310 in BRATS for the task of segmenting all tissues at once (multi-label classification). From left to right, we show the T1, T1c, T2, and FLAIR sequences, as well as the obtained segmentation in the first row. In the segmentation, the tumor tissues are: blue –necrosis, green – edema, orange – non-enhancing, and red – enhancing tumor. Each row shows the input data used for predicting each class. .... 20
- Figure 9. Qualitative results of two example patients. A: Segmentation result (yellow) of BraTumIA and the different human raters. B: Segmentation result of BraTumIA (CRET) showing blood products that were incorrectly identified by the algorithm as an enhancing tumor. .... 21

**Figure 10. Prognostic performance (C-index,  $n = 108$ /AUC,  $n = 98$ ) for all tumor sub-compartments and composed volumes for the fully automatic segmentation algorithm and manual delineations. The concordance index (CI) is a generalization of the AUC, which takes into account the time to event and censoring, and provides the probability that among two randomly drawn samples, the sample with the higher risk value has also the higher chance of experiencing an event (e.g., death). The symbol \* denotes that a CI/AUC is significantly different than random (CI/AUC = 0.5) and was determined using the “noether” method as implemented in the R survcomp package. Vertical dotted lines indicate confidence intervals for the CI/AUC values. Horizontal solid lines indicate that a performance metric (CI or AUC) is significantly higher for the indicated segmentation method as compared to its counterpart ( $p < 0.05$ , p-value from the Student t test for the comparison  $\text{cindex1} > \text{cindex2}$  as implemented in the R survcomp package). ..... 22**

## Tables

**Table 1 Summary of evaluations of the BraTumIA technology.** CE=contrast-enhancing; nCE: non-contrast enhancing; TV=Total volume (necrotic+CE+non-contrast CE); TV+=TV+edema, range=max-min..... 17



## 1 Executive Summary

The CHIC project aims at developing cutting edge ICT tools, services and secure infrastructure to foster the development of elaborate and reusable integrative models (hypermodels) in the field of cancer diagnosis and treatment, as well as larger repositories so as to demonstrate benefits of having both the multiscale data and the corresponding models readily available in the VPH domain.

In WP9, and specifically in Task 9.7, we developed novel technologies for automated longitudinal brain tumor image analysis, which are pivotal to assess tumor burden and response to treatment. The proposed technologies were designed and validated with an emphasis on technical (i.e. in-silico modelling) and clinical exploitation, as well as to provide robust tools for the emerging field of radiomics.

The presented document starts with a summary of the state-of-the-art in brain tumor segmentation as well as a short description of the current paradigms and challenges in tumor response to therapy, which motivates the development of automated tools for brain tumor analysis. Subsequent sections focus on the main findings and results obtained for the technical developments and clinical evaluations of longitudinal brain tumor segmentation. A section on comparisons with other research and industry-oriented solutions is also presented, highlighting the features and attained performance levels of the sought clinically-relevant solution. Towards the end of the document, applications to neurosurgery and radiomics, as well as preliminary findings on estimating segmentation uncertainty and enhancing interpretability of machine learning models, are presented.

## Introduction

Automated brain tumor segmentation methods are computational algorithms that yield a 3D delineation of tumor compartments from multimodal Magnetic Resonance Imaging (MRI). Their automatic nature provides important time-saving for neuro-radiologists, and compared to the standard care of GBM patients, their accuracy enables an improved tumor characterization for radiotherapy, surgical planning, drug development assessment, etc.

As described in the recent work of (Menze et al., 2015), summarizing consecutive years of the brain tumor segmentation challenges, brain tumor segmentation based on supervised learning provides tumor volumetry accuracies in the range of the inter-observer variability, which suggests that automated volumetry has potential for clinical use. However, clinical adoption of these technologies is limited due to lack of efforts to transfer this technology into dedicated and certified software tools that can be deployed in the clinical routine.

The aim of this document is to summarize findings and developments performed during the CHIC project for automated brain tumor image analysis, with focus on the longitudinal aspects of the analysis and clinical applications.

### ***1.1 Current status in the treatment of glioblastoma***

Glioblastomas are the most common and lethal primary brain tumors in humans and among the deadliest type of cancer. They exhibit a very rapid growth, are infiltrative and spread mainly along white fiber tracts, (Giese, 2003). Average survival time for glioblastoma patients is 14 months (Adamson et al., 2009). Treatment of glioblastomas includes surgery, radiation therapy, chemotherapy or a combination of these. Magnetic resonance imaging is the standard imaging technique in brain tumor diagnosis. In current clinical practice, four image sequence protocols (i.e. T1-weighted, T1-weighted with contrast agent, T2-weighted, T2-weighted with fluid attenuated inversion recovery) are commonly used to characterize sub-compartments of the tumor.

International efforts have been conducted to define advanced tumor progression criteria (Chinot et al., 2013; Wen et al., 2010). However, although volumetric tumor information is known to be important for tumor progression assessment in neuro-oncology, the lack of appropriate clinically relevant technology to quantify tumor volumetry has led these efforts to employ over-simplistic two-dimensional metrics. Such metrics have shown several limitations, such as increased measurement variability, sensitivity to imaging quality, and difficulties to assess irregularly shaped lesions, (Reuter et al., 2014; Schmitt, Mandonnet, Perdreau, & Angelini, 2013; Sorensen et al., 2001). Recently, new consensus recommendations have been produced highlighting the importance of volumetric tumor analyses (Ellingson et al., 2015). Furthermore, with the increased awareness from the scientific community on the importance of imaging biomarkers for treatment assessment and prognosis (Nicolaidis et al. 2015; Aerts et al., 2014; Bai et al., 2016; Cui et al., 2016; Gillies et al., 2016; O'Connor et al., 2015), it is of pivotal importance to develop technologies supporting and leveraging the understanding and treatment strategies of this lethal disease.

### ***1.2 State of the art in brain tumor image segmentation***

Here we provide a succinct summary of the state of the art in brain tumor image segmentation to help the reader situate the developments and achievements in light of the current state of the art.

Machine learning techniques have gained much popularity in medical image computing and brain tumor segmentation. These approaches rely on models that can be broadly classified as supervised, semi-supervised, or unsupervised learning methods, (Alpaydin, 2014; Jordan & Mitchell, 2015).

Unsupervised brain tumor segmentation methods deal with the problem of segmenting tumoral tissues through data clustering and similarity metrics. They do not need labeled data and therefore

are scalable and flexible to changes to the imaging sequence. (Juan-Albarracín et al., 2015; Prastawa & Gerig et al., 2003; Zhang et al., 2001) where clustering techniques (K-means, Expectation Maximization, Gaussian Mixture Models, Self-Organizing Maps, etc.) have been combined with shape priors (e.g. brain atlas) and post-processing steps to regularize the final segmentation. The main difficulty of unsupervised learning is associated with problems defining flexible but accurate priors, and modelling assumptions, which are confronted with the high complexity and partial knowledge of the disease and patient variability.

Semi-supervised methods aim at exploiting information from unlabeled and labeled data. The most active area of research in semi-supervised learning has been in graph-based methods (Chapelle et al. 2010; Duchenne et al. 2008; Hamamci et al. 2012; Kanas et al. 2015; Meier et al. 2014; Orbach et al. 2012; Wang et al. 2008). These methods rely on representing the image data as a graph, where nodes correspond to image voxels or high-dimensional data points, and edges connecting nodes represent the similarity among them. In graph-based methods the prediction usually corresponds to a propagation of the labels from labeled to unlabeled nodes, which is referred to as transductive inference (Vladimir, 2006). As opposed to inductive inference, where a hypothesis is inferred from training data, in transductive inference the value of the hypothesis itself (prediction function) for the available testing data is inferred (Chapelle et al., 2010).

Most modern brain tumor segmentation approaches use supervised learning methods to capture the relationship between imaging features and the classification of tumoral tissues. In supervised learning, the mapping (hypothesis building) between imaging features and target label variables is learned during an offline training phase. In this phase, images pre-annotated by an expert are used in conjunction with a set of selected or learned imaging features (i.e. feature descriptors) to derive a multidimensional classification model (Criminisi & Shotton, 2013).

The literature in supervised learning is vast. However, in medical image computing some approaches have been more extensively researched. In particular, support vector machine (SVM) (Orrù et al. 2012; Schölkopf & Smola, 2002), Random Forests (Breiman, 2001; Criminisi & Shotton, 2013), and lately, Deep learning c.f., (LeCun, Bengio, & Hinton, 2015; Schmidhuber, 2014). Supervised learning based on Random Forests (Breiman, 2001; Criminisi & Shotton, 2013) has been one of the most explored machine-learning technique for medical image computing, and in recent years, deep learning based approaches have gained popularity (Pereira et al. 2016; Urban et al. 2014).

Interestingly, a driving force for benchmarking and algorithmic development in the medical image computing community is the creation of grand challenges. These are dedicated open benchmarking competitions for best algorithmic accuracy, in which machine learning techniques have proved to be successful, e.g. (Bauer et al. 2011; Urban et al. 2014; Menze et al., 2015; Menze et al., 2015b). Similarly, dedicated workshops for machine learning and use of “Big Data”, e.g., (Bauer et al. 2013) are now well-established at the top conference venue of the field, Medical Image Computing and Computer Assisted Interventions (Miccai).

As described in the recent work of (Menze et al., 2015), summarizing consecutive years of the brain tumor segmentation challenges, brain tumor segmentation based on supervised learning provides tumor volumetry accuracies in the range of the inter-observer variability, which suggests that automated volumetry has potential for clinical use. However, it is observed from the literature that the robustness of the algorithms needs improvement in order to enhance the spatial consistency of brain tumor segmentation results, which is crucial for neuro-surgery, radio-therapy, radiomics, etc. where the spatial extent of the tumor compartments play an important role for the subsequent decision-making processes and analyses. Similarly, as mentioned in the current state of the art in glioblastoma treatment, such improvements are also essential in high-throughput studies supporting the understanding of evolutionary patterns of tumors (Bai et al., 2016; Cui et al., 2016; Sottoriva et al., 2013)(Gillies et al. 2016).

As pointed out in the recent work of (Weizman et al., 2014), it is important to analyze not only volumetric changes but also changes happening at the cellularity level and among detected tumor tissue compartments, e.g. (Galbán et al., 2009, 2011). While several approaches for brain tumor segmentation of single time points have been investigated, longitudinal brain tumor segmentation has been seldom investigated.

In (Weizman et al., 2014) longitudinal brain tumor segmentation of low-grade gliomas (LGG) was proposed. The fuzziness of tumor borders (i.e. lack of sharp edges) between consecutive scans is analyzed in order to detect areas where the tumor is changing. The user first needs to delineate the tumor and its components on the baseline image. Next, unsupervised tumor classification (GMM-based clustering) coupled to a novel tumor tissue transition model enables updating the tumor component delineations on the follow-up scan.

In (Alberts et al., 2016) longitudinal segmentation is performed through a non-parametric CRF-based tumor growth model. Single-time segmentations are unified into a final segmentation result, via graph-cut solving and by using growth and inclusion constraints modelled as infinite or zero-valued temporal graph links, which are derived from defined or learned tumor growth patterns and appearance of tumor tissues across modalities (e.g. a necrotic voxel is a subset of the tumor hyper-intensity, as defined on the FLAIR image).

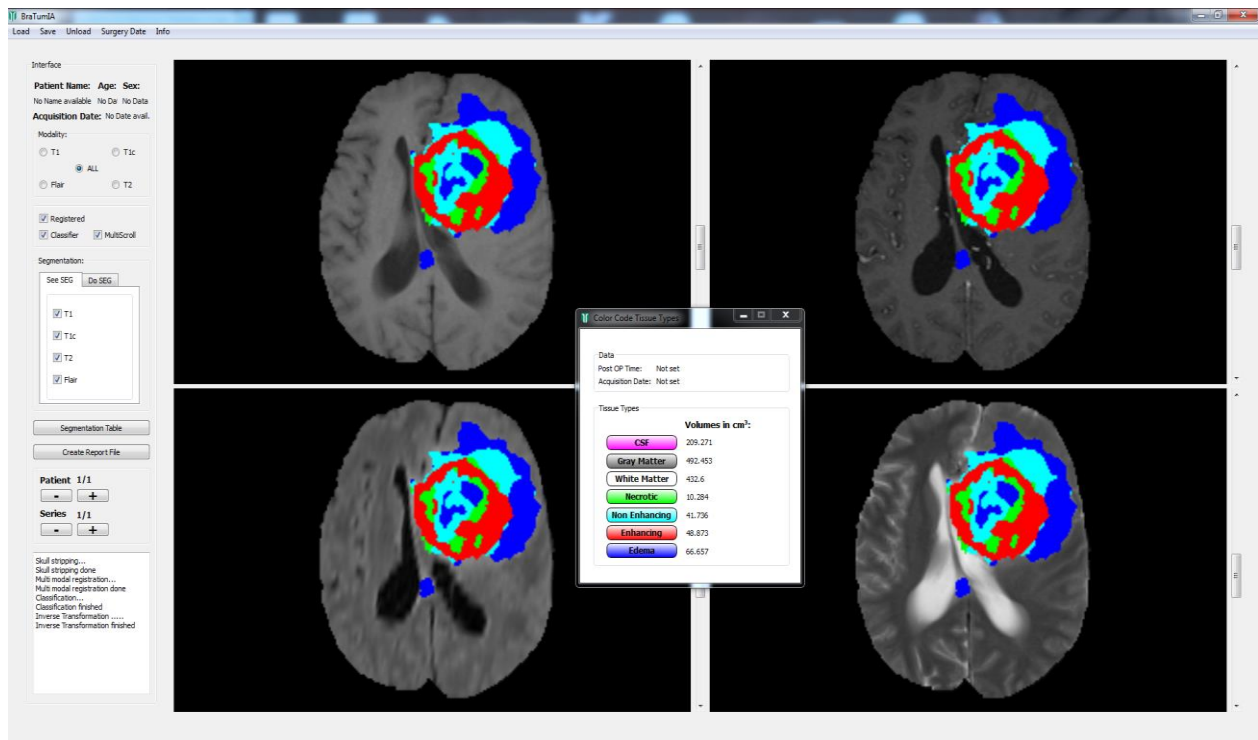
The following sections describe the developments performed during the CHIC project.

## 2 Automated Brain Tumor Segmentation

For the analysis of brain cancer, it is essential to delineate brain tumor boundaries from magnetic resonance images. In (Bauer et al., 2011) we originally presented a fully automatic technique to segment multi-modal images of grade IV patients into healthy and pathologic tissues. These two regions are further sub-classified into white matter (WM), gray matter (GM), cerebrospinal fluid (CSF) and necrotic region, active region, and edema region. The technology was then further developed in (Bauer, Porz, et al., 2014; Raphael Meier, Bauer, Slotboom, Wiest, & Reyes, 2014a; Porz et al., 2014), and has achieved good results on clinically relevant data of glioblastoma patients with a fast computation time, which makes it useful for clinical practice. Further developments of the software led us to accuracy and speed improvements, which were reflected in the results obtained at the Brain Tumor Segmentation (BRATS) Miccai Challenge, where the approach ranked in subsequent years **among the top-three approaches** (Menze et al., 2015).

The software tool, called **BraTumIA** (Brain Tumor Image Analysis), is a standalone program that performs fully automatic **brain tumor tissue segmentation** in approximately 5 minutes. In addition, the tool **also segments healthy tissues** (CSF, white matter, gray matter) and **subcortical structures**, which is useful for neurosurgery and radiotherapy purposes.

It is worth noting that our approach has reached **clinical evaluation and adoption**. *BraTumIA* has been made available at the NITRC portal (<http://www.nitrc.org/projects/bratumia/>) and has been downloaded since its release last May 2014 over 400 times.



**Figure 1** Main interface of BraTumIA (Brain Tumor Image Analysis). Four views enable visualization of the four MR sequences. After segmentation, the segmentation can be overlaid on the input images where different colors are used to differentiate the tumor layers of necrotic, enhancing core, non-enhancing core, and edema. Volumes per tumor layer can also be retrieved and exported as a text file for further analysis. The segmentation results are also available on the user's local folder for further analysis.

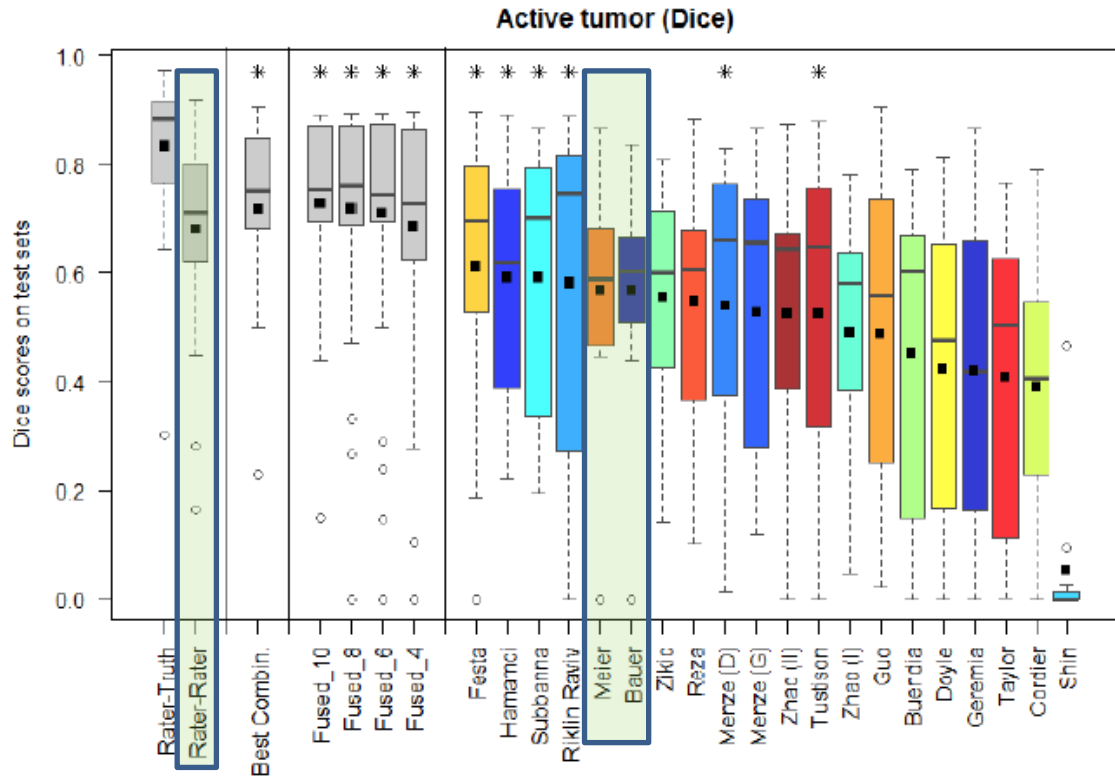
### 3 Clinical Evaluation of Automated Brain Tumor Image Segmentation

BraTumIA was evaluated clinically through a prospective study including 25 glioblastomas cases (Porz et al., 2014). Two independent expert raters performed manual segmentations, which were compared against the ones produced by BraTumIA. For analysis, the tumor was divided into three regions, namely complete tumor volume (TV) (enhancing part plus non-enhancing part plus necrotic core of the tumor); the TV+ (TV plus edema) and the contrast enhancing tumor volume (CETV).

We quantified the overlap between manual and automated segmentation by calculation of diameter measurements as well as popular image segmentation quality metrics such as dice coefficients, positive predictive values, sensitivity, relative volume error and absolute volume error. The results indicated no differences in the measured cross-sectional sum of product of diameters between automated and manual measurements. Secondly, the study showed that the estimation of TV, TV+ and CETV by the automatic method reaches a sensitivity that is comparable to the inter-observer variability, which suggests that BraTumIA yields a good accuracy for being used in the clinics, moreover the objectiveness (specially in regards to longitudinal tumor volumetry) of BraTumIA is an added value to the clinical workflow.

## 4 Comparison to other state-of-the art methods

Since 2012 the Brain Tumor Segmentation challenge hosts a multicenter dataset of GBM cases, and enables a direct comparison of automated tumor segmentation methods. In this competition BraTumIA ranked top 2 during consecutive years in 2012 and 2013 and among the top methods since then.



**Figure 2** Dispersion of Dice scores for the individual algorithms (color coded), and various fused algorithmic segmentations (gray), shown together with the expert results (also shown in gray). Boxplots show quartile ranges of the scores on the test datasets; whiskers and dots indicate outliers. Black squares indicate the mean score. Also shown are results from four “Fused” algorithmic segmentations, and the performance of the “Best Combination” as the upper limit of individual algorithmic performance. The light-grayed rectangle highlights results from BraTumIA. In comparison to other approaches it is highlighted the low variance (higher robustness) of BraTumIA, in the range of the inter-expert variability (also highlighted with a light-grayed box).

### 4.1 Evaluation of BraTumIA and BrainLab’s FDA-approved SmartBrush segmentation

BrainLab has developed a semi-automated tool for tumor delineation, called SmartBrush, mostly used by neurosurgeons. SmartBrush relies on brush strokes provided by the user to segment the tumor core; it is a fast and robust approach; however, it does not segment the complete set of tumor compartments used in clinical practice to monitor tumors. In this study we are comparing both tools and evaluating the strengths and weaknesses. The study, recently published in (Porz et al., 2016), highlights the added value of BraTumIA’s ability to segment non-contrast enhancing, as well as user-related biases yielding different tumor segmentation results, depending on user’s experience.

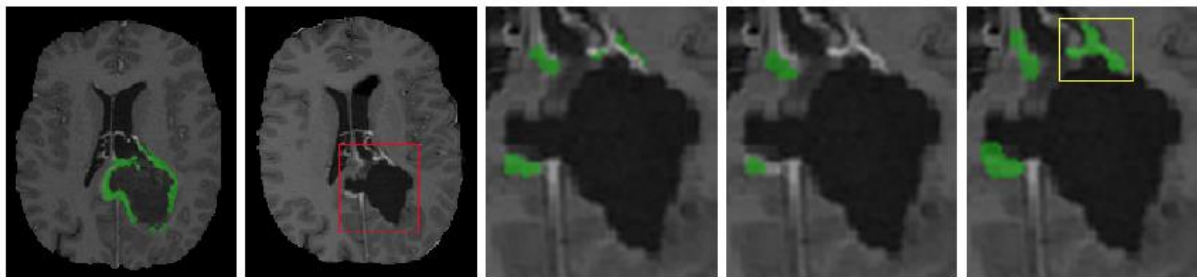


## 4.2 Summary of BraTumIA features

1. Fully automated
2. Based on Multi-sequence MRI – T1, T1c, T2, Flair
3. Segmentation in ca. 5 minutes.
4. Tumor compartmentalization: necrotic core, contrast-enhancing, non-contrast enhancing, edema.
5. Longitudinal segmentation (i.e. pre, and post-operative imaging)
6. Segmentation of white-matter, gray-matter, CSF and sub-cortical structures.
7. Exporting of label images and report of volumes for each tumor compartment in mm<sup>3</sup>.

## 5 Automated segmentation of post-operative MRI patient images.

We developed an automatic segmentation approach to segment post-operative images. The approach utilizes the pre-operative image as additional information to segment the post-operative image, and hence to compute the tumor residual and the extent of resection. This enabled us to provide detailed volumetric information about the pre- and post-operative total tumor volume and its follow-up dynamics (Meier et al. 2014). This approach corresponds to another building block to achieve a complete longitudinal brain tumor segmentation and analysis framework.



**Figure 3 Exemplary case, T1c -weighted images (green = enhancing tumor), from left to right: Preoperative image, Postoperative image (relevant area is magnified), result for the proposed semi-supervised post-operative segmentation approach, result for baseline approach, ground truth. In this specific case, part of the choroid plexus (yellow box) has been infiltrated by the tumor, which was correctly detected by our approach.**

## 6 Longitudinal Brain Tumor Segmentation

In this section two main results on longitudinal brain tumor image analysis are presented, where spatio-temporal regularization and dedicated longitudinal analyses of tumor progression/response to therapy were developed.

### 6.1 Longitudinal Brain Tumor Image Analysis of High Grade Gliomas

We developed an automatic approach to analyze longitudinal images of glioma patients. It considers four-dimensional (i.e. 3D+time) spatio-temporal information of each patient through a CRF extended to consider contiguous time points. The results showed significant improvements over isolated 3D segmentations. The outcome was presented in a conference paper (Bauer, Porz, et al., 2014) for which we were awarded a **Siemens award for best paper at the MICCAI workshop on Medical Computer Vision**. The development of longitudinal tumor segmentation has opened the possibility to monitor disease progression in an automated manner, opening possibilities to leverage treatment assessment. An example result is shown below.

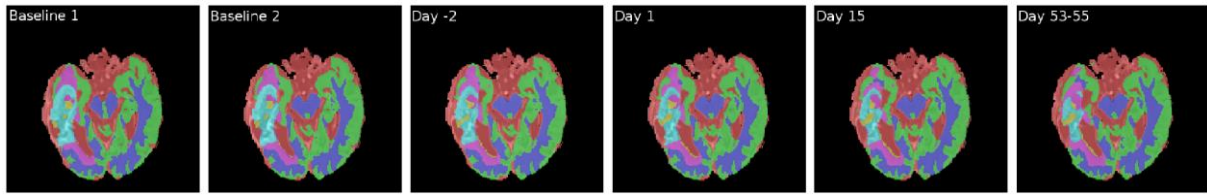


Figure 4 . Longitudinal evolution of tumor and healthy tissue compartments shown on an axial slice of one patient.

## 6.2 Clinical Evaluation of a Fully-automatic Segmentation Method for Longitudinal Brain Tumor Volumetry

We conducted a second study at our local hospital, for which we used our fully automatic segmentation method to segment longitudinal imaging studies. In a **prospective setting, a cohort of 14 patients with a total of 64 MR acquisitions, including preoperative, immediate postoperative and up to 12-month follow-up images was used.** Two expert raters segmented each case manually and BraTumIA performed the segmentation automatically. We compared the volumetric trends for contrast-enhancing tumor and for the non-contrast-enhancing, T2-hyperintense tissue (edema+non-enhancing tumor).

We found that the **volumetric trends as estimated by BraTumIA are comparable to the volumetric trends as estimated by the human raters**, which enables the use of BraTumIA for automated longitudinal tumor volumetry analysis tool (Meier et al., 2016). In addition, this study provides insights into the critical question of defining and quantifying volumetric trends, which are essential for the creation of new therapy assessment metrics that consider tumor multi-compartment volumetric changes.

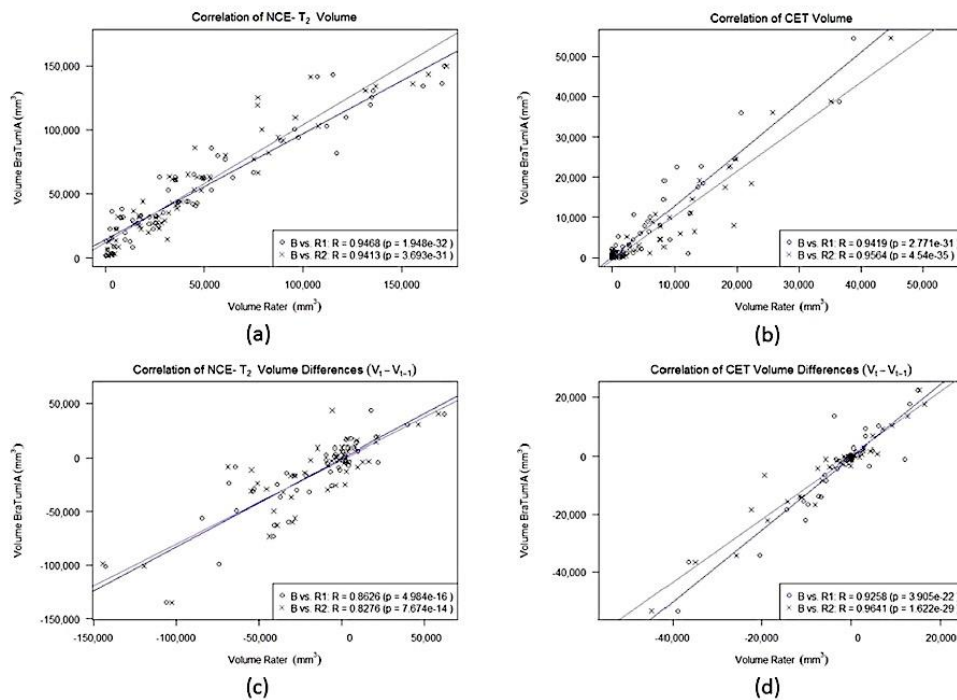


Figure 5. Absolute volumes and volume differences between consecutive time points as measured by BraTumIA and plotted against the estimates of Rater-1 and Rater-2. Strong significant correlations ( $r$ -values ranging from 0.83 to 0.96,  $p < 0.001$ ) were observed between all estimates of BraTumIA and of each of the raters (Meier et al. 2016).



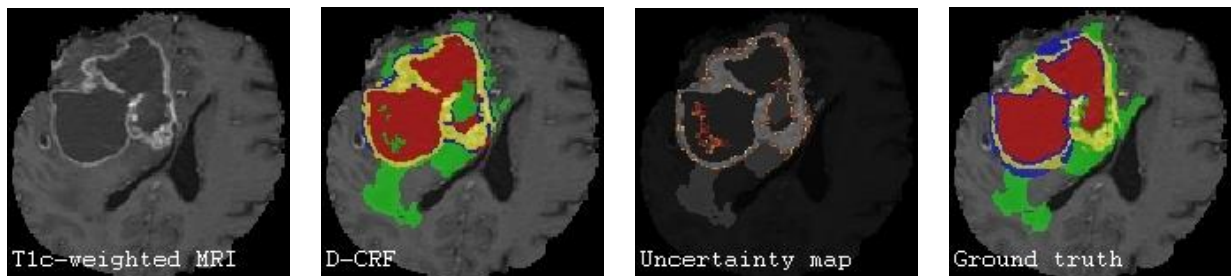
**Table 1 Summary of evaluations of the BraTumIA technology.** CE=contrast-enhancing; nCE: non-contrast enhancing T2-hyperintense tissue; TV=Total volume (necrotic+CE+non- enhancing tumor); TV+=TV+edema, range=max-min.

Study	Eligibility	Reported performance - Dice	Multicenter (M)/ Single-center (S)
<b>Porz et al. 2014</b>	N=25 Two 1.5T MRI (Siemens)  Pre-operative. Patients with newly diagnosed and histologically confirmed glioblastoma. Exclusion criteria were: incomplete image acquisition, Karnofsky performance status <70%, abnormal hematologic, renal or hepatic function, and previous cranial neurosurgery	Average Dice TV+= 0.8 Average Dice TV= 0.66 Average Dice CE= 0.63 Average absolute volume error: 20.4 ml for TV+, 14.5 ml for TV and 7.2 ml for CETV	M+S: BRATS2012+Inselspital
<b>Menze et al. 2015</b>	N= 65 Several MRI, combined 1.5T+3.0T.  Preoperative, low grade (astrocytomas or oligoastrocytomas) and high grade (anaplastic astrocytomas and glioblastoma multiforme tumors).	Average Dice CE: 0.57 Average Dice TV+: 0.69	M: Debrecen; Bern, Heidelberg, Massachusetts.
<b>Meier&amp;Knecht et al. 2016</b>	N=64 One 1.5 T MRI (Siemens)  Longitudinal imaging (pre, post and up to six time points.	Dice CE (median, range): (0.661, 0.647).  CE - Correlation to human: 0.9419.  Dice nCE (median, range): (0.746, 0.319)  nCE- Correlation to expert: 0.9468.	S: Inselspital
<b>Meier et al. BRATS Workshop 2016</b>	N=110 Several MRI, combined 1.5T+3.0T. TCIA imaging datasets. High-grade cases.	Dice CE (median, range): (0.66, 0.843)  Dice TV (median, range): (0.746, 0.902)  Dice TV+ (median, range): (0.789, 0.487)	M: TCIA multicenter.

## 7 Uncertainty Estimation in Brain Tumor Image Segmentation

We have obtained first proof-of-concept results on deriving segmentation uncertainty estimations from fully-connected CRFs. The key idea builds on an extension of the recently proposed work of (Papandreou & Yuille, 2011) where an approach called “Perturb-and-MAP” was proposed. Perturb-and-MAP allows one to efficiently sample from the underlying Gibbs distribution of a CRF through perturbing the initial energy function followed by solving the MAP estimation for the perturbed energy. The perturbation corresponds to the addition of IID random noise to the potential functions. If the perturbation density follows a Gumbel distribution and under the condition that all potentials are perturbed, it can be shown that the Perturb-and-MAP model approximates the Gibbs distribution of the corresponding random field. Our preliminary submitted work related fully-connected CRFs with Perturb-and-MAP. For fully-connected CRFs, exact MAP estimation is not feasible anymore due to the size of the graph representing the medical image volume. As indicated, efficient approximate inference in fully-connected CRFs is possible via a recently proposed method (Krähenbühl & Koltun, 2011), which is based on a mean field approximation. We adopted the Perturb-and-MAP framework to this setting of approximate MAP inference, which we termed Perturb-and-MPM.

We obtained initial results indicating that Perturb-and-MPM can yield marginals closer to the true marginal probabilities than the marginals obtained by the mean field approximation itself. Candidate segmentations can be sampled by first perturbing the original energy function and then solving the corresponding MPM problem for  $\tilde{X}_{MPM}$ . These candidate segmentations can then be used to quantify segmentation uncertainty. For every voxel, a class label histogram is generated containing the label assignments in all sampled candidate segmentations. The segmentation uncertainty corresponds to the Shannon entropy in this histogram. An exemplary uncertainty map is shown below.



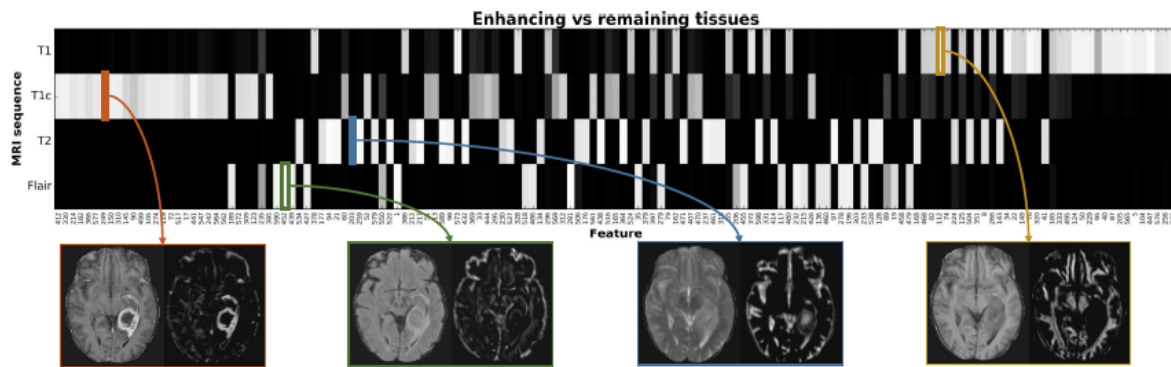
**Figure 6 Preliminary results for uncertainty estimation in fully-connected CRFs. From left to right: Contrast-enhanced T1-weighted Magnetic Resonance Sequence. Segmentation result of the fully-connected CRF (D-CRF). Uncertainty map overlaid on the segmentation result. Manual ground truth data. Note: Best seen in color.**

## 8 Enhancing interpretability of Machine Learning Models: Application to the Automated Segmentation of Brain Tumor Images

The proposed approaches rely on a supervised machine learning paradigm, in which a model is built from training datasets, and then applied on newly unseen testing cases. While much research has been dedicated to increase the accuracy and performance of these technologies, not many efforts have been given to the interpretability of the resulting models created by the training phase of supervised learning approaches. Being able to understand a given model’s behavior (for both correct

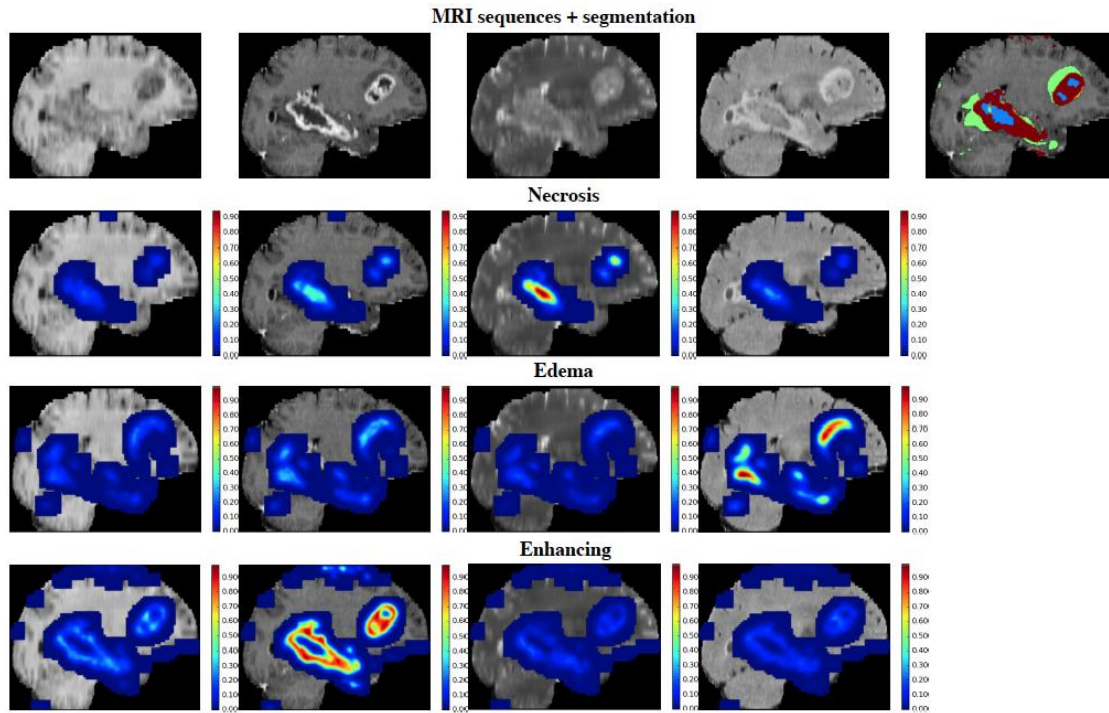
and incorrect results) is key in order to ensure that the model has correctly utilized the underlying imaging data to infer information about the disease. Furthermore, being able to interpret a model's result is also key to optimize imaging sequences and algorithmic components needed to ensure robustness of the technical solutions.

During CHIC we performed a first study (to be submitted) to enhance interpretability of machine learning models. Specifically, we developed a strategy that jointly models data (e.g. multi-sequence MRI) and objective task (e.g. labeling a brain tumor) and measures the interrelations within, allowing a human expert to elucidate what information a given model uses to perform inference (e.g. which sequence and which region in the sequence). Figure 7 shows how the proposed approach is capable of highlighting MRI sequences used by the model to characterize different areas of the tumor, in the same way as a radiologist does in practice. Such analysis allows validation of the resulting model constructed from training datasets. In this specific case, the model follows a similar data analysis scheme as a radiologist would do to segment the contrast-enhancing using the T1c image, followed by the T2 and FLAIR images, and finally the T1 image.



**Figure 7. Global interpretability of a machine learning model segmenting different tissue types.** Top: each row on the top figure corresponds to a MRI sequence (T1, T1c, T2, FLAIR); each bar corresponds to the given image feature that is mostly expressed to model the information on that given MRI sequence. Features are sorted from left to right in decreasing order of importance. Bottom: examples of expressed feature maps for each MRI sequence, highlighting the targeted brain tumor area the feature information describes.

The proposed approach has a second layer of local interpretability that enables us to check at a patient-specific level, what information is being used by the model for each segmented voxel, see Figure 8. In this specific example in Figure 8, it can be seen that segmentation of the enhancing and necrotic tissues is driven mostly by the T1c image, while segmentation of the edema is mostly driven by the FLAIR sequence. In this example one can also observe that outliers, stemming from an imperfect skull stripping operation, are mostly influenced by T1c voxels. This information can then enable us to locate sources of errors and design correction mechanisms.



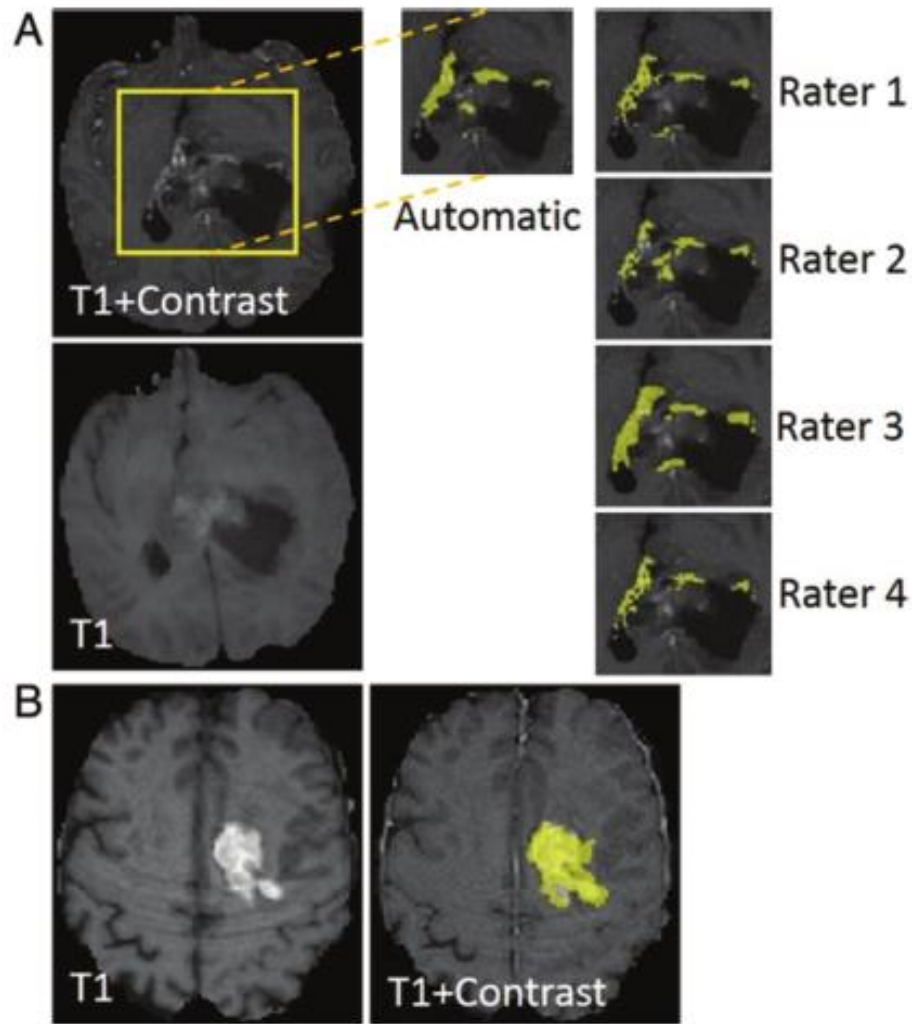
**Figure 8. Spatial local interpretability of the Challenge subject 0310 in BRATS for the task of segmenting all tissues at once (multi-label classification).** From left to right, we show the T1, T1c, T2, and FLAIR sequences, as well as the obtained segmentation in the first row. In the segmentation, the tumor tissues are: blue –necrosis, green – edema, orange – non-enhancing, and red – enhancing tumor. Each row shows the input data used for predicting each class.

In summary, the proposed approach aims at leveraging interpretability of machine learning models, needed specially in clinical settings where clinical experts need to rely and to trust a given technology for their decision-making process. These approaches are fundamental as well to ensure that their inferencing follows correct patterns extracted from the data, as well as to facilitate their improvement over time.

## 9 Applications of the Developed Technology

### 9.1 Assessment of extent-of-resection (EOR) and Residual-Tumor (RT) through automated tumor volumetry

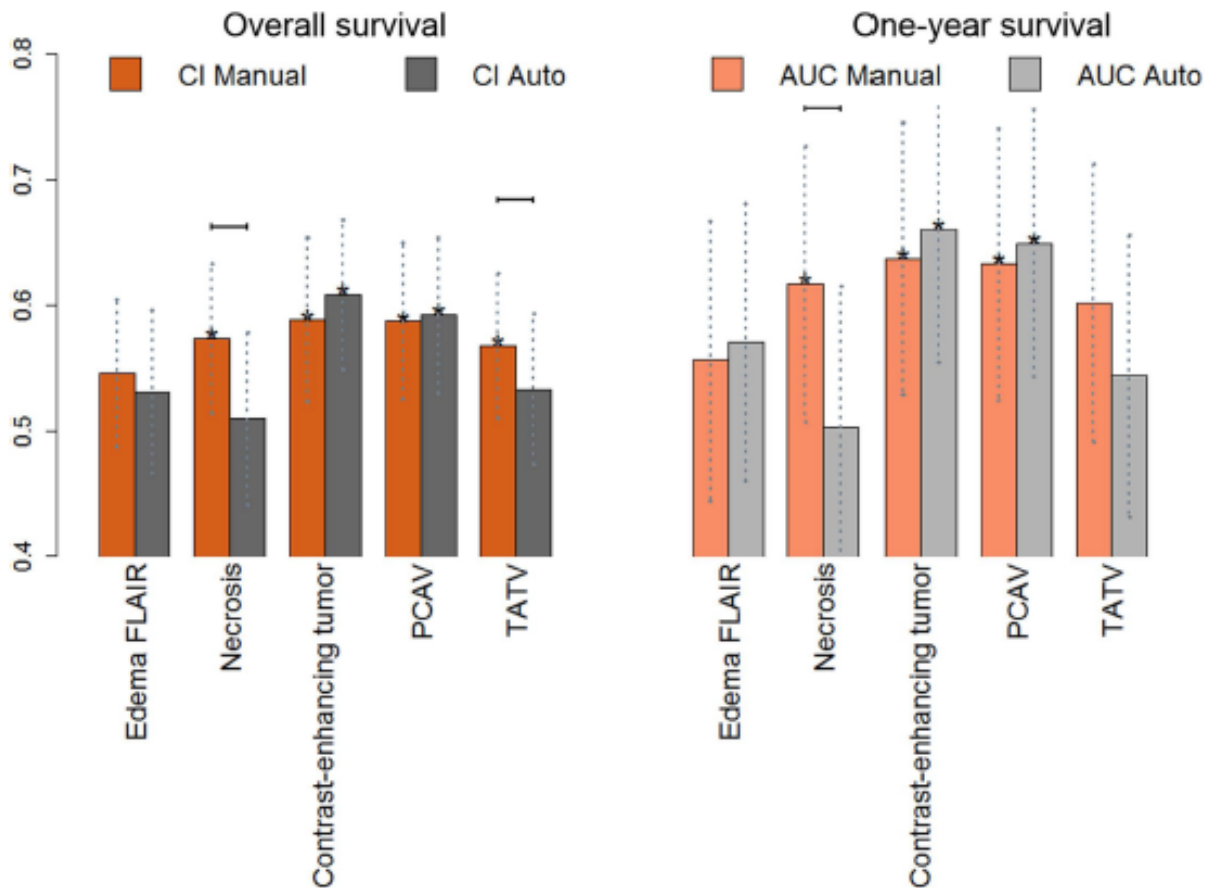
EOR and RT are currently used to assess the outcome of surgery using manual assessment on multimodal MRI. In this study we evaluated the ability of BraTumIA to quantify EOR and RT in a fully automated fashion. The main findings of the study, recently published in (Meier et al., 2017) include an agreement with other reports, supporting the use of RT over EOR (which depends on the pre-operative tumor) as well as the ability of BraTumIA to quantify RT with a similar degree of accuracy as the human expert. An exemplary case is shown in Figure 9, where the automated post-operative segmentation results are compared to four other expert raters.



**Figure 9. Qualitative results of two example patients. A: Segmentation result (yellow) of BraTumIA and the different human raters. B: Segmentation result of BraTumIA (CRET) showing blood products that were incorrectly identified by the algorithm as an enhancing tumor.**

## 9.2 Automated Volumetry for Radiomics

We performed a study together with partners at the Dana Farber Cancer Institute in Boston, in which we compared the association of tumor sub-compartment volumes (necrosis, contrast-enhancing tumor, edema) - generated by our segmentation software or by manual segmentations - with patient survival. In an analysis on 109 GBM cases from The Cancer Imaging Archive (TCIA-<http://www.cancerimagingarchive.net>)



**Figure 10. Prognostic performance (C-index,  $n = 108$ /AUC,  $n = 98$ ) for all tumor sub-compartments and composed volumes for the fully automatic segmentation algorithm and manual delineations. The concordance index (CI) is a generalization of the AUC, which takes into account the time to event and censoring, and provides the probability that among two randomly drawn samples, the sample with the higher risk value has also the higher chance of experiencing an event (e.g., death). The symbol \* denotes that a CI/AUC is significantly different than random (CI/AUC = 0.5) and was determined using the “noether” method as implemented in the R survcomp package. Vertical dotted lines indicate confidence intervals for the CI/AUC values. Horizontal solid lines indicate that a performance metric (CI or AUC) is significantly higher for the indicated segmentation method as compared to its counterpart ( $p < 0.05$ , p-value from the Student t test for the comparison  $\text{index1} > \text{index2}$  as implemented in the R survcomp package).**

We found that in case of the contrast-enhancing tumor and edema, the automatically generated volumes are equally prognostic as the manual ones, suggesting their use as a prognostic biomarker, see Figure 10. The findings from this work support the use of automated tools, needed to overcome the heavy load of manual tumor volumetry, and to effectively leverage the understanding of glioblastomas and design new prognostic biomarkers in high- throughput data analyses (Rios Velazquez et al., 2015).



## 10 Conclusions and Outlook

In summary, we have developed automated technologies to quantify and analyze the burden of brain tumors in patient suffering from high-grade gliomas. The technology is fully automated, permitting its use for high-throughput studies, such as clinical trials or large studies where automated pipelines are sought. The technology has been designed, tested and validated in conjunction with medical experts, enabling its deployment and exploitation in the clinical setting.

Along the course of the research we have identified challenges and opportunities that constitute the basis for future developments and utilization of the technology. First, robustness of the technology is a critical asset that such technologies need to feature. The proposed technology features high robustness, which is essential in longitudinal studies where response to therapy or disease progression is assessed in a relative manner with respect to a reference time point (i.e. nadir).

Radiomics is an emerging field where imaging and non-imaging patient information is combined with machine learning techniques to derive novel non-invasive imaging biomarkers that can be used for assessing patient survival (prognosis task), tumor phenotyping, association to genetic biomarkers (radiogenomics), etc. The technology developed in CHIC for automated longitudinal brain tumor image analysis brings technological solutions to foster the field of radiomics.

Furthermore, these developments open opportunities to further improve the current response to treatment assessment criteria (RANO) that relies on over-simplistic metrics of tumor size and disease progression. The combination of these technologies with advanced Radiomics, employing machine learning approaches, have then the potential to refine and personalize the analysis of patient imaging information, as opposed to the current population-level analyses. Nonetheless, we remark further research is needed to develop solutions to circumvent current issues in pseudo-response and pseudo-progression of tumors (Yang, 2015).

The proposed uncertainty segmentation estimation technology is an important advancement capable of acting as a data quality checker that can be used to eliminate spurious or faulty segmentation results that can mislead the subsequent radiomics analysis, surgical planning, assessment of surgical success (in the case of post-operative tumor segmentation), etc.

## References

- Adamson, C., Kanu, O. O., Mehta, A. I., Di, C., Lin, N., Mattox, A. K., & Bigner, D. D. (2009). Glioblastoma multiforme: a review of where we have been and where we are going. *Expert Opinion on Investigational Drugs*. Retrieved from <http://www.tandfonline.com/doi/abs/10.1517/13543780903052764>
- Aerts, H. J. W. L., Velazquez, E. R., Leijenaar, R. T. H., Parmar, C., Grossmann, P., Cavalho, S., ... Lambin, P. (2014). Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nature Communications*, 5, 4006. <http://doi.org/10.1038/ncomms5006>
- Alberts, E., Charpiat, G., Tarabalka, Y., Huber, T., Weber, M.-A., Bauer, J., ... Menze, B. H. (2016). A Nonparametric model for Brain Tumor Segmentation and Volumetry in Longitudinal MR Sequences. In A. Crimi, B. Menze, O. Maier, M. Reyes, & H. Handels (Eds.), (pp. 69–79). Cham: Springer International Publishing. [http://doi.org/10.1007/978-3-319-30858-6\\_7](http://doi.org/10.1007/978-3-319-30858-6_7)
- Alpaydin, E. (2014). *Introduction to Machine Learning*. MIT Press. Retrieved from <https://books.google.com/books?hl=en&lr=&id=NP5bBAAAQBAJ&pgis=1>
- Bai, H. X., Lee, A. M., Yang, L., Zhang, P., Davatzikos, C., Maris, J. M., & Diskin, S. J. (2016). Imaging genomics in cancer research: limitations and promises. *The British Journal of Radiology*, 20151030. <http://doi.org/10.1259/bjr.20151030>
- Bauer, S., Nolte, L.-P., & Reyes, M. (2011). Fully automatic segmentation of brain tumor images using support vector machine classification in combination with hierarchical conditional random field regularization. *Medical Image Computing and Computer-Assisted Intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention*, 14(Pt 3), 354–61. Retrieved from

<http://www.ncbi.nlm.nih.gov/pubmed/22003719>

- Bauer, S., Porz, N., Meier, R., Pica, A., Slotboom, J., Wiest, R., & Reyes, M. (2014). Interactive Segmentation of MR Images from Brain Tumor Patients. In *ISBI 2014: Proceedings of the 2014 IEEE international conference on Biomedical imaging* (pp. 862–865). Beijing.
- Bauer, S., Tessier, J., Krieter, O., Nolte, L.-P., & Reyes, M. (2014). Integrated spatio-temporal segmentation of longitudinal brain tumor imaging studies. In *Medical Computer Vision. Large Data in Medical Imaging* (pp. 74–83). Springer.
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5–32. <http://doi.org/10.1023/A:1010933404324>
- Chapelle, O., Schölkopf, B., & Zien, A. (2010). *Semi-Supervised Learning* (1st ed.). The MIT Press.
- Chinot, O. L., Macdonald, D. R., Abrey, L. E., Zahlmann, G., Kerloëguen, Y., & Cloughesy, T. F. (2013). Response assessment criteria for glioblastoma: practical adaptation and implementation in clinical trials of antiangiogenic therapy. *Current Neurology and Neuroscience Reports*, 13(5), 347. <http://doi.org/10.1007/s11910-013-0347-2>
- Criminisi, A., & Shotton, J. (2013). *Decision Forests for Computer Vision and Medical Image Analysis*. Springer Science & Business Media. Retrieved from <https://books.google.com/books?hl=en&lr=&id=nUhDAAAQBAJ&pgis=1>
- Cui, Y., Tha, K. K., Terasaka, S., Yamaguchi, S., Wang, J., Kudo, K., ... Li, R. (2016). Prognostic Imaging Biomarkers in Glioblastoma: Development and Independent Validation on the Basis of Multiregion and Quantitative Analysis of MR Images. *Radiology*, 278(2), 546–53. <http://doi.org/10.1148/radiol.2015150358>
- Duchenne, O., Audibert, J.-Y., Keriven, R., Ponce, J., & Segonne, F. (2008). Segmentation by transduction. In *2008 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1–8). IEEE. <http://doi.org/10.1109/CVPR.2008.4587419>
- Ellingson, B. M., Bendszus, M., Boxerman, J., Barboriak, D., Erickson, B. J., Smits, M., ... Gilbert, M. R. (2015). Consensus recommendations for a standardized Brain Tumor Imaging Protocol in clinical trials. *Neuro-Oncology*, 17(9), 1188–98. <http://doi.org/10.1093/neuonc/nov095>
- Galbán, C. J., Chenevert, T. L., Meyer, C. R., Tsien, C., Lawrence, T. S., Hamstra, D. A., ... Ross, B. D. (2009). The parametric response map is an imaging biomarker for early cancer treatment outcome. *Nature Medicine*, 15(5), 572–6. <http://doi.org/10.1038/nm.1919>
- Galbán, C. J., Chenevert, T. L., Meyer, C. R., Tsien, C., Lawrence, T. S., Hamstra, D. A., ... Ross, B. D. (2011). Prospective analysis of parametric response map-derived MRI biomarkers: identification of early and distinct glioma response patterns not predicted by standard radiographic assessment. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, 17(14), 4751–60. <http://doi.org/10.1158/1078-0432.CCR-10-2098>
- Giese, A. (2003). Cost of Migration: Invasion of Malignant Gliomas and Implications for Treatment. *Journal of Clinical Oncology*, 21(8), 1624–1636. <http://doi.org/10.1200/JCO.2003.05.063>
- Gillies, R. J., Kinahan, P. E., & Hricak, H. (2016). Radiomics: Images Are More than Pictures, They Are Data. *Radiology*, 278(2), 563–77. <http://doi.org/10.1148/radiol.2015151169>
- Hamamci, A., & Unal, G. (2012). Multimodal brain tumor segmentation using the tumor-cut method on the BraTS dataset. *Multimodal Brain Tumor Segmentation - BRATS*, 19–23.
- Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), 255–260. <http://doi.org/10.1126/science.aaa8415>
- Juan-Albarracín, J., Fuster-Garcia, E., Manjón, J. V., Robles, M., Aparici, F., Martí-Bonmatí, L., & García-Gómez, J. M. (2015). Automated glioblastoma segmentation based on a multiparametric structured unsupervised classification. *PloS One*, 10(5), e0125143. <http://doi.org/10.1371/journal.pone.0125143>
- Kanas, V. G., Zacharaki, E. I., Davatzikos, C., Sgarbas, K. N., & Megalooikonomou, V. (2015). A low cost approach for brain tumor segmentation based on intensity modeling and 3D Random Walker. *Biomedical Signal Processing and Control*, 22, 19–30. <http://doi.org/10.1016/j.bspc.2015.06.004>
- Krähenbühl, P., & Koltun, V. (2011). Efficient inference in fully connected CRFs with gaussian edge potentials. *Adv. Neural Inf. Process. Syst.*
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <http://doi.org/10.1038/nature14539>
- Meier, R., Bauer, S., Slotboom, J., Wiest, R., & Reyes, M. (2014a). Appearance-and Context-sensitive Features



- p for Brain Tumor Segmentation. In
- In Proceedings of MICCAI BRATS Challenge*
- .
- <http://doi.org/10.13140/2.1.3766.7846>
- Meier, R., Bauer, S., Slotboom, J., Wiest, R., & Reyes, M. (2014b). Patient-specific Semi-Supervised Learning for Postoperative Brain Tumor Segmentation. In *(In Press) Proceedings of MICCAI 2014* (Vol. 17, pp. 714–721). [http://doi.org/10.1007/978-3-319-10404-1\\_89](http://doi.org/10.1007/978-3-319-10404-1_89)
- Meier, R., Bauer, S., Slotboom, J., Wiest, R., & Reyes, M. (2014). Patient-Specific Semi-Supervised Learning for Postoperative Brain Tumor Segmentation. *Medical Image Computing and Computer-Assisted Intervention -- MICCAI 2014*.
- Meier, R., Knecht, U., Loosli, T., Bauer, S., Slotboom, J., Wiest, R., & Reyes, M. (2016). Clinical Evaluation of a Fully-automatic Segmentation Method for Longitudinal Brain Tumor Volumetry. *Scientific Reports*, 6, 23376. <http://doi.org/10.1038/srep23376>
- Meier, R., Porz, N., Knecht, U., Loosli, T., Schucht, P., Beck, J., ... Reyes, M. (2017). Automatic estimation of extent of resection and residual tumor volume of patients with glioblastoma. *Journal of Neurosurgery*, 1–9. <http://doi.org/10.3171/2016.9.JNS16146>
- Menze, B. H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., ... Van Leemput, K. (2015). The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). *IEEE Transactions on Medical Imaging*, 34(10), 1993–2024. <http://doi.org/10.1109/TMI.2014.2377694>
- Menze, B., Van Leemput, K., Riklin Raviv, T., Geremia, E., Gruber, P., Wegener, S., ... Golland, P. (2015). A generative probabilistic model and discriminative extensions for brain lesion segmentation - with application to tumor and stroke. *IEEE Transactions on Medical Imaging*. <http://doi.org/10.1109/TMI.2015.2502596>
- Nicolaidis, S. (2015). Biomarkers of glioblastoma multiforme. *Metabolism: Clinical and Experimental*, 64(3 Suppl 1), S22–7. <http://doi.org/10.1016/j.metabol.2014.10.031>
- O'Connor, J. P. B., Rose, C. J., Waterton, J. C., Carano, R. A. D., Parker, G. J. M., & Jackson, A. (2015). Imaging intratumor heterogeneity: role in therapy response, resistance, and clinical outcome. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, 21(2), 249–57. <http://doi.org/10.1158/1078-0432.CCR-14-0990>
- Orbach, M., & Crammer, K. (2012). Graph-Based Transduction with Confidence. In *Proceedings of the 2012 European Conference on Machine Learning and Knowledge Discovery in Databases - Volume Part II* (pp. 323–338). Berlin, Heidelberg: Springer-Verlag. [http://doi.org/10.1007/978-3-642-33486-3\\_21](http://doi.org/10.1007/978-3-642-33486-3_21)
- Orrù, G., Pettersson-Yeo, W., Marquand, A. F., Sartori, G., & Mechelli, A. (2012). Using Support Vector Machine to identify imaging biomarkers of neurological and psychiatric disease: a critical review. *Neuroscience and Biobehavioral Reviews*, 36(4), 1140–52. <http://doi.org/10.1016/j.neubiorev.2012.01.004>
- Papandreou, G., & Yuille, A. L. (2011). Perturb-and-MAP Random Fields: Using Discrete Optimization to Learn and Sample from Energy Models. In *Proceedings of the 2011 International Conference on Computer Vision* (pp. 193–200). Washington, DC, USA: IEEE Computer Society. <http://doi.org/10.1109/ICCV.2011.6126242>
- Pereira, S., Pinto, A., Alves, V., & Silva, C. A. (2016). Brain Tumor Segmentation using Convolutional Neural Networks in MRI Images. *IEEE Transactions on Medical Imaging*. <http://doi.org/10.1109/TMI.2016.2538465>
- Porz, N., Bauer, S., Pica, A., Schucht, P., Beck, J., Verma, R. K., ... Wiest, R. (2014). Multi-modal glioblastoma segmentation: Man versus machine. *PLoS ONE*, 9(5), e96873. <http://doi.org/10.1371/journal.pone.0096873>
- Porz, N., Habegger, S., Meier, R., Verma, R., Jilch, A., Fichtner, J., ... Mueller, W. (2016). Fully Automated Enhanced Tumor Compartmentalization: Man vs. Machine Reloaded. *PLoS ONE*, 11(11), e0165302. <http://doi.org/10.1371/journal.pone.0165302>
- Prastawa, M., Bullitt, E., Moon, N., Van Leemput, K., & Gerig, G. (2003). Automatic brain tumor segmentation by subject specific modification of atlas priors1. *Academic Radiology*, 10(12), 1341–1348. [http://doi.org/10.1016/S1076-6332\(03\)00506-3](http://doi.org/10.1016/S1076-6332(03)00506-3)
- Reuter, M., Gerstner, E. R., Rapalino, O., Batchelor, T. T., Rosen, B., & Fischl, B. (2014). Impact of MRI head placement on glioma response assessment. *Journal of Neuro-Oncology*, 118(1), 123–9. <http://doi.org/10.1007/s11060-014-1403-8>
- Rios Velazquez, E., Meier, R., Dunn, W. D., Alexander, B., Wiest, R., Bauer, S., ... Aerts, H. J. W. L. (2015). Fully

- automatic GBM segmentation in the TCGA-GBM dataset: Prognosis and correlation with VASARI features. *Scientific Reports*, 5, 16822. <http://doi.org/10.1038/srep16822>
- Schmidhuber, J. (2014). Deep learning in neural networks: An overview. *Neural Networks*, 61, 85–117. <http://doi.org/10.1016/j.neunet.2014.09.003>
- Schmitt, P., Mandonnet, E., Perdreau, A., & Angelini, E. D. (2013). Effects of slice thickness and head rotation when measuring glioma sizes on MRI: in support of volume segmentation versus two largest diameters methods. *Journal of Neuro-Oncology*, 112(2), 165–72. <http://doi.org/10.1007/s11060-013-1051-4>
- Schölkopf, B., & Smola, A. J. (2002). *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. MIT Press. Retrieved from <https://books.google.com/books?hl=en&lr=&id=y8ORL3DWt4sC&pgis=1>
- Sorensen, A. G., Patel, S., Harmath, C., Bridges, S., Synnott, J., Sievers, A., ... Weisskoff, R. (2001). Comparison of Diameter and Perimeter Methods for Tumor Volume Calculation. *J. Clin. Oncol.*, 19(2), 551–557. Retrieved from <http://jco.ascopubs.org/content/19/2/551.short>
- Sottoriva, A., Spiteri, I., Piccirillo, S. G. M., Touloumis, A., Collins, V. P., Marioni, J. C., ... Tavaré, S. (2013). Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. *Proceedings of the National Academy of Sciences of the United States of America*, 110(10), 4009–14. <http://doi.org/10.1073/pnas.1219747110>
- Urban, G., Bendszus, M., Hamprecht, F. A., & Kleesiek, J. (2014). Multi-modal brain tumor segmentation using deep convolutional neural networks. *Proceedings of MICCAI BraTS (Brain Tumor Segmentation) Challenge 2014*, 31–35.
- Vladimir, V. (2006). Transductive Inference and Semi-Supervised Learning. In *Semi-Supervised Learning*. The MIT Press. <http://doi.org/10.7551/mitpress/9780262033589.003.0024>
- Wang, J., Jebara, T., & Chang, S.-F. (2008). Graph Transduction via Alternating Minimization. In *Proceedings of the 25th International Conference on Machine Learning* (pp. 1144–1151). New York, NY, USA: ACM. <http://doi.org/10.1145/1390156.1390300>
- Weizman, L., Sira, L. Ben, Joskowicz, L., Rubin, D. L., Yeom, K. W., Constantini, S., ... Bashat, D. Ben. (2014). Semiautomatic segmentation and follow-up of multicomponent low-grade tumors in longitudinal brain MRI studies. *Medical Physics*, 41(5), 52303. <http://doi.org/10.1118/1.4871040>
- Wen, P. Y., Macdonald, D. R., Reardon, D. A., Cloughesy, T. F., Sorensen, A. G., Galanis, E., ... Chang, S. M. (2010). Updated response assessment criteria for high-grade gliomas: response assessment in neuro-oncology working group. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology*, 28(11), 1963–72. <http://doi.org/10.1200/JCO.2009.26.3541>
- Yang, D. (2015). Standardized MRI assessment of high-grade glioma response: a review of the essential elements and pitfalls of the RANO criteria. *Neuro-Oncology Practice*, 3(1), npv023. <http://doi.org/10.1093/nop/npv023>
- Zhang, Y., Brady, M., & Smith, S. (2001). Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging*, 20(1), 45–57. <http://doi.org/10.1109/42.906424>