# MICCAI 2014 Challenge Proposal

Title: **Brain Tumor Digital Pathology Challenge**

## 1. Description of the Challenge

Brain tumor diagnosis, as traditionally performed by pathologists, involves examination of tissue sections on glass slides with a light microscope. As the technology for digital imaging has advanced, there is now increasing use of digital images ("virtual slides") for pathologic analysis of surgical specimens. Automated tumor segmentation, by defining tumor regions with critical histologic features has the potential to increase both the speed and accuracy of diagnosis by pathologists and/or computer software. Key features in digital pathology diagnosis of brain tumors, which lend themselves well to computer-aided detection and automation, as means of supporting clinical decision making by pathologists, include, (a) differentiation of high and low grade gliomas, and (b) segmentation of necrotic zones – both of which have significant implications for downstream tissue sample genetic testing.

A critical dividing line in brain tumor diagnosis is separating glioblastoma multiforme (GBM, WHO Grade IV) from lower grade glioma (LGG, defined as WHO grade II or III for purposes of the TCGA), since GBM compared to LGG has worse prognosis and may be treated more aggressively by neuro-oncologists. Diffusely infiltrating astrocytic tumors with necrosis are, by definition, glioblastoma. Necrosis, therefore, is a critical histopathologic variable in separating GBM from LGG.

There will be two sub-challenges in the proposed Brain Tumor Digital Pathology Challenge:

> **Sub-Challenge I: Classification - Automated classification of LGG and GBM from a collection of 30 high resolution digital pathology slide clinical cases.**
>
> **Sub-Challenge II: Segmentation – Automated segmentation of necrotic and normal brain regions on regions of digital pathology slides from a collection of 20 GBM cases.**

This challenge and a related challenge (MICCAI-BRATS 2014 – Challenge on Multimodal Brain Tumor Image Segmentation, proposed by Menze, et. al.), will make use of data currently available through data archive resources of the National Institutes of Health (NIH), namely, the Cancer Genome Atlas (TCGA) and the Cancer Image Archive (TCIA) –see Figure 1. The proposed Digital Pathology challenge will use digital slides related to patients whose genomics data are available from TCGA. Similarly, BRATS 2014 Challenge will use clinical MRI image data, also from the TCGA study subjects. The common approaches taken by both of these Challenges are scientifically important in a number of ways - First we will be using high quality publicly available data for which the corresponding genomics data are also available. Second, targeting pathology and imaging data from the same patient cohort explores a connection between the Digital Pathology Challenge and BRATS Challenge, which joins two key drivers of clinical decision making, i.e., pathologic and radiologic diagnoses, as well as support for neurosurgical, radiotherapy and neuro-oncologic treatment planning. Third, these dual brain tumor challenges proposed for 2014 will set the stage for future MICCAI challenges which will further explore connections between multimodal data (genomics, pathology, and imaging).

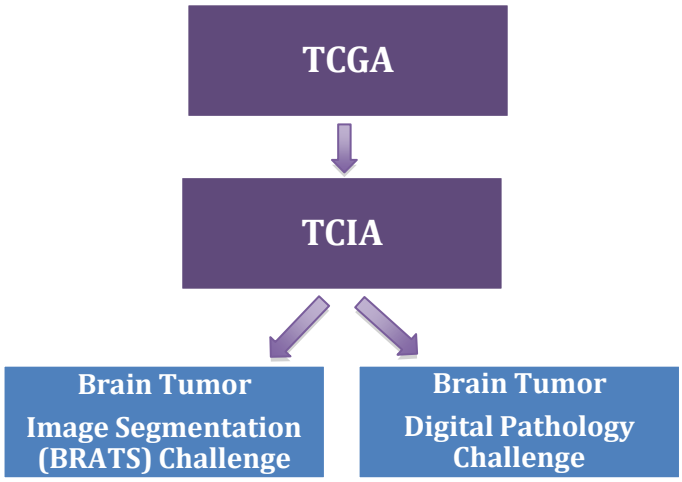


***Figure 1 –*** *The Cancer Genome Atlas (TCGA) contains tissue samples, pathology and imaging data of clinical cases of brain tumors. Anonymized imaging and digital pathology data corresponding to TCGA data will be available for the Brain Tumor Digital Pathology Challenge (this proposal) and the Brain Tumor Image Segmentation (BRATS) Challenge on the Cancer Image Archive (TCIA). Efforts will be made to use imaging and pathology data from the same clinical subjects for both brain tumor challenges.*

2. Plan

The Brain Tumor Digital Pathology Challenge will include two sub-challenges: Classification and Segmentation. For each sub-challenge there will be Training and Leader Board Phases prior to the actual challenge (Test Phase). The timeline and allocation of cases for each sub-challenge are shown in Figure 2:

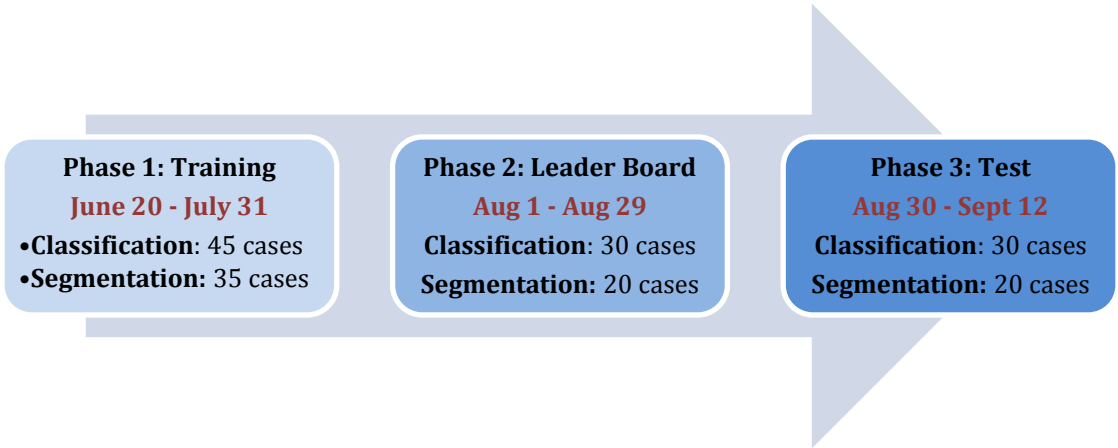


***Figure 2 –*** *Timeline and number of unique cases for each phase of Classification and Segmentation sub-challenges. GBM cases used for the Segmentation sub-challenge will be a subset of the combined LGG and GBM cases used for the Classification sub-challenge. All data will be anonymized and available on the NCI TCIA Website for download. The proposed time line may be adjusted in consideration of MICCAI registration requirements.*

In Phase 1 (Training) prospective contestants can train their algorithm to perform the required tasks. No Ground Truth or evaluation of results will be available in this phase. In Phase 2 (Leader Board) prospective contestants may submit their algorithm results for evaluation against sequestered Ground Truth data. This will help them to further optimize their algorithms. In Phase 3 (Test) contestants will submit their results for evaluation and ranking in the Brain Tumor Digital Pathology Challenge. Similar to the BRATS Challenge, we prefer to hold the Test phase "off-site" and on-line two weeks prior to the date for MICCAI Challenges. This will allow for timely download of large amounts of data (in a 24 hr time window), processing and uploading of results.

2a. Case Selection and Manual Markup

The datasets that will be used in the Challenge will be compiled by the challenge organizers using a combination of software systems as follows. The datasets will consist of digital pathology images, image regions, and annotated markups. The pathology images will be downloaded at original resolution from the TCGA repository. For case selection efficiency the organizers will use the Cancer Digital Slide Archive (CDSA, http://cancer.digitalslidearchive.net) to view and select images to be used for the Challenge – the CDSA provides a web-based graphical user interface, which makes it easy to search for images and view them. Once a set of images is selected, the case number for each image (e.g., TCGA-02-0001-01Z-00-DX3) will be used to download the image from the TCGA pathology slide repository. The Bulk Download section of the TCGA data portal will be used for this purpose. The case number will be used on the data portal to find the corresponding tissue image data files. Pathology images in the TCGA repository are stored in Aperio Scanscope format (i.e., extension .svs). Aperio provides a free software tool, called Aperio ImageScope (http://www.aperio.com/appcenter), to view, annotate, and crop images. It allows for the cropped images to be saved in TIFF or jpeg formats. Once a rectangular region in an image is selected using ImageScope, the region (tile) will be saved as a TIFF file using a lossless compression. The subregions of interest in the tile will be marked and annotated using Photoshop or iPhotoDraw, which is a free software tool for marking up images. The markups and annotations for each tile will be saved in a file for evaluating computer generated analysis results.

3. Evaluation of Computer Algorithm Results

Computer algorithms that participate in the challenge will be required to output their results as image masks and associated annotations. The resolution of the mask is the same as the source image/tile. An image mask is a 2D array of integers where integers of the same value represent the region segmented by the computer algorithm. Different integer values represent different regions. The integer value 0 (zero) means background (or the region that was not segmented by the algorithm). A region is annotated by a tuple (integer value, annotation) where the integer value corresponds to the integer value in the mask marking the region. The masks will be used to compute the metrics for comparing algorithm output with the expert markups and annotations in the segmentation sub-challenge.

Classification sub-challenge: The score for each contestant in the Classification sub-challenge will be computed as the number of correctly classified cases divided by the number of total cases. There will be guidelines, to be posted on the challenge website, outlining how issues related to a tie or lucky guess work will be resolved.

Segmentation sub-challenge: The metrics will include: (1) the amount of overlap between a region segmented by an algorithm and the region marked by the pathologist, (2) the

number and area of regions that were missed by the algorithm, (3) the number and area of regions segmented and annotated by the algorithm, but not by the pathologist.

Evaluation of individual results in the Leader Board and Test phases of the Brain Tumor Digital Pathology Challenge will be automated through a web site submission.

4. Manuscript and Poster Submissions

Participants will be required to submit short manuscripts outlining their approach and preliminary results on the training data. Test results of Phase 3 will be announced at the workshop and the top three scoring teams will be invited to give 12 min presentations of their methodology and results. All participants will be invited to submit posters for exhibition at MICCAI 2014. A joint journal publication of all participants may be prepared after the workshop. All submissions will be reviewed by organizers of the challenge.

5. Announcements and Anticipated Number of Participants

We will publicize the event through appropriate email distributions, including imageworld, machine-learning, visionlist, and NCI Cancer Imaging Program website and listserv. Based on past experience with similar challenges we expect about 10 groups (more than 20 registrants) to participate in the Brain Tumor Digital Pathology Challenge.

6. Workshop and Challenge Cluster at MICCAI 2014

A workshop entitled "Computational Clinical Decision Support and Precision Medicine in Brain Cancer" (Farahani, Jaffe, Clark) will be proposed through a separate proposal. If proposals for the workshop, BRATS 2014 and the Brain Tumor Digital Pathology Challenge are all accepted, we plan to combine all three events into a full day cluster of workshop and challenges related to brain tumor classification and segmentation in support of clinical decision making and treatment planning (see the Appendix for a Preliminary Agenda). This approach will help raise awareness of computational needs in medical imaging and computer-aided intervention within MICCAI and help connect various drivers of clinical decision support systems.

7. MICCAI Resource Needs

Space for 60+ people with one projector and one screen (or multiple large screen display monitors), wireless Internet access. Due to the clinical relevance of topics in the Brain Tumor Workshop and Challenge Cluster we request holding these events at the Harvard Medical School. This will encourage participation by clinicians and their research staff.

8. Organizers & Major Contributors:

Daniel J. Brat, Emory University (dbrat@emory.edu)
Larry Clarke, National Cancer Institute (lclarke@mail.nih.gov)
James Davis, Stony Brook Cancer Center (James.Davis@stonybrookmedicine.edu)
Keyvan Farahani, National Cancer Institute (Farahani@nih.gov)
John Freymann, Leidos Biomedical Res, Inc(freymanj@mail.nih.gov)
Carl Jaffe, Boston University (carljaffe@gmail.com)
Justin Kirby, Leidos Biomedical Res., Inc, MD (kirbyju@mail.nih.gov)
Tahsin Kurc, Stony Brook Cancer Center (tahsin.kurc@stonybrook.edu)
Miguel Ossandon, National Cancer Institute (mo193x@nih.gov)
Joel Saltz, Stony Brook Cancer Center (Joel.Saltz@stonybrookmedicine.edu)

**Contact: Keyvan Farahani (NCI): farahani@nih.gov ; 240-276-5921**

**Appendix:** Preliminary Agenda for the Brain Tumor Workshop, Imaging, and Digital Pathology Challenges offered as a 1-Day cluster of related events

# MICCAI 2014- Workshop and Challenges in Brain Tumors

**Workshop Title: Computational Clinical Decision Support and Precision Medicine in Brain Cancer: The Value of Open Science Grand Challenges**

Goal: To present and discuss requirements and resources for open science development of systems for clinical decision support and precision medicine in brain cancer diagnosis and therapy based on Big Data, including genomics, pathology, and imaging.

**8:30 am – 12:00 pm Workshop**

[Chairs: Clarke (NCI), Farahani (NCI), Jaffe (BU)]

8:30-8:45 am Introduction

8:45-10:00 am Invited talks (3-4) – on computation and/or path correlation, plus

- Report: NCI 2013 Workshop on Imaging and Genomics
- Open science platforms for assessment of technologies

10:00-10:30 am Break

10:30-11:30 am Proffered papers

11:30 am-12:00 pm Presentation of NCI resources: TCGA, TCIA, HubZero, etc.

**1:00 pm – 5:30 pm Brain Tumor Challenges**

**1:00 – 3:00 pm Brain Tumor Image Segmentation Challenge (BRATS)**

[Chairs: Kalpathy-Cramer (MGH), Menze (ETH), Reyes (Bern)]

1:00 – 1:30 pm Presentation of results by chairs and discussion of results

1:30 – 2:15 pm Presentations by top 3 challenge winners (12 min each + 3 min discussion)

2:15 – 3:00 pm General discussion

3:00 – 3:30 pm Break

**3:30 – 5:30 pm Brain Tumor Digital Pathology Challenge**

[Chairs: Saltz (Stony Brook), Brat (Emory), Gilbertson (MGH)]

3:30 – 4:00 pm Presentation of results by chairs and discussion of results

4:00 – 4:45 pm Presentations by top 3 challenge winners (12 min each + 3 min discussion)

4:45 – 5:30 pm General discussion and wrap-up

5:30 pm Adjourn

**Scientific Committee**: D. Brat, L. Clarke, J. Davis, K. Farahani, J. Freymann, J. Gilbertson, C. Jaffe, J. Kalpathy-Cramer, J. Kirby, T. Kurc, B. Menze, S. Mercer, M. Ossandon, M. Reyes, J. Saltz