The Digital Pathology Challenge has two sub-challenges:

**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.

Each sub-challenge will have three phases; *training, leader board, and test*. In the training phase, contestants will use the datasets to train their algorithm to perform the required tasks. No Ground Truth or evaluation of results will be available in this phase. In the leader board phase, contestants may submit their algorithm results for evaluation against Ground Truth data. This will help them to further optimize their algorithms. In the Test phase, contestants will submit their results for evaluation and ranking in the challenge.

**Input and Output Datasets:**

In the classification sub-challenge, whole slide tissue images will be used. The training set will have 45 images, the leader board set will have 30 images, and the test set will have 30 images. Contestants will download the images from the NCI TCIA (The Cancer Image Archive) site.

When an image is processed by a classification algorithm, the output will be a label indicating whether the image is LGG or GBM. What we expect from each contestant in the classification sub-challenge is a file that contains <image file name, classification label> tuples with the following format:

challenge\_image1.svs LGG

challenge\_image2.svs GBM

challenge\_image3.svs GBM

challenge\_image4.svs LGG

…

In the segmentation sub-challenge, rectangular tiles/regions extracted from whole slide images will be used. There will be 35 image tiles in the training set, 20 image tiles in the leader board set, and 20 image tiles in the test set. Contestants will download the image tiles from the challenge web site (our web site).

When an image tile is processed by a segmentation algorithm, the output will be a mask indicating the segmented necrosis regions – all the other regions in the tile will be considered normal tissue regions. A tile may contain more than one necrosis regions. Each contestant will be asked to submit a text file containing the dimensions of the tile and the region id label for each pixel in row-major order with C-style indexing (i.e., pixel indices will start from 0). Each segmented region should be given a different id (1,…,K). The normal tissue region will have the id of 0 (zero). The format of the file will be as follows:

Tile\_xdim Tile\_ydim // N M

Pixel\_region\_id // pixel (0,0)

Pixel\_region\_id // pixel (0,1)

Pixel\_region\_id // pixel (0,2)

…

Pixel\_region\_id // pixel (N-1,M-1)

For example, assume a tile of 4x5 pixels has two necrosis regions:

00000

01100

22200

02200

The mask file will have the following content:

4 5

0

0

0

0

0

0

1

1

0

0

2

2

2

0

0

0

2

2

0

0

Each contestant will upload a zip file containing the masks for all the image tiles that were processed by his/her algorithm. If a contestant wishes to submit results for multiple algorithms, the contestant will be asked to submit separate zip files, each containing the results for a segmentation algorithm. The web interface will allow the contestant to enter some metadata as a free-text note about each submission; e.g., the metadata may give a name to the algorithm, describe the algorithm, etc. The zip file will also contain a mapping file to associate mask files with image files as follows:

Mask1\_file.txt Image1\_file.tif

Mask2\_file.txt Image2\_file.tif

…