

## < RT704 Assignment 2 >

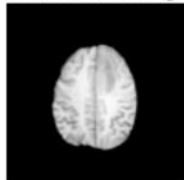
Glioma is the most primary brain tumor in the human brain. Glioma is divided into glioblastoma(HGG) and lower grade glioma(LGG) with HGG being an aggressive and life-threatening tumor. In this assignment, we will classify HGG/LGG of multi-modal brain MRI images via several machine learning algorithms.

Download MRI data. This dataset contains training, validation and testing data for classification. For each subject, `img.npy`, `seg.npy` and `label.npy` were prepared. `Img.npy` contains **[4,H,W]** size array which represents **4 types of MRI images(T1, T2, T1ce, T2-Flair)**. `Label.npy` contains classification label (**0: LGG and 1: HGG**). `Seg.npy` is **pixel-level annotation of tumor**. It contains subregions that are heterogeneous, each with a different pattern on the brain MRI scan. Though each subregions were annotated with **1,2, and 4** in the `[H,W]` array, just use them as a single label using `array=np.clip(array, 0, 1)`.

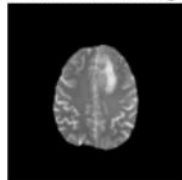
Check below examples. You can load the data from `.npy` file using this `np.load(file_path)`.

```
number of train labels: 222  
number of valid labels: 29  
number of test labels: 84
```

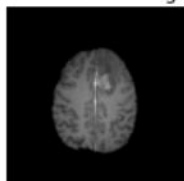
1 th channel image



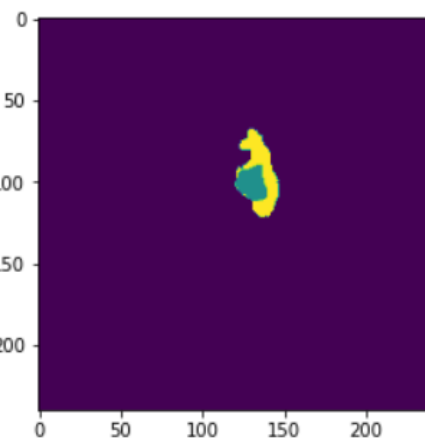
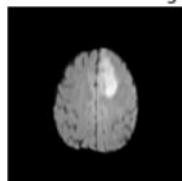
2 th channel image

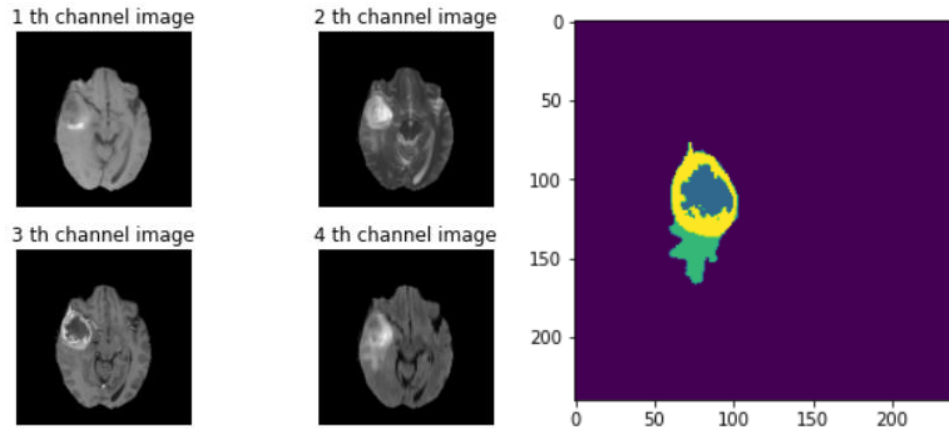


3 th channel image



4 th channel image





You can find more related information at <http://braintumorsegmentation.org/> (optional)

[40 pt] 1. Train a CNN-based HGG/LGG classification model and test it. Use a simple CNN network as there are not that many data samples for training. Concatenate the image and mask array and use it as an input. Report your observation.

[40 pt] 2. Train logistic regression, neural network, and random forest classifiers using radiomics features and test them. Use pyradiomics library for feature extraction (see <https://www.radiomics.io/pyradiomicsnotebook.html>) and scikit-learn library for random forest model.

(see <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html>). The pyradiomics lib require nrrd file format. You need to save each type of MRI with nrrd file first. Report your observation.

[20 pt] 3. Report the difference between the deep-learning based CNN model and the machine learning algorithms.

**Submit your report with the codes on LMS by 11/20. When you submit, make your zip filename “HW2\_yourfirstname.zip”**