Yong Ik Jeong
EECS 349 Project

Motivation

Glioblastoma multiforme or GBM is a grade IV tumor that is found in the brain, and it is considered to be highly malignant with a very low chance of survival for the patient. It is difficult to find treatment, because there are many types of GBM tumors and certain treatment only work for a small subset of patients. In order to study the tumors and post-process them, region of interest or ROI have to be drawn on the medical image of the tumor by someone like a radiologist manually. And this process takes a lot of time and induces intra- and inter-reviewer variability. In this project, the goal was to use machine learning to automatically segment the GBM tumor in MRI images of the brain.

Solution

For this task, the MRI images were first pre-processed by co-registering the different types of images in each case and normalizing them. This was done to minimize misalignment of images and allow differently scaled cases to be used for analysis. The co-registration was done using a widely accepted protocol Statistical Parametric Mapping. And the normalization was done in such a way that the same type of tissue is scaled to the same value for every case using normal white matter, gray matter and cerebrospinal fluid as reference. The scaling was separate for each type of image.

The main focus of this task was on using supervised learning, namely, k-Nearest Neighbors and Support Vector Machines. The problem to solve was binary classification using continuous variables, so k-Nearest Neighbors and Support Vector Machines were good choices to use. Initially, k-means clustering was also tried, but it wasn't performing as well as the others so this option was taken out.

The input features used for this task were the pixel intensity of three types of MR images (namely, T1-weighted image, T1-weighted image with contrast agent, and T2-weighted image), and the mean and standard deviation of corresponding voxel's neighborhood. The output feature was a binary tumor mask drawn manually. Figure 1 shows an example case of the normalized three types of MR images. A total of nine inputs and one output was formed as follows:

$$GBM_{i,j} = f(T1si_{i,j}, T1Csi_{i,j}, T2si_{i,j}, T1m_{i,j}, T1Cm_{i,j}, T2m_{i,j}, T1std_{i,j}, T1Cstd_{i,j}, T2std_{i,j})$$

where for every pixel *i* in case *j*

- *GBM* is the output label,
- *T1si* is T1-weighted signal intensity,
- *T1Csi* is the T1-weighted with contrast agent signal intensity,
- *T2si* is the T2-weighted signal intensity,
- *T1m* is the average *T1si* of the pixel's surrounding,
- *T1Cm* is the average *T1Csi* of the pixel's surrounding,
- *T2m* is the average *T2si* of the pixel's surrounding,
- *T1std* is the standard deviation of *T1si* of the pixel's surrounding,
- *T1Cstd* is the standard deviation of *T1Csi* of the pixel's surrounding, and
- *T2std* is the standard deviation of *T2si* of the pixel's surrounding,

Training and Testing

A total of 114 cases were used for this project (from the Cancer Imaging Archive), and each case consisted of the three types of images, tumor mask, and a brain mask. The brain mask was used to limit the pixels to within the brain, and only the slice with the largest tumor cross-sectional area was chosen. Every image was also reduced to a size of 128x128. This was necessary in order to shorten the training time and memory usage, as otherwise it was taking forever (in terms of days) and unfeasible. While 114 cases was probably not enough to capture and train on all kinds of tumors, each pixel corresponded to

one instance so there were plenty of data points; for example, the number of instances in the training set was 674,268 and test set had 239,018 instances.

The training set and test set were split by 75% and 25% of total cases. Using the training set, a 10-Nearest Neighbor model and a SVM model were trained. The k value was chosen after performing a 10-fold CV multiple times on the training set with different k values. For SVM it was found that a Radial Basis Function kernel was doing better than a Linear kernel using 10-fold CV. Then predictions were made using the test set and success was measured by the overall accuracy $(\mathrm{TP} + \mathrm{TN})/\mathrm{Total}$ across all instances and Dice-coefficient $(2 * \mathrm{TP})/(2 * \mathrm{TP} + \mathrm{FP} + \mathrm{FN})$ comparing the tumor region across all instances.

Results

Both the 10-Nearest Neighbor model and the SVM model performed better than a ZeroR (accuracy: 91%, Dice-coefficient: 0). The 10-NN model had an overall accuracy of 94.5% while the SVM model had 94.9%. On the other hand, the Dice-coefficient for the 10-NN model was 0.53 and 0.59 for the SVM model. The Dice-coefficient is the more important measure for this task, so SVM seemed to perform better than 10-NN. By looking through the cases visually, prediction seemed to be good for certain types of tumor while bad for other types. However, it's hard to say what the distinguishing difference is. The best case had a Dice-coefficient of 0.87, while the worst case had 0.04. Figure 2 shows an example case of a good prediction. Additionally, normalization improved the average Dice-coefficient; on a 10-NN model, training on non-normalized data gave a coefficient of 0.41.

Suggestions for Future Work

Since this task was for a course project, time was limited and it was difficult to implement other algorithms that are outside the scope of this course. More sophisticated algorithms like Markov Random Field would be useful to take into account the dependence of pixels to its neighborhood. Convolutional Neural Network would also be a good option to try. Also including other types of MR images, such as diffusion, would be useful as they all provide different information. Unfortunately for this project, not many types of MR images were available. And lastly, getting a powerful computer would definitely be good for trying many different options, etc.

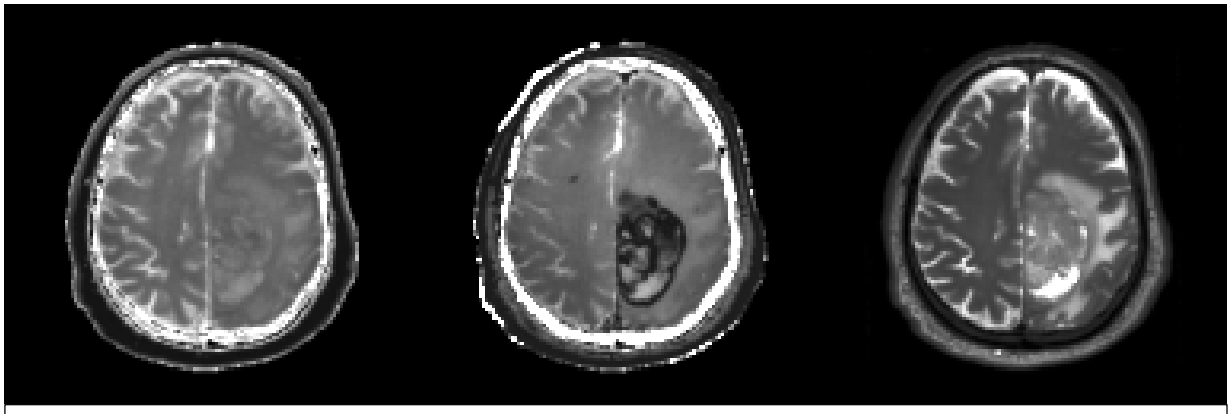

Figure 1: Example of normalized version of T1-weighted image, T1-weighted image with contrast agent and T2-weighted image (from right to left).

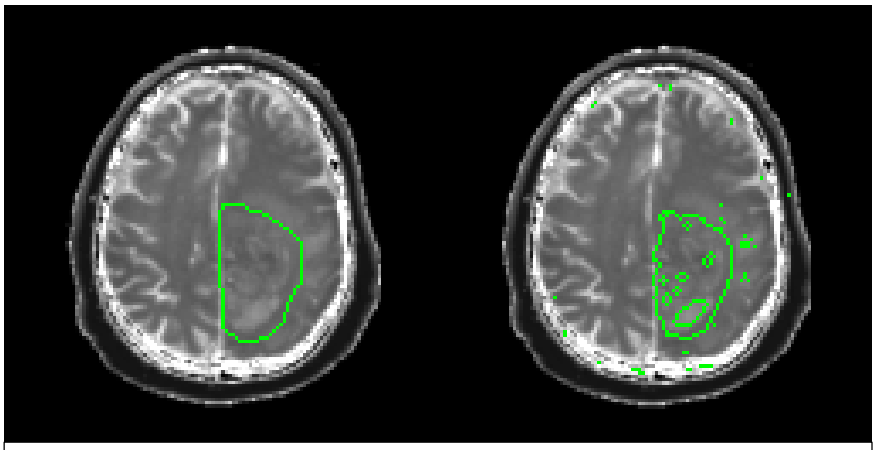

Figure 2: Example of a case that was predicted well with a Dice coefficient of 0.87 (same case as in Figure 1). The left image is the hand drawn tumor ROI and the right image is the predicted tumor ROI.