BRAIN TUMOR SEGMENTATION
CS6347 Project Group 3

Jie Liu jxl164830@utdallas.edu
Xiaobo Ma xxm160330@utdallas.edu
Hemeng Tao hxt160430@utdallas.edu
Shasha Jin sxj152630@utdallas.edu

1 INTRODUCTION

Gliomas are the most common primary brain malignancies, with different degrees of aggressiveness, variable prognosis and various heterogeneous histological sub-regions, i.e. peritumoral edema, necrotic core, enhancing and non-enhancing tumor core. This intrinsic heterogeneity of gliomas is also portrayed in their imaging phenotype (appearance and shape), as their sub-regions are described by varying intensity profiles disseminated across multimodal MRI scans, reflecting varying tumor biological properties. Due to this highly heterogeneous appearance and shape, segmentation of brain tumors in multimodal magnetic resonance imaging(MRI) scans is one of the most challenging tasks in medical image analysis[1].

To date, human radiologist segmentation is still the golden standard of segmentation. However, with the exploding size of image data, manual segmentation of all images is no longer a applicable method. In order to tackle this problem, we propose a statistical learning pipeline in this project that takes extracted features from image into a Nave Bayes classifier to segment and classify tumors into their sub-categories.

Our main goals are as follow:

- Set up a model that can process large amount of data in reasonable time. The training process should finish within hours and inference should finish within minutes.
- Find appropriate feature extraction methods and conduct experiments to compare their performance.
- Analyze the strengths and weaknesses of different feature extraction methods and reach out to the reasons behind.

[1] http://braintumorsegmentation.org/

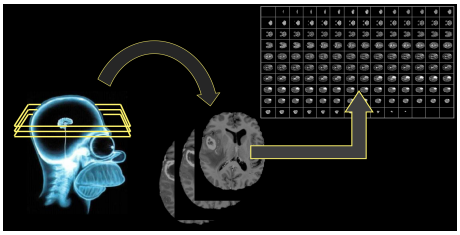

Figure 1: Basic MRI workflow

2 DATA

2.1 DATA DESCRIPTION

All data was provided by the 2016 MICCAI BraTS Challenge [4], which consists of 220 high-grade glioma(HGG) cases and 54 low-grade glioma(LGG) cases. This dataset is available at https://www.smir.ch/BRATS/Start2016. Both HGG and LGG datasets contain four different MRI pulse sequences and each case is comprised of 155 brain slices from one patient, for a total of 620 images of size 240×240 per patient. See Figure 1 to get a better understanding of MRI scan.

Professional manual segmentation is provided as ground truth labels for each image. There are 5 segmentation labels: '0' means the pixel is normal with no tumor, while '1' to '4' are labels represents 4 sub-categories respectively. Figure 2 is an example of a scan with the ground truth segmentation, 4 colors are assigned to each sub-category (1-green, 2-yellow, 3-blue, 4-red).

There are multiple radio frequency pulse sequences options in MRI scan. To generate this dataset, four different sequences has been used: Fluid Attenuated Inversion Recovery (FLAIR), T1, T1-contrasted(T1C), and T2. Each pulse sequence exploits the distinct chemical and physiological characteristics of various tissue types, resulting in contrast between the individual classes. Im-

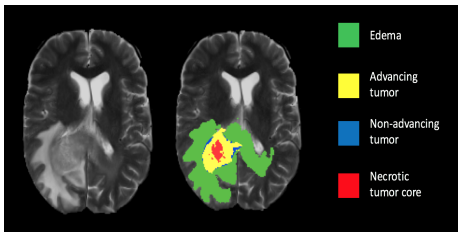

Figure 2: Ground truth on a T2 Weighted scan

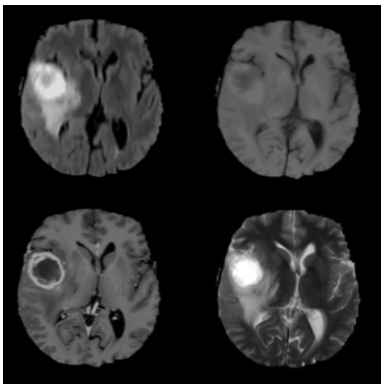

Figure 3: FLAIR(top left), T1(top right), T1C(bottom left), T2(bottom right)

ages of the same brain taken with different pulse sequences are shown in Figure 3. The dataset comes already pre-processed, which involves skull stripping and co-registration. Co-registration is crucial, as it aligns images for all four pulse sequences such that the voxel found in the image coordinates $[x, y, z]$ in all four pulse sequences will point to the same coordinates in real space.

2.2 DATA SELECTION

Since it is a very large dataset, also considering the limitation of the project scale, we decided to do the experiments only on the HGG portion. One fact to be noticed is that MRI data usually contains artifacts introduced either by inhomogeneity in the magnetic field or small movements made by the patient during scan time. Sometimes a bias will occur and affect the segmentation results particularly in computer-based models. After carried some experiments on different pulse sequences, we found that the bias in T1 and T1C images is very high. Although some bias correction techniques are available [7], we decided to dismiss the images from T1 and T1C pulse sequence and focused only on the T2 and FLAIR pulse sequence. Moreover, considering the fact that the majority of the brain is always located in the center of a 3-dimension space, which indicates that the first and last several slices of the scan images are practically worthless black pixels, only 10 equally space separated slices $(31, 41, 51, ..., 121)$ are selected.

As we are utilizing Naive Bayes classifier, we encode all MRI images into samples and each sample represents one pixel in one image. So the total number of samples are calculated as: number of patients $\times$ number of pixels per patient, which in our case we have $220 \times (240 \times 240 \times 10)$, approximately 120 million samples(a huge burden on the computer memory).

3 FEATURE EXTRACTION

Although features from this pre-processed dataset are invariant to uniform scaling, orientation, illumination changes, and partially invariant to affine distortion, considering the size of our dataset, extracting useful features is still difficult.

Multiple feature detectors have been used in our project.

3.1 FAST

Based on the fact that all images are in gray scale which contains simple texture information, we started with a simple feature extraction algorithm: FAST (Features from Accelerated Segment Test) [8] .

The traditional FAST algorithm mainly consists of following steps:

- Select a pixel in the image which is to be identified as an interest point.
- Draw a circle of n pixels around the pre-selected pixel under test(Figure 4).

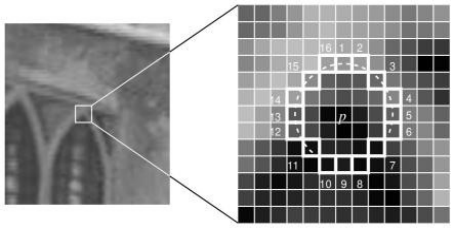

Figure 4: FAST algorithm example

Instead of using FAST algorithm to extract corners in the image, we apply FAST to get information of each pixel point. The features we choose for each pixel are:

- Energy:
$$\sum_{(i,j)} p(i.j)^2$$
where (i,j) are pixels around the selected point.

- Contrast:

$$\sum_{(i,j)} |p(i,j) - p|$$

where p is the pixel value of the pre-selected point, and p(i, j) are the values corresponding to the surrounding pixels.

3.2 LBP

Local Binary Pattern (LBP) [5] is another simple but very efficient texture operator which labels the pixels of an image by thresholding the neighborhood of each pixel and considers the result as a binary number.

The LBP algorithm mainly contains following steps:

- Divide the brain images into cells, and each cell contains 16×16 pixels.
- Derive the gray scale and rotation invariant operator by defining pixel T in a local neighborhood of a monochrome image as the joint distribution.

$$T = t(g_c, g_0, ..., g_{p-1})$$

where gray value g_c corresponds to the gray value of the center pixel and $g_p, p \in [0, p-1]$ correspond to the gray values of P equally spaced pixels on a circle of radius R.

- Achieve invariance with respect to the scaling of the gray scale by considering just the signs of the differences instead of their exact values:

$$T = t(s(g_0 - g_c), s(g_1 - g_c), ..., s(g_{p-1} - g_c))$$

where

$$s(x) = \begin{cases} 1, & x > 0 \\ 0, & otherwise \end{cases}$$

- Transform into a unique $LBP_{P,R}$ number that characterizes the spatial structure of the local image texture(Figure 5) by assigning a binomial factor 2^p for each $sign(s(g_p - g_c))$, :

$$LBP_{P,R} = \sum_{k=0}^{p-1} s(g_k - g_c) \times 2^k$$

- Compute the histogram for each cell, the histogram can be seen as a 256-dimensional feature vector. Then concatenate (normalized) histograms of all cells. This gives a feature vector for the entire brain image. The histogram for one brain scan image is shown in Figure 6.

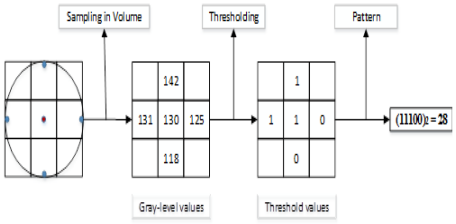

Figure 5: LBP example

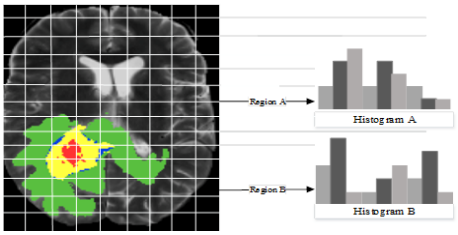

Figure 6: LBP histogram example

3.3 SIFT

Scale-invariant feature transform(SIFT) [3] is an algorithm in computer vision to detect and describe local features in images. It is quite an involved and complicated algorithm. This process is done for different octaves of the image in Gaussian Pyramid. It is represented in Figure 7.

The basic steps for SIFT algorithm are as follow:

- Constructing a scale space—the initial preparation: internal representations of the original image is created to ensure scale invariance. This is done by generating a 'scale space'.
- LoG Approximation: The Laplacian of Gaussian is great for finding interesting points (or key points) in an image. But it is computationally expensive. In order to control the complexity, the representation created in the previous step is used as an approximation.
- Finding keypoints: With the super fast approxima-

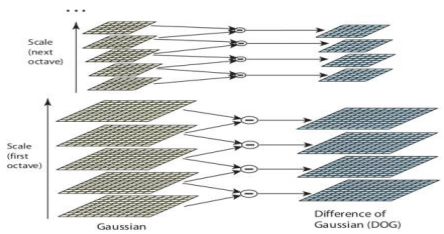

Figure 7: SIFT octave structure

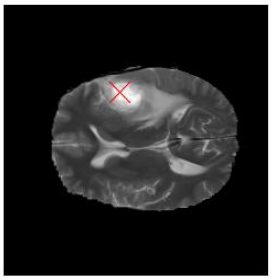

Figure 8: SIFT point

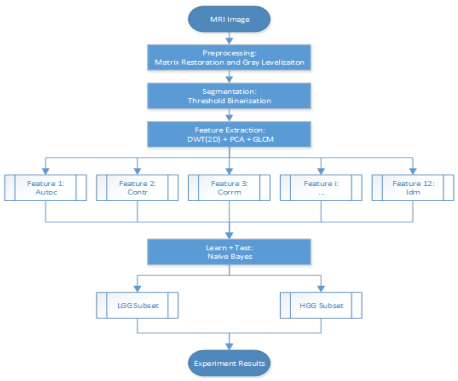

Figure 9: GLCM workflow

tion, the key points are located. These are maxima and minima in the Difference of Gaussian image we calculate in step 2.

- Assigning an orientation to the keypoints and generate SIFT features.

We treat the keypoints with least scale as our SIFT point shown in Figure 8. The SIFT point represents the location of the tumor. We construct the 6 dimensional features in the following way:

- Pixel value
- Pixel value difference between each pixel and the SIFT point
- Hamming distance between each pixel and the SIFT point
- x value
- y value
- slice number

3.4 GLCM

A statistical method of examining texture that considers the spatial relationship of pixels is the gray-level co-occurrence matrix (GLCM) [6], also known as the gray-level spatial dependence matrix. The GLCM functions characterize the texture of an image by calculating how often pairs of pixel with specific values and in a specified spatial relationship occur in an image, creating a GLCM, and then extracting statistical measures from this matrix. The basic workflow of GLCM feature extraction is shown in Figure 9:

- Construct 2-D DWT of the input binary images. Discrete Wavelet Transform(DWT) is a powerful implementation of the Wavelet Transform (WT) using the dyadic scales and positions (more information is available on Wikipedia[2]). In case of 2D images, the DWT is applied to each dimension separately.
- Apply Principle Component Analysis(PCA) [2] to reduce the feature dimensions. PCA is an efficient tool to reduce the dimension of a data set consisting of many interrelated variables while retaining most of the variations. It is achieved by transforming the data set to a new set of ordered variables according to their variances or importance.
- Calculate Gray Level Cooccurrence Matrix from PCA coefficients. GLCM is defined as a 2-dimensional histogram of gray levels for a pair of pixels, which are separated by a fixed spatial relationship. An image GLCM is computed using a displacement vector d, defined by its radius a and orientation e.

Notice that one GLCM is created for one image. So this time we are trying to construct features for images not pixels. 12 statistic feature values are also computed, such as Contrast, Correlation, Energy, Homogeneity, Mean, Standard Deviation, Entropy, Variance, Smoothness, Kurtosis, Skewness, Inverse Dierence Moment (IDM) or Homogeneity. The statistic features are combined with GLCM feature so that we get the features for each image.

4 EXPERIMENT AND DISCUSSION

Experiments are accomplished based on the selected data. Four different sets of extracted features are re-

[2]https://en.wikipedia.org/wiki/Discrete_wavelet_transform

garded as the input for Naive Bayes classifier separately. 80% of the data are used to train the model while the rest 20% are used for test purpose. Since the dataset is very biased, i.e. the majority of pixels in the images are normal points (not tumor, labeled 0), and almost all of them can be correctly classified, there is no surprise that the overall classification accuracy is over 90%. However, these normal points are of no interest in our goals, so we only evaluate the classification accuracy on the tumor pixel points, whose ground truth label is $[1, 2, 3, 4]$. The result are shown in Table 1.

Table 1: Experiment results

Accuracy(%)	FLAIR	T2
FAST	39.16	25.59
LBP	35.75	23.37
SIFT	38.43	36.25
GLCM	45.23	60.45

We start our discussion by the feasibility of applying Naive Bayes classifier on this brain tumor segmentation task. As we can tell from Table 1, Naive Bayes is not a suitable classifier. This is because all the features are treated as independent of each other under the Naive Bayes model, but actually there may be very strong relationships among them so that cannot be captured by Naive Bayes model. According to the result published by the BraTS (Multimodal Brain Tumor Segmentation Chanllege), the state of art classifier model on this task is achieved by the convolutional neural network [1]

FAST algorithm is widely used in corner points detection applications. But in our case, considering the complicated structure of human brains as well as the vague edges between two tumor sub-types, corner points in the gray scale scan images may or may not correspond to the edge of the tumor tissue, and therefore hurts the accuracy of the result.

LBP is in a similar situation as to FAST method, LBP can detect local features that corresponds to a block of pixels, but when it turns to extract features for each pixel point, the information we can get from LBP method is very limited.

The problem with SIFT method we adopted in our model is that although the SIFT point we found is always on the tumor, it may not locate on the center of the tumor. In other words, SIFT features based on the SIFT point can result in very bad performance if the SIFT point is far off the target from the center of the tumor. In this situation, the probability of misclassifying the tumor pixel points that far away from the SIFT point is very high

GLCM has the best performance since the feature extracted are much more sophisticated.Gray level co-occurrence matrix has been proved to be a powerful basis for use in texture classification. Various textural parameters calculated from the gray level co-occurrence matrix help understand the details of the overall image content. This property of GLCM is a good match for our segmentation task.

References

[1] Mohammad Havaei, Axel Davy, David Warde-Farley, Antoine Biard, Aaron Courville, Yoshua Bengio, Chris Pal, Pierre-Marc Jodoin, and Hugo Larochelle. Brain tumor segmentation with deep neural networks. *Medical image analysis*, 35:18–31, 2017.

[2] Ian Jolliffe. *Principal component analysis*. Wiley Online Library, 2002.

[3] Tony Lindeberg. Scale invariant feature transform. *Scholarpedia*, 7(5):10491, 2012.

[4] Bjoern H Menze, Andras Jakab, Stefan Bauer, Jayashree Kalpathy-Cramer, Keyvan Farahani, Justin Kirby, Yuliya Burren, Nicole Porz, Johannes Slotboom, Roland Wiest, et al. The multimodal brain tumor image segmentation benchmark (brats). *IEEE transactions on medical imaging*, 34(10):1993–2024, 2015.

[5] Timo Ojala, Matti Pietikäinen, and Topi Mäenpää. Gray scale and rotation invariant texture classification with local binary patterns. In *European Conference on Computer Vision*, pages 404–420. Springer, 2000.

[6] L-K Soh and Costas Tsatsoulis. Texture analysis of sar sea ice imagery using gray level co-occurrence matrices. *IEEE Transactions on geoscience and remote sensing*, 37(2):780–795, 1999.

[7] Nicholas J Tustison, Brian B Avants, Philip A Cook, Yuanjie Zheng, Alexander Egan, Paul A Yushkevich, and James C Gee. N4itk: improved n3 bias correction. *IEEE transactions on medical imaging*, 29(6):1310–1320, 2010.

[8] Deepak Geetha Viswanathan. Features from accelerated segment test (fast), 2009.