ECE 4250: Digital Signals and Image Processing

# Automatic Segmentation of Brain MRI scans

Sonu Kapoor
Spring 2020
Cornell University

## Introduction and Background

MRI scans are crucial to understand the anatomy of the brain and identify abnormalities in the brain structure. Proper labeling - also known as segmentation - of various parts of the brain is a key element in this venture. Labelling of the MRI scans would traditionally be done manually, with the help of experts in this subject area. However, the process of manual segmentation is time-consuming and expensive, due to the dependency on these experts. Given the size of the current population that might need a brain MRI scan, it is not feasible to manually label all of these images, thus creating a need for an automated method of segmenting these images.

Automatic segmentation, in an ideal world, would identify the different regions of the brain and label the pixels using a predetermined guide. The algorithm "trains" itself using samples of manually segmented images, and applies a series of transformations to the image to automatically segment the MRI scan. This, however, introduces challenges regarding the accuracy of these automatically segmented images. The Sørensen–Dice Index is used to measure this accuracy using the following formula:

$$\mathrm{Dice} = \frac{2|A \cap M|}{|A| + |M|}$$

Where A and M are the automatic and manual segmentations of the image respectively.

This project is a program that would use a series of training and validation images to automatically segment a group of testing images. The goal is to maximise the Dice coefficient, using numerous methods that are to be discussed in this report.

## Methods

A brief overview of the algorithm used is presented below. The process can be divided up into the following sections, and various alterations can be made in each section to refine the algorithm. These alterations are mentioned below and discussed in further detail in the 'Discussion and Results' section of this report.

1. <u>Extracting the images</u>

The automatic segmentation algorithm requires a series of training and validation images of the original MRI scans as well as the manual segmentations of these scans. The middle coronal slice is extracted using simple matrix operations and saved for further use. The automatic segmentation algorithm has three main steps: an image transform function, an optimisation function, and a label fusion strategy.

2. Image Transform

The transform function outputs an altered version of the input image using linear interpolation. These alterations could include affine transformations like scaling, shearing, and translating, as well as more complex piecewise affine or non-parametric transformations. Here, initially, a simple 4 parameter transform (scale, angle, x and y translation) was applied, which was later modified to take 6 parameters - a 2x2 affine matrix and x and y translation. This transform function is applied to the training (moving) images in order to look more similar to the validation (fixed) images.

3. Optimisation

The loss value between these images can be computed as the sum of squared differences between the fixed and transformed moving images. The goal of the optimising function is to minimise this loss in order to achieve the highest accuracy segmentation possible. Two built-in functions were tested for optimisation - the Nelder-Mead method as well as Powell's method. The results of these methods are further discussed below. Additionally, the input images also underwent histogram equalisation and Min-Max Normalisation to further improve the result of the optimisation function.

4. Transforming the Segmentations

Once the minimum loss is achieved, the transformation parameters associated with this loss can be applied to the manual segmentation of the moving image to create the automatic segmentations of the fixed image. Now, in the case of transforming the segmentations, simply using linear interpolation would not be suitable as the pixel values denote specific labels. The pixel values here need to be discrete, and thus, two interpolation strategies were tried - Nearest Neighbour and Linear Interpolation of the one-hot encoded image. These are discussed in detail as well.

5. Label Fusion

The process above works for a single moving image. To achieve higher accuracy, multiple moving images can undergo the same process to then obtain a series of automatic segmentations for the same fixed image. The different versions of the segmentations can be combined using a label fusion strategy. Some of the strategies mentioned are majority-vote, global atlas weighted mode, and local atlas weighted mode.

6. Computing and Improving the Accuracy

The accuracy of these automatically computed segmentations can be calculated using the Dice index formula, which takes the ratio of the intersection and union of the Regions of Interest in the manual and automatic segmentations. By observing the results of this index, the user can further tweak the algorithm to maximise the Dice coefficient. For this project, the Regions of

Interest were the left and right Cerebral Cortex, and the left and right White Matter. The corresponding fill codes for these regions were provided by the instructors of this class.

## Results

The table below provides a quick summary of the various changes made to the algorithm along with the resulting Dice Indices. A discussion of these changes is provided after the table. The code submitted is the one with the highest Dice Index (Kaggle score) for the Testing subjects.

| No. | Image Transform | Optimisation Method | | | Segmentation | Label Fusion | Dice Index | |
|---|---|---|---|---|---|---|---|---|
| | | | Normalisation | | | | | |
| | No. of Parameters | Minimiser | Histogram | MinMax | Interpolation Strategy | Strategy | Validation | Testing (Kaggle) |
| 1 | 4 | NM | No | No | NN | MV1 | 0.653 | N/A |
| 2 | 6 | NM | No | No | NN | MV1 | 0.624 | N/A |
| 3 | 6 | Powell | No | No | NN | MV1 | 0.646 | N/A |
| 4 | 6 | Powell | Yes | No | NN | MV1 | 0.692 | 0.6745 |
| 5 | 6 | NM | Yes | No | NN | MV1 | 0.651 | N/A |
| 6 | 6 | Powell | Yes | Yes | NN | MV1 | 0.688 | 0.68013 |
| 7 | 6 | Powell | Yes | Yes | NN | MV5 | 0.707 | 0.709 |
| 8 | 6 | Powell | Yes | Yes | LOH | MV5 | 0.707 | 0.709 |
| 9 | 6 | Powell | Yes | No | LOH | MV3 | 0.709 | 0.704 |
| 10 | 6 | Powell | Yes | Yes | NN | MV3 | 0.706 | N/A |
| 11 | 6 | Powell | Yes | Yes | LOH | GAWM | 0.709 | 0.705 |
| 12 | 6 | Powell | Yes | Yes | LOH | LAWM1 | 0.721 | 0.704 |
| 13 | 6 | Powell | Yes | Yes | LOH | LAWM5 | 0.717 | 0.712 |

* 'NM' denotes Nelder Mead, NN denotes Nearest Neighbour, 'LOH' denotes Linear One Hot, 'MV#' denotes Majority Voting in a # by # neighborhood, 'GAWM' denotes Global Atlas Weighted Mode, and 'LAWM#' denotes Local Atlas Weighted Mode in a # by # neighborhood

## Discussion

This section provides a detailed discussion about the various changes made to the algorithm along with their effect on the accuracy and Dice Index that is highlighted in the table above.

1. Improved Transform Function

**4 parameter vs. 6 parameter geometric transform**
Initially, a four parameter geometric transform was implemented in the algorithm, with input variables: scale, angle of rotation, horizontal translation, and vertical translation. However, a 6-parameter affine transform felt like a better choice as it offered more movement, such as shearing. On one side it was possible to get more accurate results but on the other, there was a higher likelihood of the optimising function getting stuck in a local loss minimum instead of a global one.

**Piecewise Affine Transforms and  Non-Parametric Transform**
The use of the piecewise affine and non-parametric transforms would allow for higher flexibility of movement in the moving image, and thus, result in a higher Dice index. However, these types of transforms would be incredibly difficult to implement and extremely time-consuming to optimise, given the number of variables.
2. Upgraded Optimisation Method

**Powell's vs Nelder-Mead Minimiser**
The Dice Indices with Powell's method were consistently higher than those with Nelder-Mead. However, it was noted that Powell's method of optimisation took a longer time than the Nelder-Mead method. Thus, there was a negative correlation between accuracy and time.

**Histogram Equalisation**
Implementing Histogram Equalisation drastically improved the Dice index. This is because the contrast of the images was better adjusted and more accurate optimal transformation parameters could be found by the optimisation function.

**Min-Max Normalisation**
The result of Min Max Normalisation had unclear results. The change in Dice index was not consistently higher or lower for any particular case. Thus, the use of this normalisation method was a trial-and-error. A possible reason for this uncertainty is that it limited the intensity values and thus resulted

3. Alternate Interpolation Strategy for Image Segmentation Transform

**Nearest Neighbour Interpolation vs. Linear Interpolation of One-hot Encoded Image**
Initially, the Nearest Neighbour Interpolation method was used to transform the manual segmentation of the moving image. As discussed earlier, the pixel of interest took the value of the closest neighbouring pixel. In order to remove any errors caused by this interpolation, an alternate interpolation method was introduced. First, the original segmentation image is turned into a one-hot encoded image for the different intensities - so a 256x256 array was now a 256x256xN array where N is the number of unique intensity values. In the $k^{th}$ layer of this array, the value of each pixel was set to 1 if the intensity at that point equalled k, and 0 otherwise. Then, each of these layers underwent the geometric transform with the optimised parameters using linear interpolation. Now, each of these layers depicted a probability of having that particular intensity. The resulting transformed segmentation image was then obtained by assigning each pixel with the highest probability intensity at the particular point. There was no discernable difference in the usage of these two implementation methods, so these methods were used interchangeably.

4. Enhanced Label Fusion

**Inclusion of Neighborhood Pixels in the Modal Calculation**
Including pixels from the neighbourhood helped eliminate the amplification of any noise present in the system. Different sizes of neighborhoods (1x1, 3x3, 5x5, 10x10) resulted in different Dice indices. It was expected that the 3x3 neighbourhood would be best but it turned out that a 5x5 neighborhood was the most optimal. These observations are not fully exhaustive in the above table due to space constraints, but can be confirmed with quick tests.

**Majority Vote vs. Weighted Mode**
Majority Vote is a version of Weighted Mode where all of the images are weighted equally. This might not be ideal as images with higher loss values might affect the result of the label fusion and negatively impact the Dice coefficient. A solution to this problem is weighing each of the images according to their loss values. The scaled used on the image was S = 10*exp(-6*NormalisedLoss) + 1 where NormalisedLoss was the loss value divided by the maximum loss value among the images. This scale ensured that a smaller loss value was rewarded and a higher one was punished to a greater degree. The sample array over which the mode was calculated depended on this scale - pixels belonging to a particular image were repeated S times. Thus, the mode was now weighted.

**Global vs. Local Atlas Weighted Mode**
The previous method is a Global Atlas Weighted Mode, and one of the biggest limitations is that the loss value for any given image is non-uniform across the pixels. To incorporate this factor, a

Local Atlas Weighted Mode was implemented. The primary difference of this method is that the scale depends on the loss value of a 5x5 neighbourhood of pixels as opposed to the entire image. This proved to result in a more accurate label fusion.

## Ethical Considerations

Machine learning is proving to be increasingly useful in medical imaging, as it speeds up the process of segmentation and decreases the expenses involved. This means that imaging and image processing resources are now more readily available to the medical industry. However, this raises some serious ethical questions.

One such issue is the detection of false positives due to machine error, and the effect on the treatment of patients. Irregularities caused by automatic segmentation might appear as severe health issues and lead to inaccurate diagnoses. The resulting unnecessary medical intervention might be expensive and even life-threatening. It is, therefore, important to have routine manual checks to limit the number of inaccuracies in the imaging data.

Another issue is the usage of these now low-cost and abundant resources. Since hospitals would now have a lower cost-basis attached to the imaging department, there might be a surge in the level of usage of these resources. This goes hand-in-hand with the previous issue, where the rate of misdiagnosis increases. Besides the additional treatment cost, it is possible that the hospital would still charge hefty fees to the patient or the patient's insurance for the imaging itself. There would be a higher profit margin as the cost decreases and the number of 'customers' increases. However, this practice would be considered ethically ambiguous as it would result in an unnecessary utilisation of resources.

## References


- ❖ Prof Mert Sabuncu's Lecture Notes from the class
- ❖ Some discussion of methods with classmates Anja Samardija and Kumail Al Hamoud, as well as the teaching assistants of the course.