BIANCA Training on Cleveland Clinic Data

MRI Lab @ NYU Abu Dhabi
Zunair Viqar

10th April, 2022

Overview

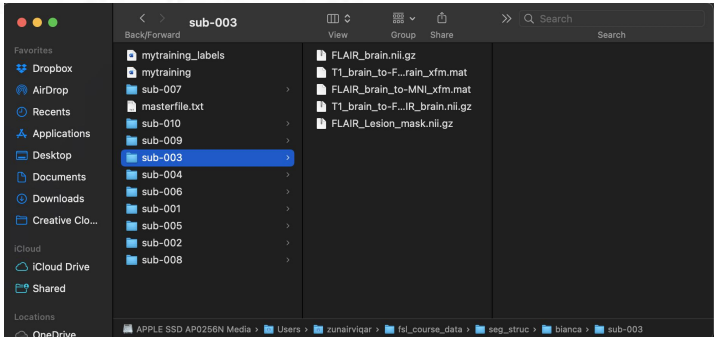
- After using the existing training dataset for BIANCA in their training practical, we decided to manually segment our own dataset in order to ensure the BIANCA model is better suited to our needs
- We collected data for multiple patients, resulting in a total of 18 Flair & T1 Images.
- These 18 Scans were then Manually Segmented using FSLEyes to produce our training dataset
- From the 18 Scans, 16 were taken for training purposes and then the remaining 2 were taken for testing purposes.
- I will explain the entire process in further detail in the coming slides.

Preparing the Data

- As the first step, I followed the Training Practical that required us to have the following data points:

  - `FLAIR_brain.nii.gz` : main structural image, brain extracted
  - `FLAIR_Lesion_mask.nii.gz` : binary manual lesion mask for the subjects used to train BIANCA
  - `FLAIR_brain_to-MNI_xfm.mat` : transformation matrix from subject space (main structural image) to standard space (optional). This is to be able to use spatial features (MNI coordinates)
  - `T1_brain_to-FLAIR_brain.nii.gz` : additional input (optional). Other modalities that can help the lesion segmentation (e.g. T1), all registered to the main image. Click here to see how it was obtained.

- Then, I looked at the Training Data from the practical to have a closer look:



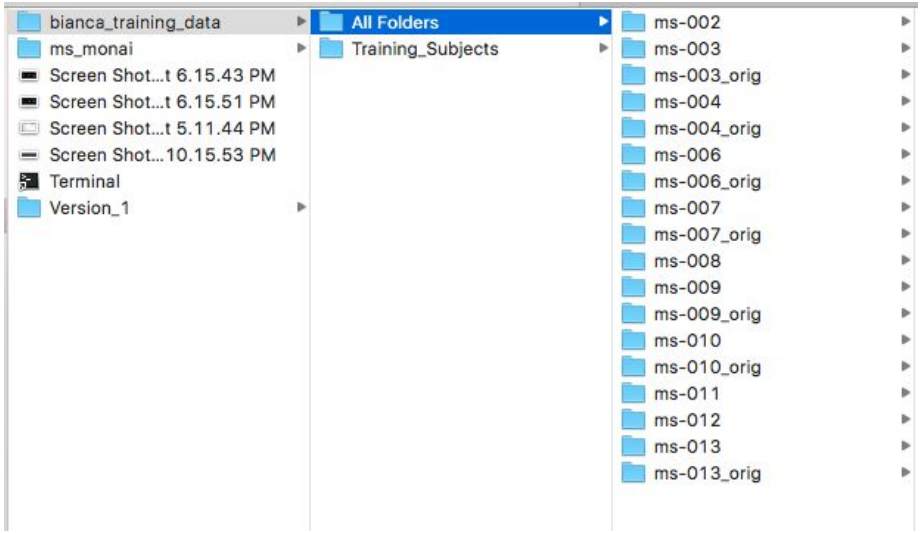
- After this, I made a list of the things that were required:
  - Flair Brain Image (We already had)
  - Flair Brian Lesion Mask (We created this)
  - T1 Brain Image (We already had)
  - Flair Brain to MNI for .mat (Yet to be extracted)
  - T1 Brain to Flair Brain (Yet to be extracted)

First Step: Resampling for Same Dimensions (1)

- However, before proceeding towards extracting any of the required images as mentioned earlier, we first had to make sure that all of the Flair Images and their corresponding masks were in the same dimensions

- Therefore, I chose the most frequently occurring dimension - 256x256x40 - since most of our images were in this dimension and would ultimately lead to the least amount of reduction in data.

- Inside of bianca_training_data, I created two folders, 'All Folders' and 'Training_Subjects'. All Folders contained all the folders for each subject with and without being resampled. For the subjects that required a resampling, their original images are stored in the **{subject_name}_orig** folder, and for those that did not require any modification, their folder names are as it is.

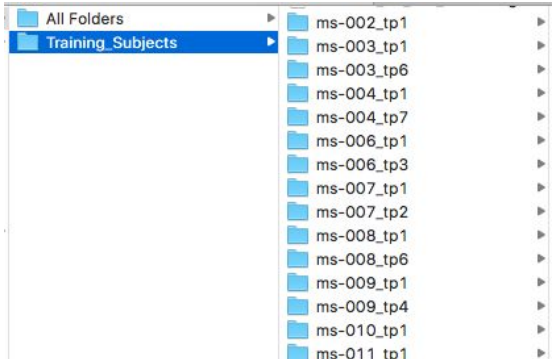


First Step: Resampling for Same Dimensions (2)

- Now, in order to make sure the dimensions were the same, I had to run the following two commands simultaneously;
  - flirt -in **{input image}** -ref **{ms-002 flair brain image}** -out **resampled_image.nii.gz** -omat **resampling.mat** -interp nearestneighbour -dof 6
  - flirt -in **{mask for the same input image}** -ref **{ms-002 flair brain image}** -out resampled_mask.nii.gz -applyxfm -init **resampling.mat** -interp nearestneighbour -dof 6

- And then simply deleted the resampling.mat since it was no longer required. Then, I also renamed the **resampled_image.nii.gz** to **ms_0xx_tpx_flair_brain.nii.gz** and **resampled_mask.nii.gz** to **ms_0xx_tpx_flair_lesion_mask.nii.gz** and then copied them into the folder.

- Once this was done for all the subjects, I copied all the folders with the resampled data into Training_Subjects folder and split the different timepoints of each subjects in different folders. Now this is where we will do any other data extraction.



MRI Lab @ NYU Abu Dhabi
Zunair Viqar

Second Step: Flair Brain to MNI for .mat

- Now that we have all the images over the same dimensions, we can move forward with the step of extracting the .mat file needed to train BIANCA by registering the flair brain image over the sample MNI Image that we had used earlier in our previous version using the following steps:

    - First, I copied the **MNI152_T1_2mm_brain.nii.gz** file onto the Training_Subjects folder

    - Then I ran the following command for registration purposes:
        - flirt -in **{flair brain image}** -ref MNI152_T1_2mm_brain.nii.gz -out **{flair brain to MNI}** -omat **{flair brain to MNI .mat}** -bins 256 -cost normcorr -searchrx -180 180 -searchry -180 180 -searchrz -180 180 -dof 7 -interp nearestneighbour

    - This command was run over all the flair images across the 17 subject files from ms-002-tp1 to ms-013-tp1. (17 subjects because I trained the model on 16 and tested on the 17th, ms-013_tp3 is left for another trial)

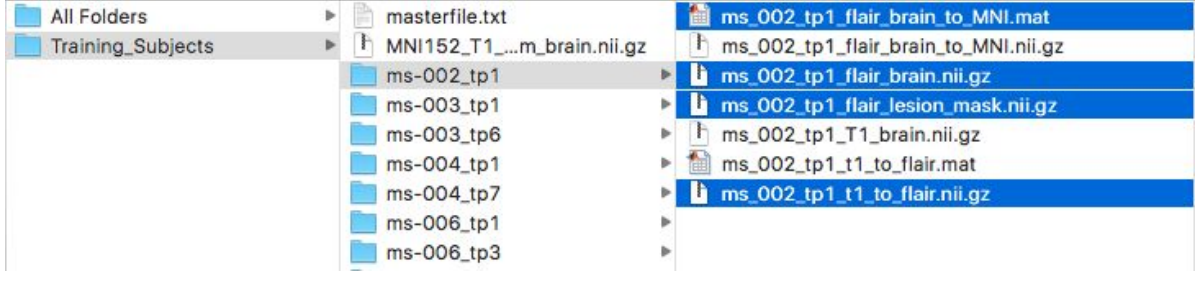
- As a result we obtained the **ms_0xx_tpx_flair_brain_to_MNI.mat** for file for every subject

Third Step: T1 Brain to Flair Brain

- Doing the same as we did in the Second Step, this time I used the T1 Brain Image and then registered it onto the Flair Brain images across all the subjects using the following command:

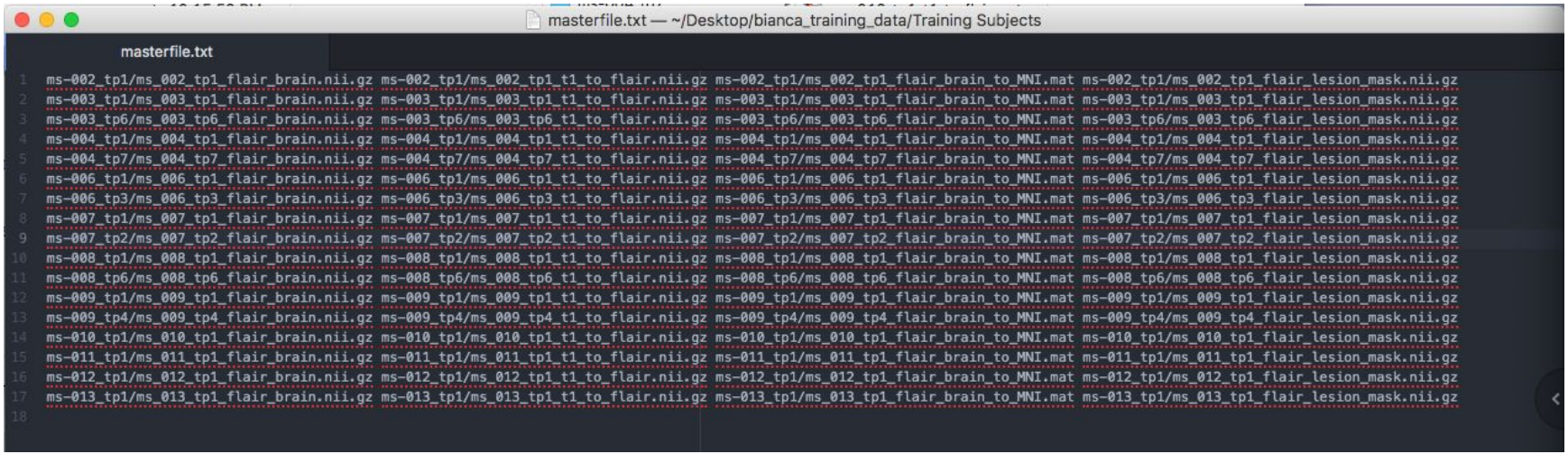
**flirt **
**in {T1 Brain Image} **
**-ref {Flair Image for the same subject} **
**-out {T1 to Flair Image} **
**-omat {.mat file for the T1 to Flair Image} **
-bins 256 -cost normcorr -searchrx -180 180 -searchry -180 180 -searchrz -180 180 -dof 7 -interp nearestneighbour

- After running the previous command across all the 17 subjects, we had the following files in the folder for each subject, from ms-002-tp1 to ms-013-tp1 (highlighted images are the ones to be used by BIANCA, others are by-products):



Fourth Step: Masterfile

- Once we had all the files, I moved onto creating a **masterfile.txt** which included all the four files that were highlighted on the screenshot at the bottom of the previous slide. The text files looks as follows:



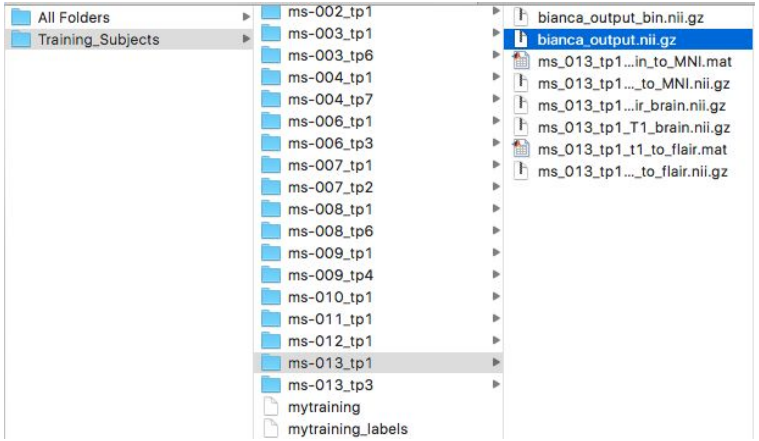
- Even though ms-013_tp1 does not have any flair mask, the file is still mentioned as it is a requirement to Train BIANCA and the standard for creating the masterfile. The intention of the last one is to be used for testing purposes.

Fifth Step: Train BIANCA

- Now that we have all the data to Train BIANCA, and it was linked in the Masterfile, I went forward and trained BIANCA using the following command:

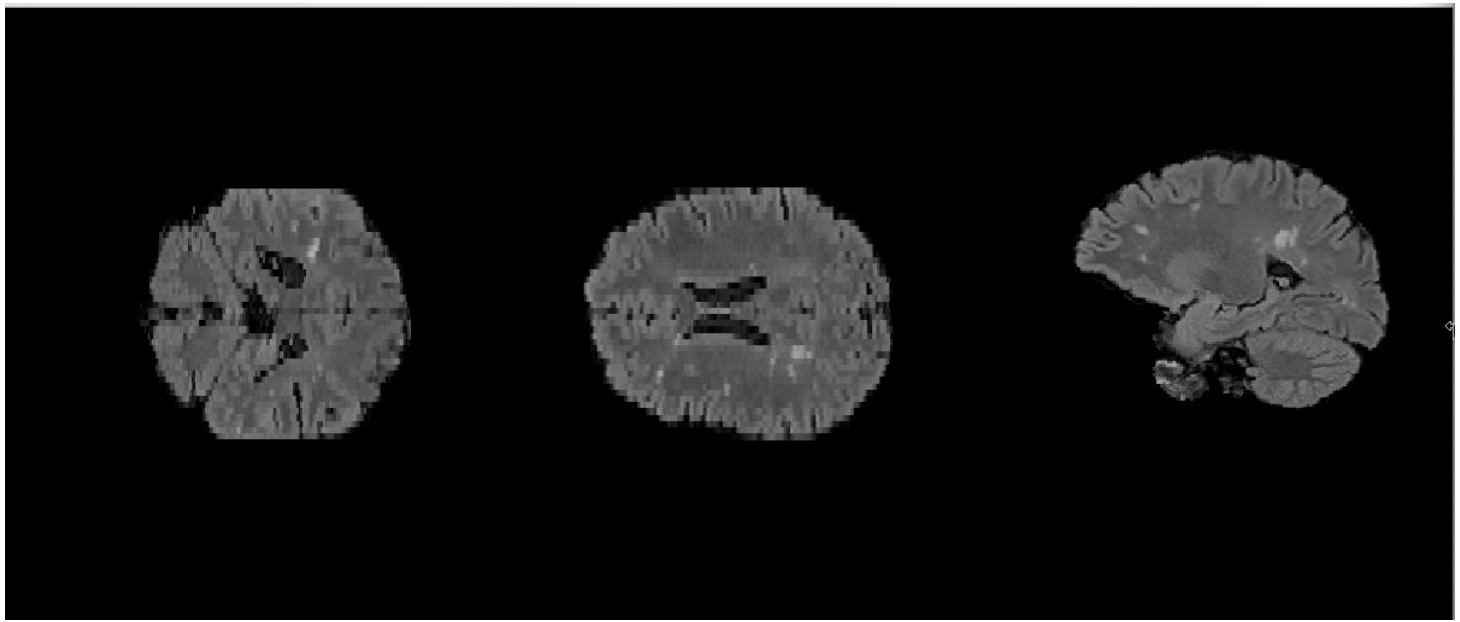
```
bianca --singlefile=masterfile.txt --trainingnums=1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16 --labelfeaturenum=4 \
--querysubjectnum=17 --brainmaskfeaturenum=1 --featuresubset=1,2 --matfeaturenum=3 \
--trainingpts=2000 --nonlespts=10000 --selectpts=noborder -o ms-013_tp1/bianca_output \
--saveclassifierdata=mytraining -v
```

- After it took some time to run, I obtained a lesion mask for ms-013_tp1 and also the training_labels and the training file which is the trained model that can be used on any other subjects to come in the future.



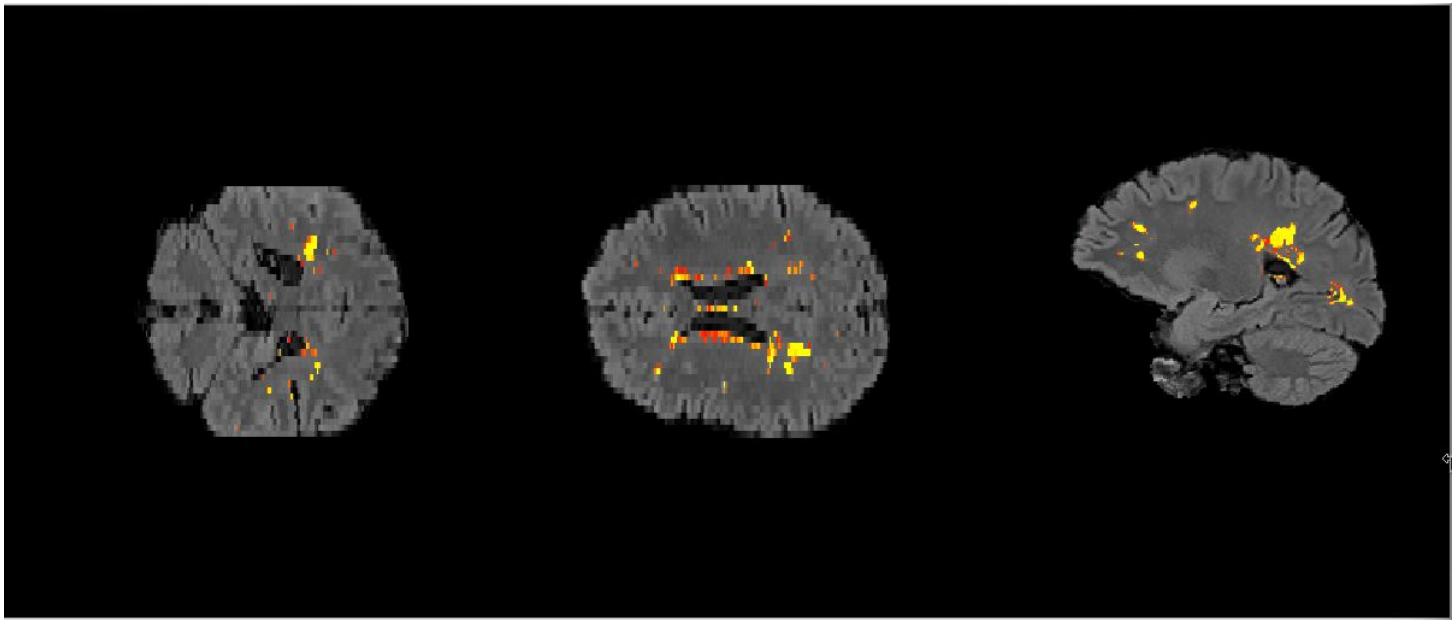
Results:

- After putting a threshold of 0.7 on the bianca output, the result seemed very decent
- Original Slice:



Results:

- BIANCA Segmented Slice:



- A decent result with no excessive registrations or resampling!