Hippocampal Volume Quantification in Alzheimer's Progression

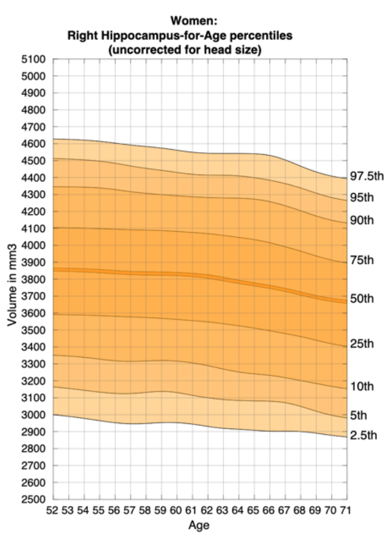
Validation Plan
Rafael Castro

Algorithm Description

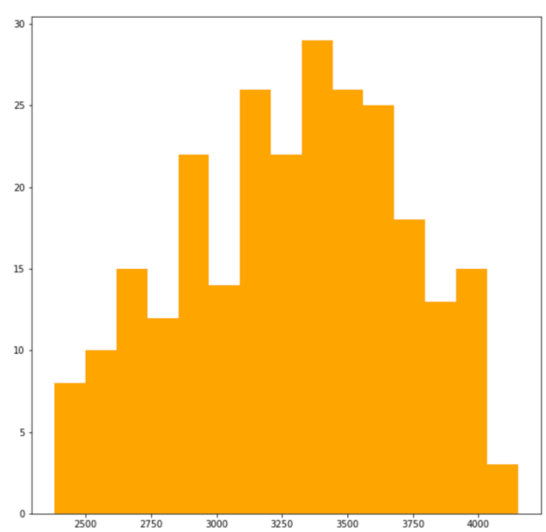
1. Background and Intended Use

Background
Alzheimer's disease (AD) is a progressive neurodegenerative disorder that results in impaired neuronal (brain cell) function and eventually, cell death. AD is the most common cause of dementia. For patients exhibiting early symptoms, quantifying disease progression over time can help direct therapy and disease management. A radiological study via MRI exam is currently one of the most advanced methods to quantify the disease. In particular, the measurement of hippocampal volume has proven useful to diagnose and track progression in several brain disorders, most notably in AD. Studies have shown a reduced volume of the hippocampus in patients with AD.

According to Nobis et al., 2019, the volume of hippocampus varies in a population, depending on various parameters, within certain boundaries, and it is possible to identify a "normal" range taking into account age, sex and brain hemisphere.



Data volume ranges from 2200 to 4500 mm3

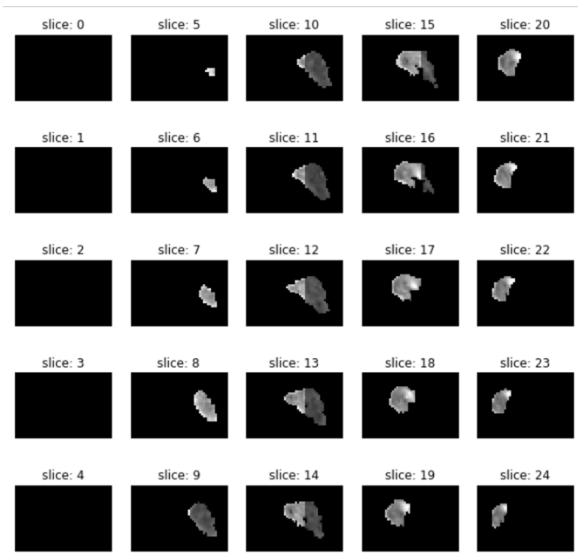


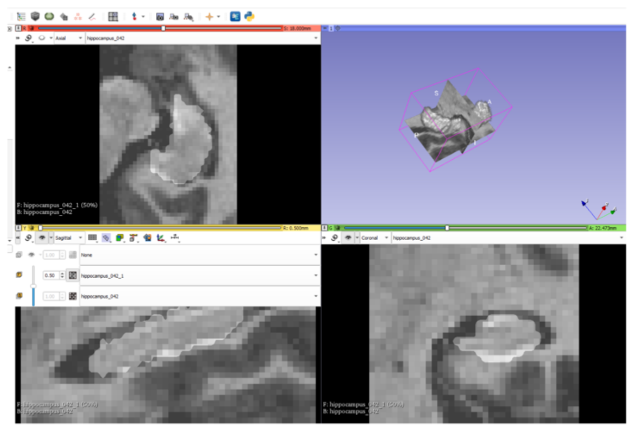
Hippocampal Volume Quantification in Alzheimer's Progression

We are using Hippocampus Dataset (Medical Decathlon competition) stored as a collection of NIFTI files, collecting MRI scans of different age, gender, and brain hemisphere, with one file per volume, and one file per corresponding segmentation mask. The original images here are T2 MRI scans of the full brain. As noted, in this dataset we are using cropped volumes where only the region around the hippocampus has been cut out. This makes the size of our dataset quite a bit smaller allowing us to have reasonable training times.

Intended Use

The algorithm is intended to use in determining the early stages of Alzheimer's disease progression through the monitoring of volume size of the hippocampus using the data mentioned above. However, there is one problem with measuring the volume of the hippocampus using MRI scans, though - namely, the process tends to be quite tedious since every slice of the 3D volume needs to be analyzed, and the shape of the structure needs to be traced. The fact that the hippocampus has a non-uniform shape only makes it more challenging. That is where our algorithm comes to scene.





Training Model and Performance

We used UNet Architecture for our algorithm, and we measured the performance after 10 Epochs (looking forward to optimize our training loss) using Tensorboard as showed below (Fig. 1.1):

Fig.1.1

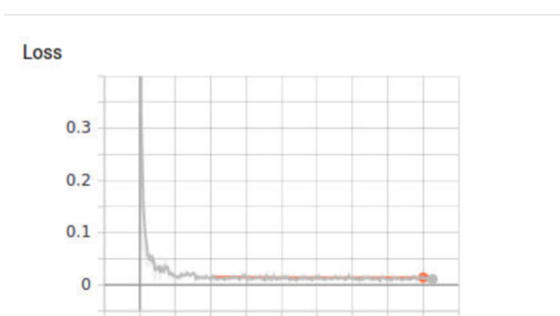


Fig.1.2 **Dice and Jaccard** Index
(for sample files – recorded for each record)


```
{
  "volume_stats": [
    {
      "filename": "hippocampus_044.nii.gz",
      "dice": 0.9033179869820607,
      "jaccard": 0.8236826867400115
    },
    {
      "filename": "hippocampus_381.nii.gz",
      "dice": 0.8990432200593863,
      "jaccard": 0.8166017380881031
    },
    {
      "filename": "hippocampus_251.nii.gz",
      "dice": 0.8118313404657017,
      "jaccard": 0.6832627118644068
    },
    {
      "filename": "hippocampus_221.nii.gz",
      "dice": 0.5810841460648674,
      "jaccard": 0.40952685421994883
    },
    {
      "filename": "hippocampus_337.nii.gz",
      "dice": 0.7961264016309888,
      "jaccard": 0.6613039796782387
    },
    {
      "filename": "hippocampus_225.nii.gz",
      "dice": 0.8633926656658804,
      "jaccard": 0.7596226415094339
    },
    {
      "filename": "hippocampus_229.nii.gz",
      "dice": 0.9217975371821525,
      "jaccard": 0.8549391871848117
    },
    {
      "filename": "hippocampus_101.nii.gz",
      "dice": 0.8718622300058377,
      "jaccard": 0.7728331177231565
    },
    {
```

Hippocampal Volume Quantification in Alzheimer's Progression

Clinical Impact Performance

We present a set of metrics for validating 3D image segmentation that were selected based on a literature review of papers in which 3D medical image segmentations are evaluated: Dice and Jaccard Index, both of the metrics measure the same aspects and provide the same system ranking. Therefore, it does not provide additional information to select both of them together as validation metrics. True Positive Rate (TPR), also called Sensitivity and Recall, measures the portion of positive voxels in the ground truth that are also identified as positive by the segmentation being evaluated. Analogously, True Negative Rate (TNR), also called Specificity, measures the portion of negative voxels (background) in the ground truth segmentation that are also identified as negative by the segmentation being evaluated. The impact of false positive and false negatives. However, under the clinical impact perspective, these two measures are not common as evaluation measures of medical image segmentation because of their sensibility to segments size, i.e. they penalize errors in small segments more than in large segments.

6. FDA Validation Plan

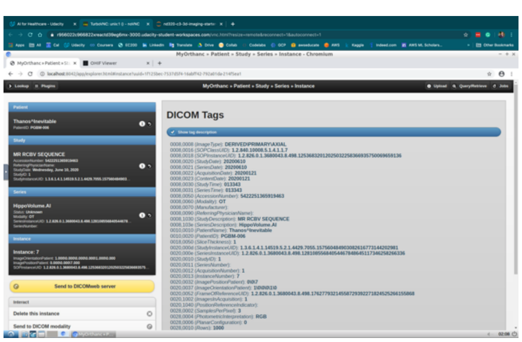
Patient Population Description for Validation Data Set

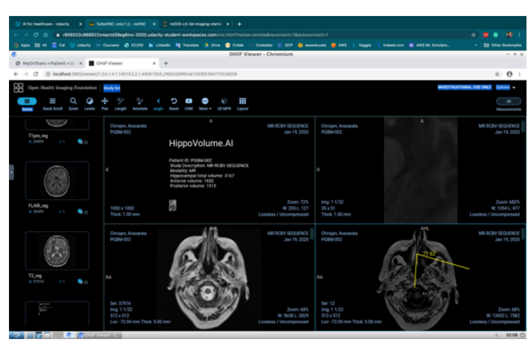
Ideal dataset (Niftii/Dicom files) would be collected from MRIs (body part: Head) ***where the hippocampus volume ranges between 2,200 and 4,500mm3)*** in similar ***gender and ages proportion***. ***No prevalence of other cerebral conditions (as tumors) preferably.*** A balanced Data Set Split, conserving real-life prevalence is also mandatory. Error in labeling might be a limitation if occurred.

Ground Truth Acquisition Methodology

We define **Silver Standard** as our ***optimal* ground truth acquisition methodology** since Hippocampal Volume Quantification in Alzheimer Progression is a challenging task that relies on the availability of experts. However, as we stated, this task is very demanding in relation to time taking and helping to visualize files at Imaging 3D software is a key factor, as well as the potential improvements for the algorithm in terms of segmentation (See Fig. 2 to illustrate examples, including the algorithm):

Fig. 2 (Screenshots with '**HippoVolume.AI'** numbers for **Study1**)





References

1. Udacity. AI for healthcare (Lesson 3) URL https://classroom.udacity.com/nanodegrees/nd320
2. Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool https://www.ncbi.nlm.nih.gov/
3. Calculator and report for HipoFit http://www.smanohar.com/biobank/calculator.html