

# Hippocampal Volume Quantification for Alzheimer's Progression:

## Segmentation with a U-Net Convolutional Neural Network

**Name:** Devika Pace

**Name of Device:** HippoCrop\_Unet

## Algorithm Description

### 1. General Information

#### **Intended Use Statement:**

The HippoCrop\_Unet AI algorithm is intended to aid physicians, namely radiologists, by automatic segmentation and quantification of hippocampal tissue in 3-D volume MRI renderings.

The algorithm is intended to help diagnose and track any progression of Alzheimer's Disease and other forms of Dementia in patients displaying symptoms of disease, including issues with short and long term memory as well as learning. Quantification will help guide disease management.

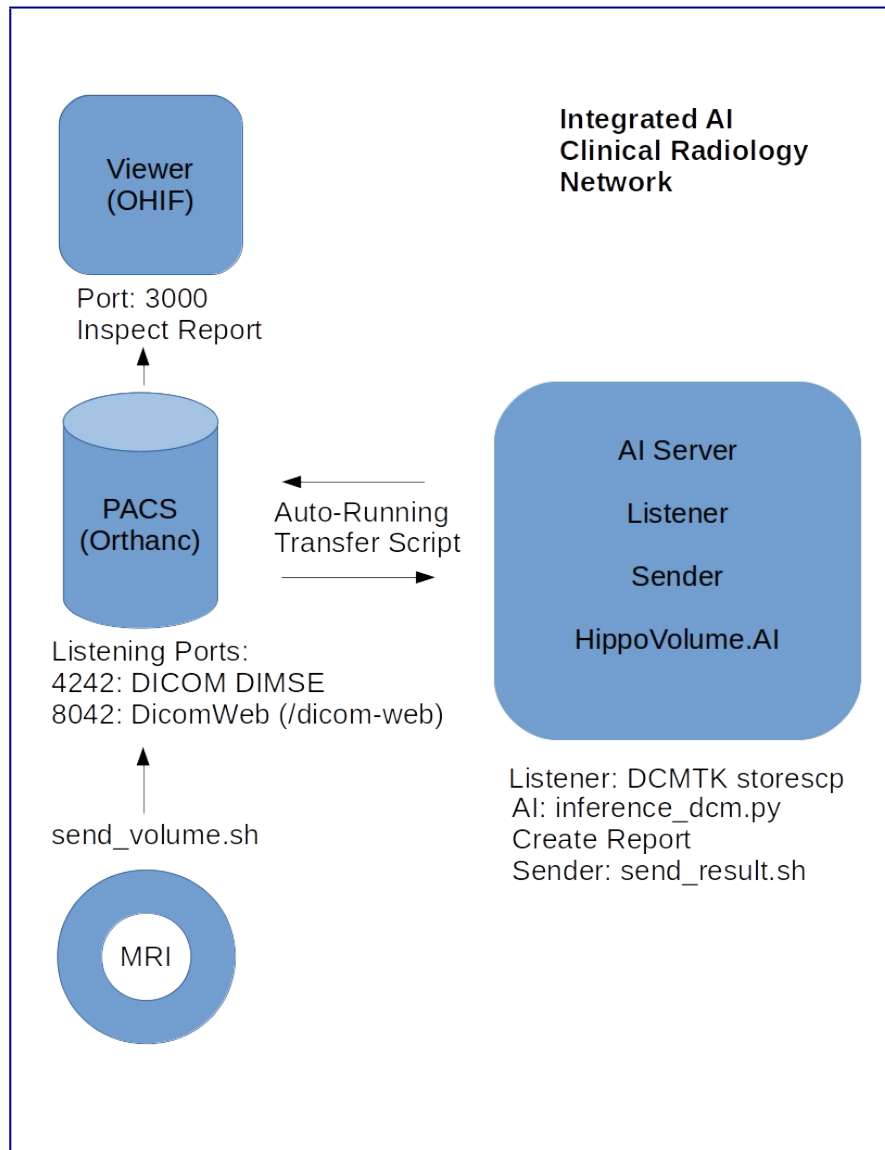
### 2. Algorithm Design and Function

#### **System Design:**

The HippoCrop\_Unet is an end-to-end AI system which features a machine learning algorithm that integrates into a clinical-grade viewer and automatically measures hippocampal volumes of new patients, as their studies are committed to the clinical imaging archive. An MRI volume will be pushed to a Picture Archive and Communication System (PACS) server over Digital Imaging and Communications in Medicine (DICOM) services. Each volume will be auto-transferred from the PACS server to an AI server, which will screen for volumes designated 'HippoCrop', subsequently running inference, and returning a results report to the PACS server for viewing through an Open Health Imaging Foundation (OHIF) Viewer.

The OHIF Viewer is a zero-footprint medical image viewer provided by the Open Health Imaging Foundation. It is a configurable and extensible progressive web application with out-of-the-box support for image archives which support DICOMweb. The PACS server may be simulated with Orthanc, a free, open-source, lightweight DICOM server.

Figure 1. AI Clinical Radiology Network:

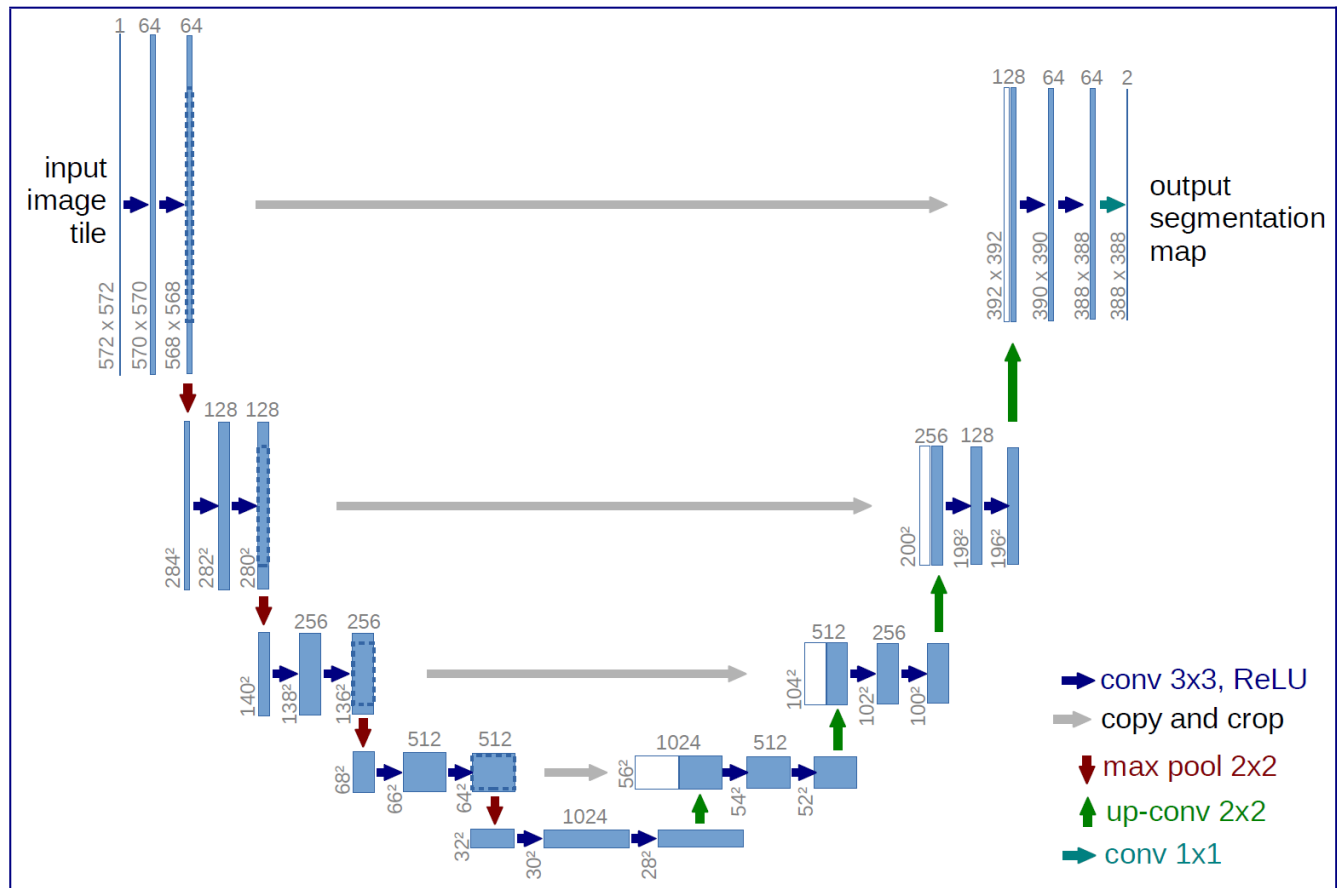


### U-Net Architecture:

We utilize the U-net CNN architecture as described by authors in ([U-Net: Convolutional Networks for Biomedical Image Segmentation, Ronneberger et al, May 2015](#)).

The U-Net differs from a standard CNN architecture in that it combines the location information from the downsampling path with the contextual information in the upsampling path. The purpose of this expanding path is to enable precise localization combined with contextual information from the contracting path. In the upsampling path, skip layers are added from contraction path, concatenating upsampled layer with the corresponding cropped feature map from the contracting path. The result combines localisation and context necessary to create good segmentation map. For a deeper understanding, see the webpage on U-net created by one of the authors of the original paper, Olaf Ronneberger ([here](#)).

Figure 2: U-Net Architecture. Source: U-Net: Convolutional Networks for Biomedical Image Segmentation, Ronneberger et al, May 2015.



### 3. Dataset

#### Collection of Training Data:

The hippocampus training data was gathered from the Medical Decathlon competition, found at <http://www.medicaldecathlon.com>. This dataset is stored as a collection of 260 Neuroimaging Informatics Technology Initiative (NIfTI) files, with one file per image volume, and one file per corresponding segmentation mask. The original images are T2 MRI scans of the full brain. This dataset utilizes cropped volumes where the region around the hippocampus has been cut out.

The following details regarding this dataset acquisition are described here: ([A large annotated medical image dataset for the development and evaluation of segmentation algorithms, Simpson et al, Feb 2019](#)). From the paper, 'The dataset consisted of MRI acquired in 90 healthy adults and 105 adults with a non-affective psychotic disorder (56 schizophrenia, 32 schizoaffective disorder, and 17 schizophreniform disorder) taken from the Psychiatric Genotype/Phenotype Project data repository at Vanderbilt University Medical Center (Nashville, TN, USA).' As well, 'Structural images were acquired with a 3D T1-weighted MPRAGE sequence (TI/TR/TE, 860/8.0/3.7 ms; 170 sagittal slices; voxel size, 1.0 mm<sup>3</sup>). All images were collected on a Philips Achieva scanner (Philips Healthcare, Inc., Best, The

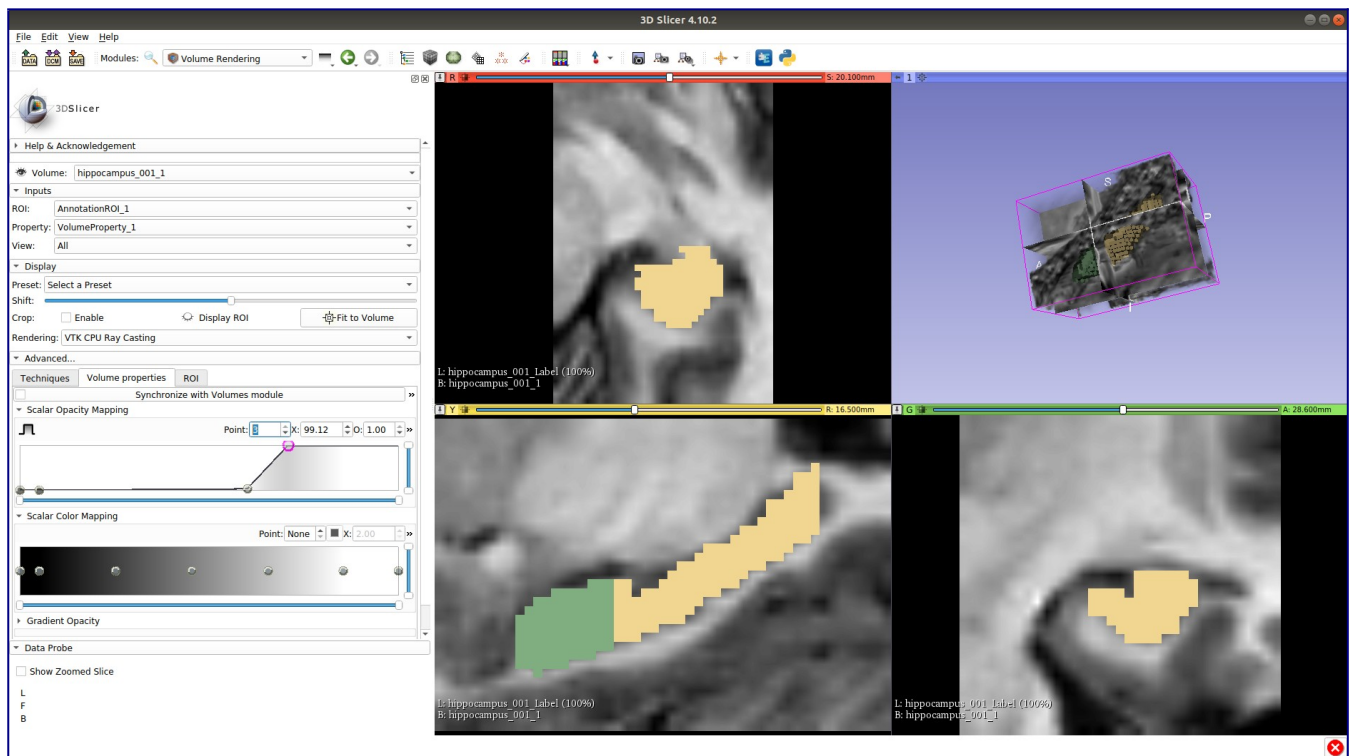
Netherlands). Manual tracing of the head, body, and tail of the hippocampus on images was completed following a previously published protocol (see reference). For the purposes of this dataset, the term hippocampus includes the hippocampus proper (CA1-4 and dentate gyrus) and parts of the subiculum, which together are more often termed the hippocampal formation. The last slice of head of the hippocampus was defined as the coronal slice containing the uncus apex. The resulting 195 labeled images are referred to as hippocampus atlases. Note that the term hippocampus posterior refers to the union of the body and the tail.'

Further, all data were made available online under Creative Commons license CC-BY-SA4.0. The data was reformatted from DICOM to NIfTI, an open standard supported by National Institutes of Health (NIH), and as such doesn't require proprietary software. This serves to ensure open access, consistency and interoperability.

### Labeling of Training Data:

Medical image segmentation seeks to extract anatomical regions of interest from a medical image or series of images. Semantic segmentation aims to automatically associate each image pixel with a classification label. The brain tissue was classified as non-hippocampal (background), anterior hippocampal, or posterior hippocampal.

Figure 3: Sample slices in 3D-slicer showing axial, sagittal, and coronal views:



*Figure 4. Sample volume depicting Sagittal Slices:*

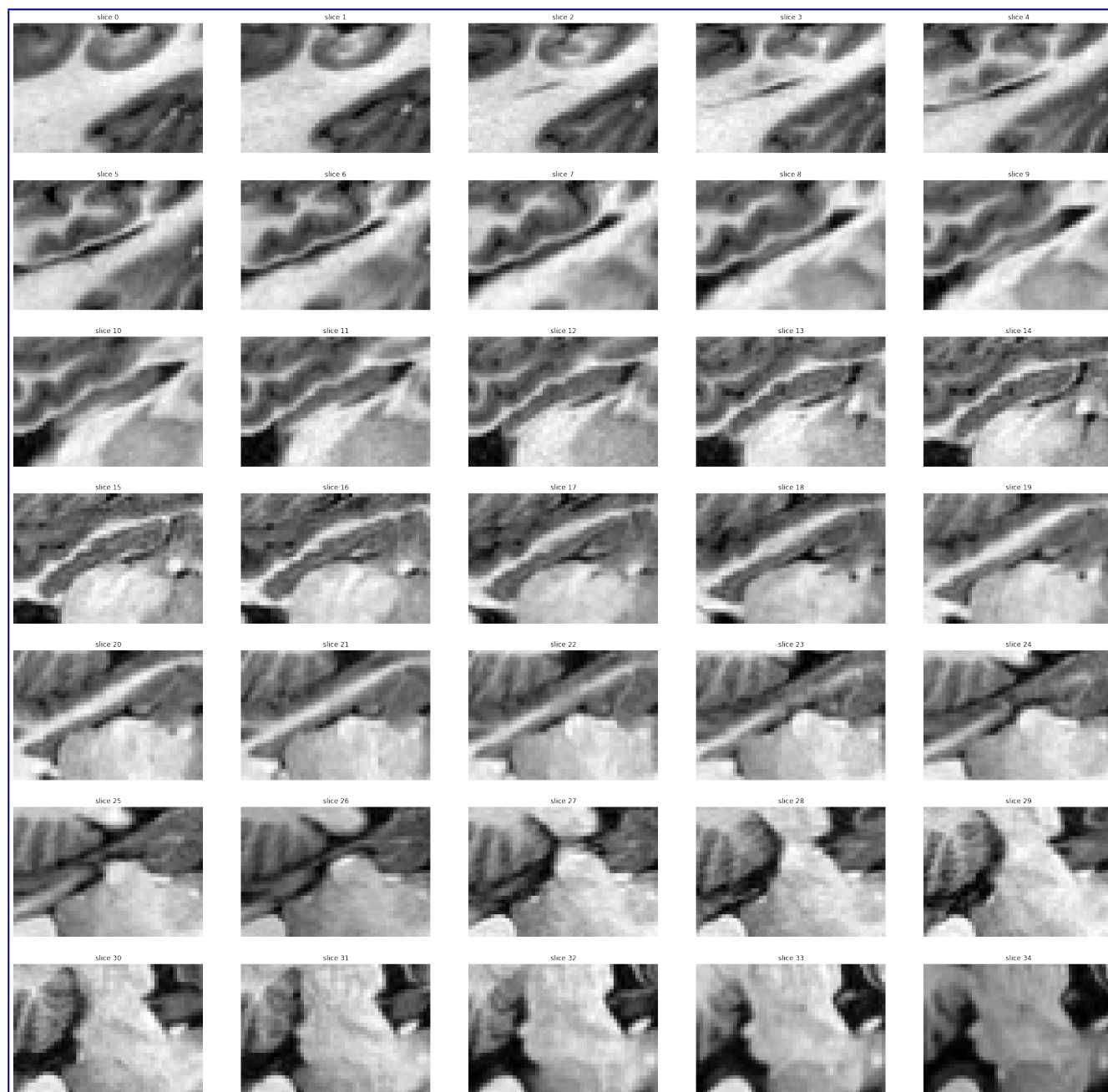
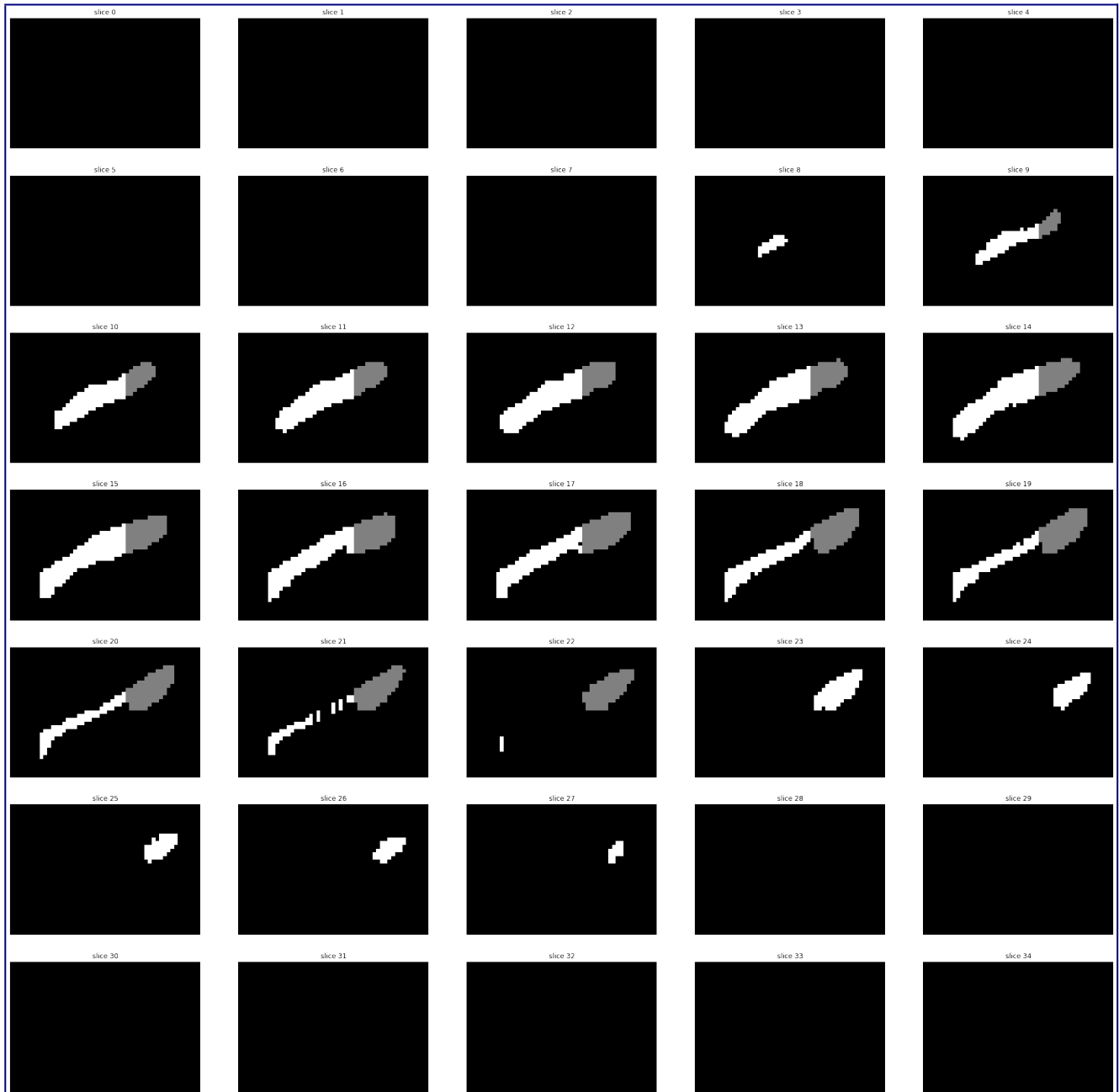


Figure 5. Sample hippocampal segmentation showing anterior and posterior areas:



## Splitting the Dataset

The algorithm was trained by splitting the dataset into train, validation, and test sets. The data was randomly split resulting in 70% of data for training, and 15% for for each validation and test sets.

## 4. Algorithm Training

We performed out training with the following hyperparameter configuration:

Loss Function: CrossEntropyLoss

Optimizer: Adam

Learning Rate: 0.0002  
Batch Size: 8  
Patch Size: 64

Preliminary results are visualized using Tensorboard in the following figures.

Figure 6. TensorBoard Training Loss:

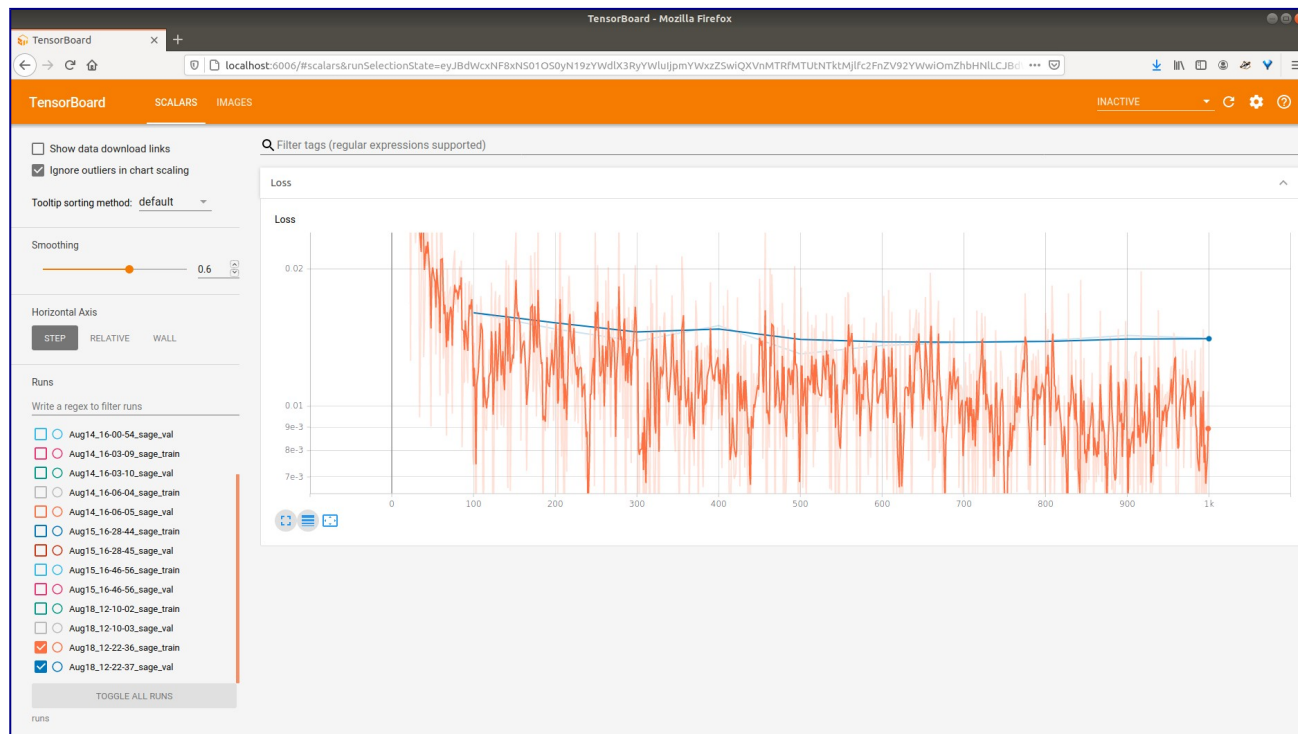


Figure 7. TensorBoard Images and Masks:

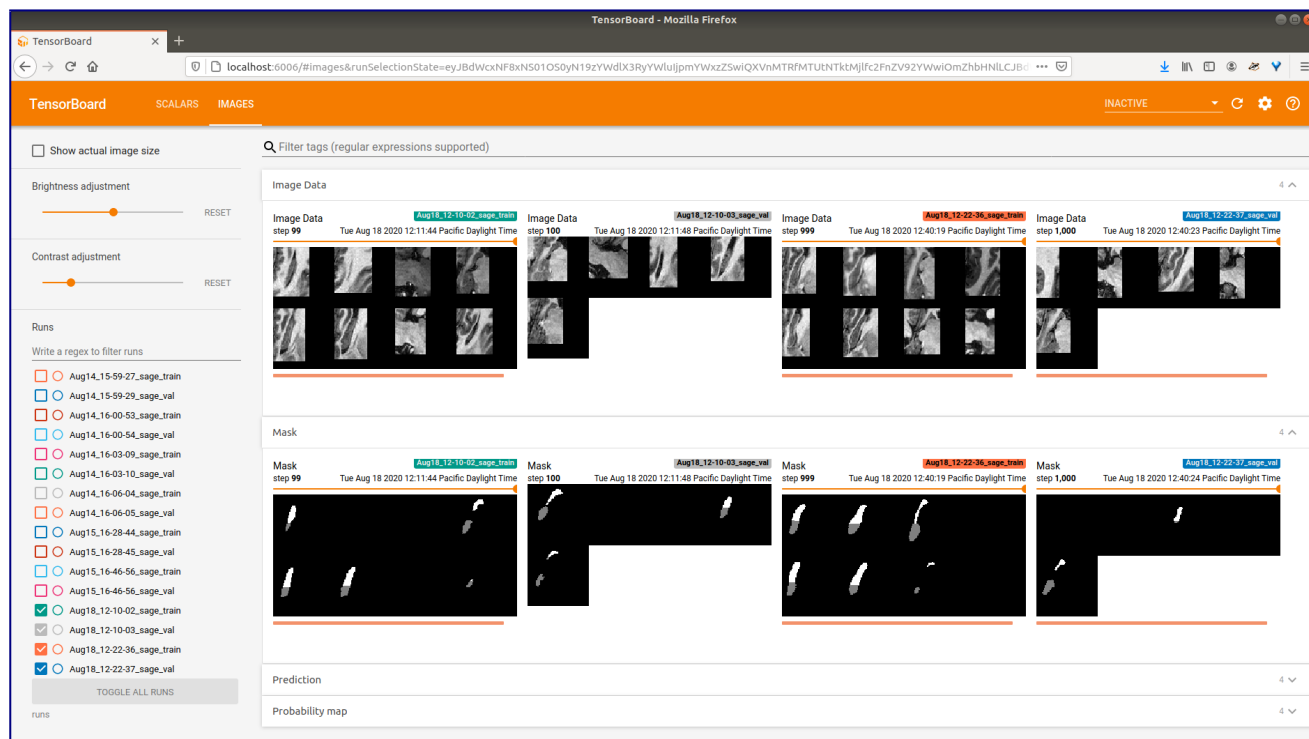


Figure 8. TensorBoard Masks and Predictions:

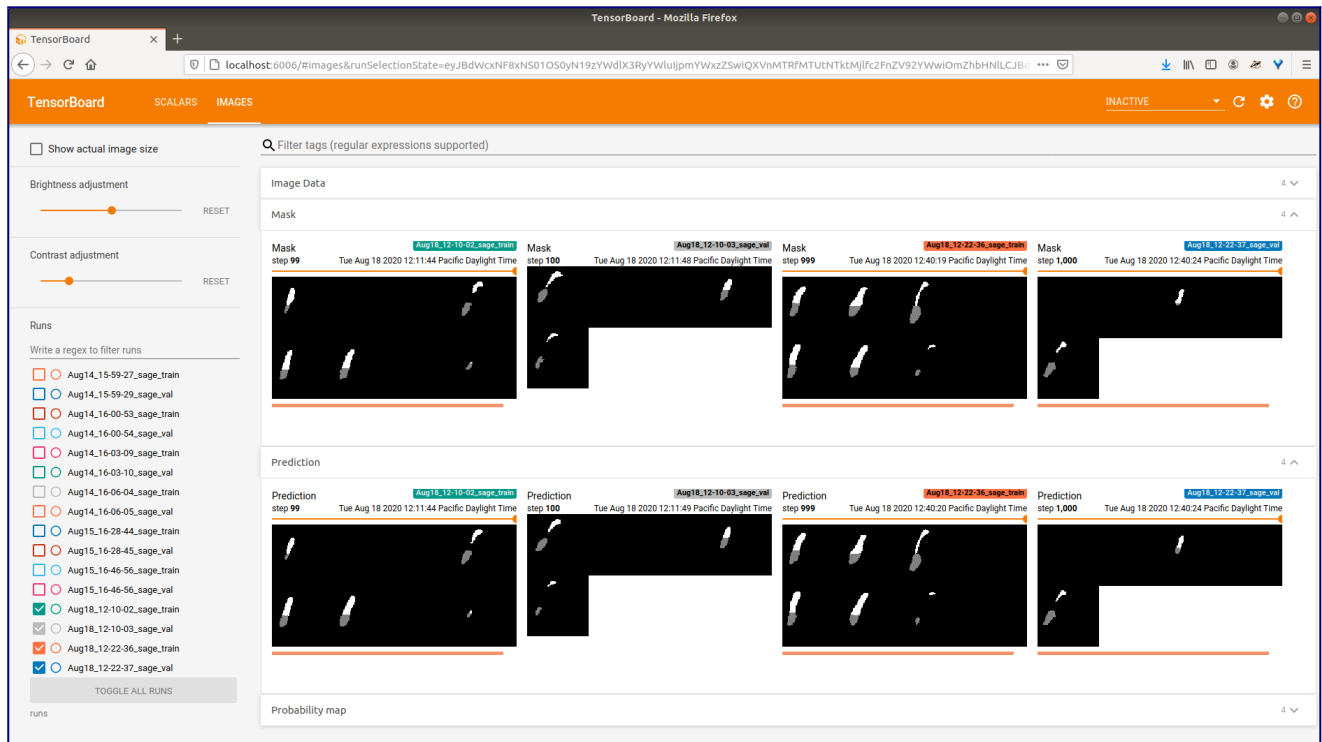
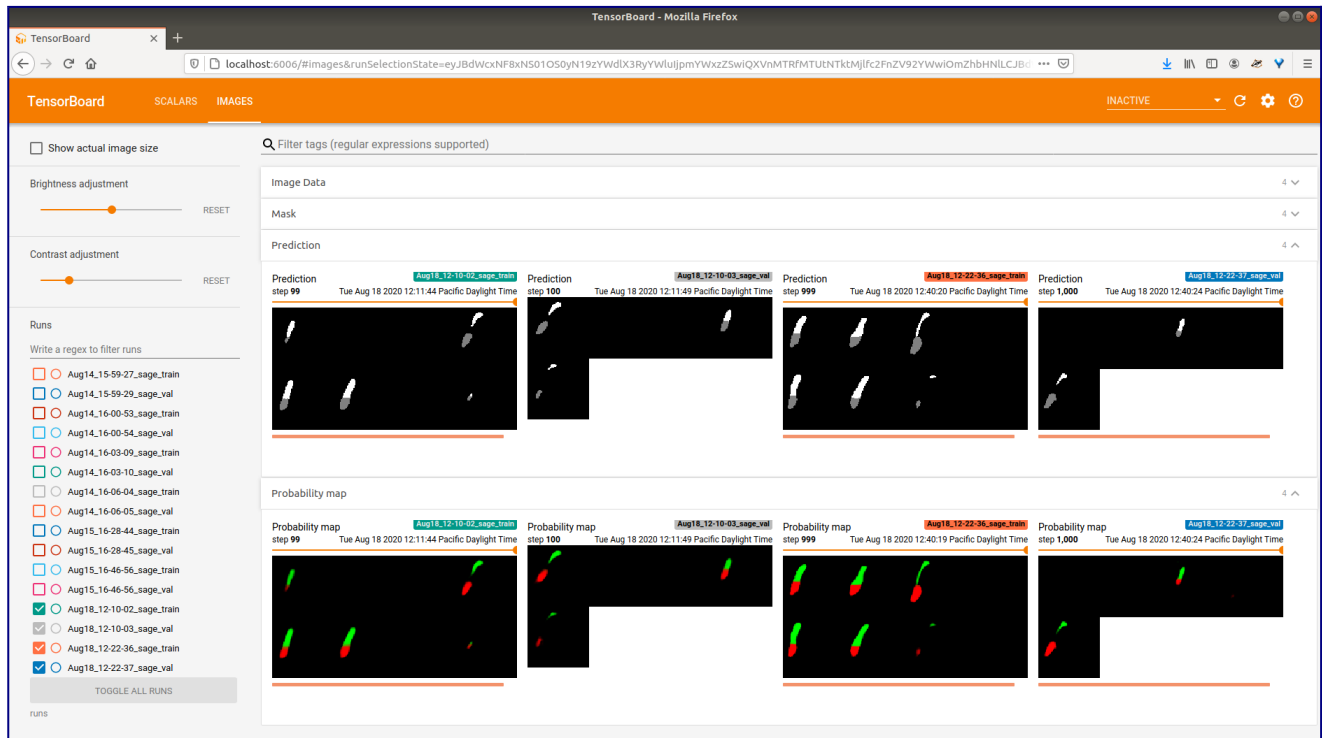


Figure 9. TensorBoard Predictions and Probability Maps:





## Performance Metrics

The final training performance from the algorithm was measured with the Dice Similarity Coefficient and the Jaccard Index. These are two formulas that measure how well our predictions overlap with our ground truth.

Given two volumes X and Y,

The Dice Similarity Coefficient (DSC) can be derived as a 2 fold factor of the volume intersection over the sum of volumes:

```
X = (X > 0)
Y = (Y > 0)
intersection = np.sum(X*Y)
volumes = np.sum(X) + np.sum(Y)
if volumes == 0:
    return -1
```

```
DSC = 2.*float(intersection) / float(volumes)
```

The Jaccard Index (JI) can be derived as the intersection of the volumes over the union:

```
X = (X > 0)
Y = (Y > 0)
intersection = np.sum(X*Y)
union = np.sum(X) + np.sum(Y) - intersection
```

```
if union == 0:
    return -1
```

```
JI = float(intersection)/float(union)
```

## Training Results

Results for training are seen below in final recorded json file.

*Figure 9. JSON Performance Metrics:*

```
{
  "volume_stats": [
    {
      "filename": "hippocampus_318.nii.gz",
      "dice": 0.8694209891435465,
      "jaccard": 0.7690050680181382
    },
    {
      "filename": "hippocampus_150.nii.gz",
      "dice": 0.9107251675807434,
      "jaccard": 0.8360839160839161
    },
    {
      "filename": "hippocampus_068.nii.gz",
      "dice": 0.9243956797531288,
      "jaccard": 0.8594198278610137
    },
  ],
}
```

```

{
  "filename": "hippocampus_393.nii.gz",
  "dice": 0.8964670577001773,
  "jaccard": 0.8123609394313968
},
{
  "filename": "hippocampus_130.nii.gz",
  "dice": 0.9117997616209773,
  "jaccard": 0.8378970427163198
},
...

{
  "filename": "hippocampus_250.nii.gz",
  "dice": 0.9193620509763005,
  "jaccard": 0.8507586206896551
},
{
  "filename": "hippocampus_328.nii.gz",
  "dice": 0.9318269354217487,
  "jaccard": 0.8723557692307692
},
{
  "filename": "hippocampus_007.nii.gz",
  "dice": 0.9263341067285383,
  "jaccard": 0.8627768773635872
},
{
  "filename": "hippocampus_042.nii.gz",
  "dice": 0.8322307488753639,
  "jaccard": 0.7126671198731022
},
{
  "filename": "hippocampus_249.nii.gz",
  "dice": 0.9279416235937975,
  "jaccard": 0.8655700510493477
}
],
"overall": {
  "mean_dice": 0.9079577116344708,
  "mean_jaccard": 0.8324087577513897
}
}

```

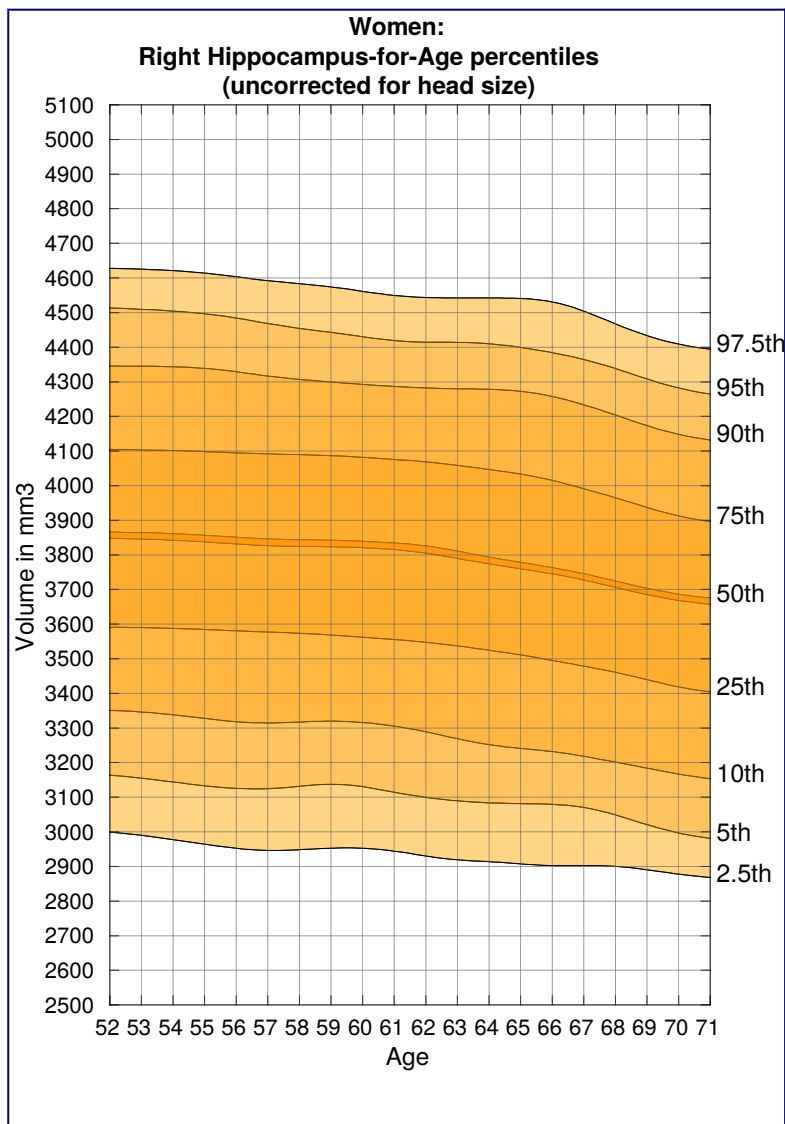
## 5. Algorithm Performance Standard

### Clinical Significance and Ground Truth:

Resulting segmentation and quantification by the AI algorithm will serve to guide radiologists given the performance metric of correct overlap as described above. Radiologists should be the final arbiters to verify the results of the algorithm.

For reference guidance, ([research from the UK Biobank](#)) has shown normative ranges for hippocampal volume per gender and age and one would expect ranges of measurement to be commensurate in range given any factoring for necessary disease states.

Figure 10. Hippocampal Quantification Metrics: Hippocampal volume across age: Nomograms derived from over 19,700 people in UK Biobank, Nobis et al, 2019



### Device Limitations:

#### Algorithmic Limitations:

**Patient History and Demographics:** This algorithm was not specifically tested on those with a prior history of brain illness, including neurological illness, head injury, or substance abuse dependence. Racial or ethnic demographics as well are not noted in the dataset so any differential performance in this regard is unknown.

#### Computational Limitations:

Model training as well as testing and inference was performed utilizing an Nvidia RTX 2060 GPU with 6GB of onboard RAM. It is suggested to utilize HippoCrop\_Unet on a system with comparable hardware.