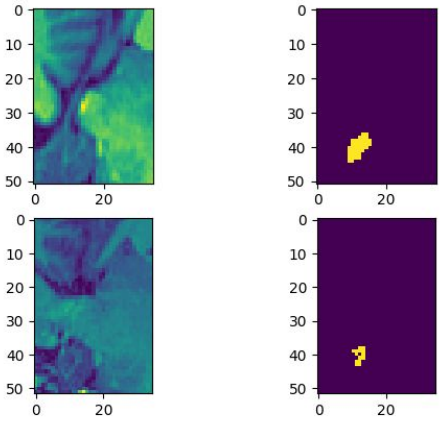
CLINICAL EXPLANATION OF ALGORITHM

Alzheimer is a neurodegenerative disorder that tends to affect the hippocampal volume in the brain. In 2015, there were approximately 29.8 million people worldwide with 4–5% reporting early onset Alzheimers. In 2015, the condition resulted in about 1.9 million deaths as per statistics(https://en.wikipedia.org/wiki/Alzheimer%27s_disease). While the medical community has advanced leaps and bounds to treat effectively challenging diseases, Alzheimers is also a financially draining disease (https://www.sciencedirect.com/science/article/abs/pii/S0755498205838825?via%3Dihub). It is hence important to have automated systems in place for those who either cannot afford personalised examination and treatment or those who have no access to advanced healthcare facilities.

Quantification of hippocampal volume has been a potential biomarker for Alzheimer's disease(https://www.ncbi.nlm.nih.gov/pubmed/19251758). Hence, segmentation of the hippocampus from MRI scans can be an indicator for Alzheimer's disease. This is used in the following manner. On the left hand side is the original image obtained from a patient scan and on the right hand side(called the mask) highlights the hippocampal area and volume abnormalities are flagged for further inspection.



The architecture takes in the 3 dimensional image, slices it, accounts for every small aspect of the image and localises the area of abnormality by classifying and distinguishing it.
The results hence obtained are as follows:

Dice scores: Since the hippocampal area is uneven, we need to account for the overlap in the region in order to report accuracy. So, dice score is defined by size of the overlap of the two segmentations divided by the total size of the two objects.
Performance: ~0.90

Jaccard Score: Is defined by the ratio of area of overlap between predicted segmentation and annotated mask values by area of union between the two values.
Performance: ~0.80

Additional Statistics:
Specificity: The ratio of true negatives to the total rate of true negatives and false positives.

Sensitivity: The ratio of true positives to the total rate of true positives and false negatives.

TECHNICAL REPORT

Training process requirements:

The model was trained on CPU, intel i7 four core processor with 16GB RAM.

Training process suggestions:

Model architecture: Experimenting with Variants of UNet is advised.
Vanilla UNet is as described here: https://arxiv.org/abs/1505.04597
ChannelUNet: https://www.frontiersin.org/articles/10.3389/fgene.2019.01110/full
H-DenseUNet: https://arxiv.org/pdf/1709.07330.pdf

Data pipeline: The pipeline for Electroencephalography has been documented by Harvard here: https://www.frontiersin.org/articles/10.3389/fnins.2018.00097/full. Similar fine tuning methodologies are likely to be relevant here. Since biomedicine is a safety critical application, introducing transparency in the pipeline and prediction methodologies is preferable.

Loss functions: Metrics can be optimised as per https://arxiv.org/pdf/1911.01685.pdf.

Data augmentation: pixel level data augmentation with GAN can help.
https://ieeexplore.ieee.org/document/8683590
The techniques can be improved as stated here: https://openreview.net/pdf?id=rkBBChjiG
Using one shot learning's learned transformations for augmentation could work too.
http://www.mit.edu/~adalca/files/papers/cvpr2019_brainstorm.pdf
Other types include: local shifts, color/projection transformations, compressions and copying, could be other augmentation techniques to explore.

What kind of data augmentations would NOT add value?
Based on this data: https://dl.acm.org/doi/pdf/10.1145/3341016.3341020
Local copying (copy data with different backgrounds) doesn't seem to yield substantial results.