HippoVolume.AI - Validation Plan

HippoVolume.AI is a piece of medical software centered on a deep neural network model that is trained to segment MRIs for hippocampi. The software uses the segmentation to compute hippocampus volume and generates a report that can be integrated into a suitable clinician's workflow.

Intended Use

The intended use of this software is to assist a radiologist in measuring hippocampus volume while reading a brain MRI. It must be paired with out from the tool HippoCrop in order to function properly.

Indications for use: One promising application of this software is in the diagnosis and progression-tracking of Alzheimer's disorder. Hippocampus volume has been shown to be useful for this; see [1], [2], and [3].

Training Data

The training data was obtained from a repository at Vanderbilt University Medical Center. Patients were recruited from the community in and surrounding Nashville, TN. Some patients were healthy and some had a psychotic disorder, but all were free from significant neurological illnesses or head injuries. See [5] for more details.

Even though the training images were all obtained via the same medical imaging device (a Philips Achieva scanner), we would like to claim that our software functions at acceptable levels on output from any brain MRI as long as there is sufficient contrast to make out the relevant structures.

For later validating the model's performance, we will request more brain MRIs from a variety of devices. We will request that the images be from patients who do not have any significant neurological illnesses or head injury, and we will request that the contrast be sufficient to make out the hippocampi. We plan to use the HippoCrop tool to generate input for our model.

Ground Truth

Ground truth for training was obtained by manual tracing of the posterior and anterior parts of the hippocampus. The tracing was done following the protocol found in [4].

For later validating the model's performance, we will request ground truth labels manually traced by qualified specialists according to the protocol found in [4].

Model Performance

The current model's perforance was assessed using a Dice coefficient. We obtained a mean Dice coefficient of 0.9 on validation data with the Dice coefficient not going below 0.8.

For later validating the model's performance, we will

- *apply HippoCrop to the MRI data obtained as described above,*

- *generate predicted segmentations by applying our model,*

- *and compute the Dice coefficient comparing the predicted segmentation to the ground truth obtained as described above.*

If the mean dice coefficient is above 0.8 then the software can be considered to be performing acceptably.

References

[1] Edit Frankó, Olivier Joly, and Alzheimer's Initiative. Evaluating alzheimer's disease progression using rate of regional hippocampal atrophy. *PLoS ONE*, 8:e71354, 08 2013.

[2] C.R. Jack, R.C. Petersen, Y.C. Xu, P.C. O'Brien, G.E. Smith, R.J. Ivnik, B.F. Boeve, S.C. Waring, E.G. Tangalos, and E. Kokmen. Prediction of ad with mri-based hippocampal volume in mild cognitive impairment. *Neurology*, 52(7):1397–1397, 1999.

[3] Seok Moon, Boram Lee, and Young Choi. Changes in the hippocampal volume and shape in early-onset mild cognitive impairment. *Psychiatry investigation*, 15, 04 2018.

[4] J.C. Pruessner, Li Min, Serles WW, Pruessner MM, Louis Collins, Kabani NN, Sonia Lupien, and Alan Evans. Volumetry of hippocampus and amygdala with high-resolution mri and three-dimensional analysis software: Minimizing the discrepancies between laboratories. *Cerebral cortex (New York, N.Y. : 1991)*, 10:433–42, 05 2000.

[5] Amber L. Simpson, Michela Antonelli, Spyridon Bakas, Michel Bilello, Keyvan Farahani, Bram van Ginneken, Annette Kopp-Schneider, Bennett A.

Landman, Geert Litjens, Bjoern Menze, Olaf Ronneberger, Ronald M. Summers, Patrick Bilic, Patrick F. Christ, Richard K. G. Do, Marc Gollub, Jennifer Golia-Pernicka, Stephan H. Heckers, William R. Jarnagin, Maureen K. McHugo, Sandy Napel, Eugene Vorontsov, Lena Maier-Hein, and M. Jorge Cardoso. A large annotated medical image dataset for the development and evaluation of segmentation algorithms, 2019.