For our method, we pre-train a state-of-the-art deep neural network segmentation model DeepLab-v3+ on T1 MRI scans from the ADNI and OASIS public datasets to predict FreeSurfer automated labels. We then fine-tune on the T1-weighted, T1-IR, and T2-FLAIR scans using the training labels, and finally fine-tune using the testing labels.

- Is your algorithm automatic or semi-automatic? Describe the required user input, and the average time spent per scan, for semi-automatic algorithms.
  - The user inputs are T1 scans only and the average time per scan is less than 15 seconds.
- Which sequences are used by your algorithm? Only the thick-slice T1-weighted scan, or the IR, FLAIR and thin-slice T1 scan as well?
  - We have trained models to only take the thick-slice T1-weighted scan as an input, or to use all three modalities (thick-slice T1-weighted, IR, and FLAIR scans); for this evaluation we use only T1 scans.
- List the overall structure of the algorithm in a step-wise fashion and describe each step of the algorithm in detail. Include pre- or post-processing steps, when required.
  - We first normalize the entire brain scan volume to have mean 0 and standard deviation 1. We then train fine-tune our pretrain DeepLab-v3+ model using transverse 2D patches from the MRBrainS18 dataset
- Mention limitations of the algorithm. Was the algorithm specifically designed to segment only certain types of scans (e.g. scans of healthy volunteers or patients with specific pathology)? Was it optimized to work for scans with thick or thin slices? Was your algorithm specifically designed for a certain goal (e.g. volume measurement to quantify atrophy in patients with Alzheimer's disease)? Etc.
  - This algorithm was pre-trained on data from both healthy patients as well as aged patients with dementia. The goal was to build a model that would segment new scans robustly even in the case of patient brain atrophy. We hope to extend this by adding in additional data from patients with other pathologies, for example tumors.
- Was the algorithm trained with example data other than the training data provided for the MRBrainS challenge? If so, describe the characteristics of the training data.
  - Yes, the model was pre-trained with around 7 K input T1 scans from the ADNI and OASIS datasets to predict corresponding FreeSurfer segmentations. Importantly, the pre-training was done for less than one epoch, and could be continued further for improved results.
- Which labeled structures were used to fine-tune your algorithm (e.g. cortical gray matter and basal ganglia were used separately or combined labeled as gray matter)?
  - We used Cortical gray matter, Basal ganglia, White matter, White matter lesions, Cerebrospinal fluid in the extracerebral space, Ventricles, Cerebellum, and Brainstem labels to fine-tune at first, before fine-tuning on the final test classes.
- If the algorithm has been tested on other databases, you could consider including those results.
  - We have tested this model on the Brats15 tumor detection dataset and will present results in a later publication
- What is the average runtime of your algorithm, and on which system is this runtime achieved?
  - The algorithm takes about 15 seconds to segment an entire brain volume on an GeForce 1080 Ti GPU.