

Introduction

Stroke is the leading cause of adult disability worldwide, with up to two-thirds of individuals experiencing long-term disabilities. Large-scale neuroimaging studies have shown promise in identifying robust biomarkers (e.g., measures of brain structure) of stroke recovery. However, analyzing large datasets is problematic due to barriers in accurate stroke lesion segmentation. Manually-traced lesions are currently the gold standard for lesion segmentation, but are labor intensive and require anatomical expertise. While algorithms have been developed to automate this process, the results often lack accuracy. Newer algorithms that employ machine-learning techniques are promising, yet these require large training datasets to optimize performance.

About data

We have ATLAS (Anatomical Tracings of Lesions After Stroke), an open-source dataset of 220 T1-weighted MRI scans with manually segmented lesions and metadata. We note that this data is relatively diverse, given the data collection across 9 research sites worldwide.

If a particular subject had multiple lesions, for each additional lesion, there would be three additional lesion masks (e.g., *_LesionSmooth_stx.nii.gz, *_LesionSmooth_1_stx.nii.gz). In general, lesions were ranked based on size where the largest lesion, *_LesionSmooth_stx.nii.gz, was considered the main lesion. And The rest of the smaller masks, *_LesionSmooth_1_stx.nii.gz, was considered the sub mask.

This large, diverse dataset can be used to train and test lesion segmentation algorithms and provides a standardized dataset for comparing the performance of different segmentation methods.

The data that we use for our models:

We have 3 models that their results are in 3 classes: background, main mask, sub mask.

The best result (Dice coefficient) of test dataset for our models on 44 patients:

models	background	Main mask	Sub mask

Resnet	0.998	0.515	0.0251
segnet	0.996	0.409	0.000

The highest accuracy of Resnet model for main mask and sub mask are 85.7% and 66.4%, respectively

The highest accuracy of segnet model for main mask and sub mask are is 77.5% and 0.0%, respectively