Research Proposal Group 3 – Multi-institute training

Goal:

The purpose of our multi-institute training task is to determine if training on a single, or a variation of data from various institutes, will result in better training performance. This will be performed by applying the model to a train/test sampled from each institute independently, and from a combination of these institutes. Furthermore, we suggest training the model on one institution and then perform evaluation on another institution. This will give information about how well the model is able to generalize on new data. Thereby determining if data from a certain institute holds more promise for wide-spread applicability. We will train on each institute and then separately evaluate on the data of both of the other institutes. This could also provide us with a minimum base-line for training performance.

We would also like to investigate the effect of compounding the training process, and investigating if there is an effect on the order of institutes trained, to the eventual performance. To do so, we will be utilizing an 80% training and 20% validation set, consistent throughout the study. Even when training on multiple institutes, the cumulative amount of training and test data should be inter-operable across the different experiments, while maintaining the class balance between the two sets in all experiments.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Test/Train | Utrecht | Ams | Sing | Utr/ams | Ams/sing | Utr/sing | Utr/ams/sing |
| Utrecht |  |  |  |  |  |  |  |
| Ams |  |  |  |  |  |  |  |
| Sing |  |  |  |  |  |  |  |
| Utr/Ams |  |  |  |  |  |  |  |
| Ams/sing |  |  |  |  |  |  |  |
| Utr/sing |  |  |  |  |  |  |  |
| Utr/ams/sing |  |  |  |  |  |  |  |

Differences between the data should be kept in mind. So, when we collect data for training/ testing, we need to ensure that the number of patients is the same. We will concede that the number of images between the institutions will be different. Keep that in mind when we explain training/testing performance. We should maintain the number of slices per institution in order to best maintain the integrity of the MRI data/technique for each institi. And thereby determine which institi. Has the better and more generalizable MRI method for WMH segmentation. This will be done by controlling for the number of patients. Since the number of slices per image is nothing but a proxy representative of the MRI technique, we will best represent the data of each MRI technique by considering the totality of the images per patient, instead of performing quality control on a slice-by-slice level.

For single institi. training: 16 training and 4 test.

Two instit… from each instit. 8 training and 2 test.

\*3 instit….from each instit. 5 training and 1 test.

Note. Number of slices per patient will differ based on specific institution. However we only control for number of patients. Also, make sure training/test set is the same across the different experiments. Do this by defining which images we will use to ensure that the same patients are being considered when we compare different institutions.

Adjust x and y dimensions for slices. Keep the number of slices. For picking which slices… white/black threshold. Do we normalize images?

Hypothesis:

It is expected that the data from Singapore will be more different form that of Utrecht and Amsterdam, simply by taking proximity and assuming collaborations in methodology by virtue of this proximity. We suspect that evaluating on Utrecht will give better results when training on Amsterdam than when training on Singapore.

Notes/questions to be considered:

* When training in all institutes compared to just one institute, we will ensure that the training datasets are of similar size
* Is the mask significantly different between institutions?
* Federated-learning is a similar challenge to multi-institute training. You take the model to different institutes instead of asking from data from different institutes. Might be too involved to apply, will only be looked into if we have enough time.