Research Plan Group 3 – Multi-Institute training

**Goal**

The goal is to see how the performance of the model on different test sets varies based on training set used when both the training and test set is comprised of data from one, two or three institutions.

**Data**

Since we are using data from different institutions we will normalize the pixel size in the x and y direction of the slices using zero-padding. The number of slices per patient will not be altered even though different institutions provide a different number of slices per patient. We maintain the number of slices per institution in order to best maintain the integrity of the MRI data/technique for each institute and thereby determine whether it is possible to train a model that generalizes well between different institutes.

The data will be manually split into a training and test set based on the degree of WMH present in the scan. This will result in both the training and the test set for each institution having samples with a relatively high amount of WMH and a relatively low amount of WMH. When a patient is decided to be used for testing it will always be used for testing and never for training.

**Notes from Matthijs:**

* The model is made in such a way that it work for the Utrecht data and size.
* Look into sliding window inference.
* If we are applying zero padding do it for all the images
* Usually do a random split meaning it should roughly be randomly distributed
* Don’t use metric just separate randomly. Use zero padding to make sure all the sizes are the same.

**Datasets**

The table below shows how the different training and test sets are comprised from the data that is available.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Train set | | | Test set | | |
| Set name | Utrecht | Amsterdam | Singapore | Utrecht | Amsterdam | Singapore |
| Utrecht | 16 | 0 | 0 | 4 | 0 | 0 |
| Amsterdam | 0 | 16 | 0 | 0 | 4 | 0 |
| Singapore | 0 | 0 | 16 | 0 | 0 | 4 |
| Utr\_Ams | 8 | 8 | 0 | 2 | 2 | 0 |
| Utr\_Sing | 8 | 0 | 8 | 2 | 0 | 2 |
| Ams\_Sing | 0 | 8 | 8 | 0 | 2 | 2 |
| Utr\_Ams\_Sing | 5 | 5 | 5 | 1 | 1 | 1 |

* Possibly also try a run with all the data

The eight patients used from Utrecht dataset in the Utr\_Ams and Utr\_Sing dataset are the same eight patients. This is also the case for the Amsterdam and Singapore patients in the other two institutions datasets. The five patient from each institution in the three institution training set are chosen from the eight patients of the reduced dataset used for two institution training. A similar approach is used when creating the two and three institution test datasets.

The validation set is automatically created from the training set using train\_test\_split with a 0.1 split. This split results in two patients per dataset ending up in the validation set. The data is split at a patient level in order to prevent slices from one patient ending up in the training and validation set.

**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 |  |  |  |  |  |  |  |

There are seven different datasets that will be used to train a model which are then each tested on seven different test datasets.

**Model**

The U-net model provided will be used without any alterations. When training the model we will monitor the loss curve to see if training went well.

Data loaders

The dataloading part of the provided code is changed. There is now a seperate data loader for the training and test set which allows us to load the train and test data from different directories after it is divided into the different datasets. It also allows for an easy change of test data in order to quickly test a trained model on a variety of test dataset.

Because of some nan values present in the dataset from the Amsterdam institute the array of the images are subjected to the np.nan\_to\_num() to set the nan values to zero.

ptl.seed\_everything() is used to set the seed for everything to 15 in order to have the same randomness for training of all the models. This results in not having a different initialization when training the different models which improves the comparability of the results.

The optimizer used is the adam optimizer and the dice loss function is used for the loss.

**Results**

The Dice loss achieved on the test dataset will be used to evaluate the results.