HC701 – Assignment # 3

Due on 11:59pm, 24/04/2023

In this assignment, you must implement the following task and submit the code along with a brief report summarizing your findings. You must submit a PDF for your work. Submitting the code is highly recommended.

Task 1: Liver Tumor Segmentation (10/10 marks)

This dataset includes 130 CT scans along with segmentation masks for liver and liver tumor. It was part of MICCAI and ISBI 2017 challenges. For this task, we changed the data and file structure, so you must use the split of training and testing data from Moodle. You cannot use the dataset splits on Kaggle. Furthermore, you **must not** use any pre-trained model from the Kaggle or other online sources. You can only use random initialization or ImageNet pretrained models.

Task 1.1 (4/10 marks)

Using this data, train a 2D Convolutional Neural Network (CNN) to segment the **liver and liver tumor** while fulfilling the following:

- You can use any architecture.
- You can choose any data augmentation techniques you want.
- You must use a maximum of 50 epoch to train one experiment.
- You must use random initialization or ImageNet initialization only.
- Explore different measures of performance (loss) during training e.g., Validation Accuracy, Validation Dice score, Validation IoU, Validation Cross Entropy, weighed CE, focal loss, Hausdorff distance, etc.
- The maximum number of different experiments you are allowed to run is 5.
- You are encouraged to use a validation set from the training set.

Report the solutions you are proposing (Architecture, main hyper-parameters, augmentations, performance measures you used) for the five experiments.

Task 1.2 (2/10 marks)

Apply the trained models on the testing set. Report Dice and Jaccard on the 5 experiments (mean and standard deviation).

Task 1.3 (2/10 marks)

Explain why you think the top performing experiment has achieved better results than the other 4 experiments.

Task 1.4 (2/10 marks)

Describe how you would go about deploying the best model you trained for two hospitals in a real-life scenario. Each hospital has multiple CT scanners and you do not have to address any specific regulatory requirements. You should support your deployment proposal with a flow chart that shows how the data will flow and where each data processing step will happen. You should address one of the following areas:

1. Collaborative and Distributed Learning.

2. Continual Learning.
3. Domain Generalization.