Lab Notebook

9/12 - Gained access to HPC and spent time understanding SLURM submissions for job scheduling.

11/12 - Attempted training with the nnUNet framework using Task\_002 example.

23/01 - Converted KiTS19 into 3D nifti format. Considering the differences between left/right VTTs. Discussed the possibility of training separate models or combining them. Also considered using vein segmentation points.

24/01 - Gained access to GPU nodes, which made training much faster.

26/01 - Attended a meeting with NAXIVA Clinical trial team. They suggested that I run an ablation study with nnUNet as the baseline, and split the data into a training and test set. I should also look into auto3dseg. The team discussed how tumors in veins and blood clots around them can be difficult to differentiate and suggested forecasting the response using radiomic features.

27/01 – Explored editing .nii.gz files and resolved issues with pytorch version while training nnUNet models. Set the modality to MRI rather than CT in the nnUNet dataset.json file to prevent the quantitative approach from being used.

28/01 - Studied data augmentation literature and found that the data augmentations used by nnUNet are suitable for NAXIVA images. Investigated uncertainty quantification.

29/01 – Compared GPU training speeds to the nnUNet team baselines on A100s. The runtimes were identical – everything seems to finally be running correctly. Pre-training models were frozen and fine-tuned. Developed Dice score evaluation code.

02/02 - Conducted dicom and png to nifty conversion for CHAOS.

04/02 - Tested nnUNet cascade, but the NAXIVA images were too small to use. Decided to use full resolution nnUNet instead.

10/02 - Conducted AMOS for only larger organs, but it did not significantly improve the results. Explored gradcam for visualizing where the network is looking to extract features / create segmentations.

11/02 - Learned how to use pyradiomics to investigate radiomic features.

13/02 - Developed a complete pyradiomics analysis pipeline.

17/02 - Developed a full analysis pipeline for both Dice evaluation and radiomics. Easier to find larger VTTs with less localization burden and higher contrast.

19/02 - Conducted a radiomics comparison between ground truth and segmentations and found that the nnUNet was identifying more spherical VTTs with lower volume, indicating under-segmentation. Radiomic features, size, and image quality were identified as relevant factors for VTT identification.

20/02: Attended meeting with Ines, Hania, and Mireia to discuss VTT segmentation. Discussed patient splitting and implications of using cubic kernels for making spherical segmentations. We also discussed using automatic segmentation to predict response rather than just making segmentations. Might be suitable even if the segmentations are not of very high quality.

21/02: Discussed definition of responder for Ines's analysis.

22/02: Attended update meeting with NAXIVA clinical team. They seemed to be happy with the direction, and they emphasised looking at automatic segmentation response compared to manual segmentations. Figured out how to save model checkpoints and decided to save a checkpoint every 10 epochs. Will test to see how different models compare epoch by epoch. Can't save every checkpoint due to large file size.

24/02: Started quantifying edge contrast. Read into clustering by radiomic features.

26/02: Pure AMOS inference on NAXIVA data is lacklustre. Struggling with IVC segmentations due to distortions in anatomy and out-of-dataset images. Also seeing a lot of left/right kidney confusion. Will fix with post-processing.

28/02: Combined CHAOS and AMOS datasets for kidney, liver, and spleen.

29/02: Worked on visualizing softmax and other utilities.

01/03: Worked on postprocessing scripts and housekeeping of scripts.

02/03: Met with Hania (and Ines) to show her how to use the radiomics scripts and discuss the congruency in our projects and suggestions for analysis directions.

03/03: Worked on cropping to field of view and pre-cropping algorithms based on kidney, liver, and spleen segmentations.

06/03: Tried using 4dnifti/multiple channels to input CHAMOS segmentations as an extra layer into nnUNet, but this does not seem to be possible.

09/03: Presented to lab. Examined volume/length/Mayo correspondence between segmentation and ground truth. Decided on models (pre-cropped/2-stage/baseline nnUNet) for ablation study but still need to determine exact configurations (ensembling, post-processing, etc.).

13/03: Conducted image quality physics work on contrast, SNR, resolution, Wiener spectrum, etc. Worked to find literature for this type of analysis.

14/03: Created noise segmentations in the image and checked that they are all Gaussian background noise. Worked on Wiener spectrum code.

17/03: Examined post-processing algorithms, e.g. region expansion by watershed.

21/03: Compared coronal and axial scans. Found that the variability on the same VTT/date is too large to use for the same model without degrading performance.

22/03: Tested options for region expansion. Will explore a neural network-based region expansion, probably using nnUNet, in the future since other options did not perform well. Will also look into distance-based loss metrics for this.

28/03 - Conducted ensembling and post-processing tuning work. Organized data on the HPC. Filtered the largest connected component for CHAOS+AMOS segmentations to enhance pre-cropping.

29/03 - Found a way to add data into nnUNet input as an extra channel, which I will use for the 2-stage nnUNet.

01/04 - Ran KiTS (CT) inference docker on NAXIVA images and discovered very poor kidney segmentations due to the large difference between CT and MRI. Using KiTS as pretraining will not be viable.

02/04- Worked on and trained the two-stage nnUNet model (initial segmentation, synthetic data to refine) for the next few days.

06/04 - Converted the convoluted conversion process into one .py script and converted new scans (2 coronal + 9 axial for reference), mainly for Hania.

07/04 - Fixed bugs in the ensembling and physics codes and the one slice offset in the new conversion code.

08/04 - Attempted to limit overfitting by playing around with batch sizes and learning rates, but it only trained slower and ended up with slightly worse performance. Best validation selection outperformed normal early stopping, probably because it takes models from throughout the training process.

11/04 - Analyzed the correlations of radiomics / image quality metrics to segmentation quality.

13/04 - Qualitative analysis of the segmentation.

14/04 - Met with Hania to discuss radiomics, qualitative elements of the segmentation. There is no significant difference between cancer in kidney/VTT other than location. Could use kidney segmentations to exclude these cases. We discussed different kinds of geometrical response, which would be interesting to see if they are apparent in automatic segmentations.

18/04 - Evaluated segmentation quality using AUC ROC & AUPRC.

21/04-26/04 - Started the initial draft of the project report and noted the contents.

28/04 - Met with Ines and Hania. Submitted an abstract for MIUA. Discussed write-up structure and relevant information for imaging data. Analyzed the segmentation utility for a radiologist using the AUGMENT framework.

29/04-03/05 - Wrote the bulk of the report from the notes. Created many plots and formatted the report.

04-06/05 - Wrote and edited the MIUA abstract and sent it to coauthors for review.

07/05 – Final editing of the report, adding a few plots.

11/05 - Presented to the lab and discussed using spherical kernels, the effects of resolution, the applicability to semi-automatic segmentations, and ovarian cancer. Also compiled, cleaned and uploaded all of the code I wrote during the project.