Image Segmentation

Exercise 2.3
Introduction to Deep Learning in Computer Vision

October 2025

This exercise forms the first part of what you will present in your project for Segmentation. The second part will be introduced next week.

In this exercise, you will be working with two different datasets for image segmentation:

- The skin lesion segmentation dataset PH2

- The retinal blood vessel segmentation dataset DRIVE

Your task is to design a generic segmentation architecture that you apply to both tasks and validate your algorithm using multiple metrics for segmentation performance, as well as a toy ablation study illustrating how different choices of a specified parameter impact the performance of your network at your given task.

Your tasks are as follows:

1. Locate the two datasets at /dtu/datasets1/02516 and describe them in terms of size, pre-assigned splits, etc. You will have to make choices on how to create training/validation/test splits based on how the datasets are structured from scratch, and you will have to write data loaders for both datasets. Use the data loader given in the segmentation exercise as a starting point to do this.

   NB! The retinal blood vessel dataset contains two different types of masks: The retinal vessel masks, and a field-of-view-mask that separates out the non-black part of the image. You should identify the retinal vessel mask and make sure to use this one for segmentation.

2. Implement a simple encoder-decoder segmentation CNN and implement the following metrics for validating segmentation performance: Dice overlap, Intersection over Union, Accuracy, Sensitivity, and Specificity. You should use this set of metrics for each of your tasks below; please comment on the strengths and weaknesses of the different performance measures, both generally and for the specific cases.

   Report your results with respect to the metrics on both datasets.

3. Implement a U-net architecture to segment the images, and train and test it using the same train/validation/test splits as above. Please report the five performance measures on all three splits. *Tip: You may have to adapt the U-net architecture to a new size image. Choose your resampled image size wisely to avoid losing too much resolution, but still have a well-functioning U-net. Sketch your U-net and the size of the feature image of each layer, and verify that the shape at each layer in your implementation matches.*

   NB! Beware of your memory consumption if you choose to use the retinal images at full resolution – please downscale your batch sizes to make sure you don't absorb all

the memory on the GPUs. If your jobs are memory intensive, please make sure only one group member is running one at a time.

4. Perform an ablation study where you try to improve the segmentation by testing different loss functions (cross entropy, focal loss, cross entropy with positive weights to make up for class imbalance). Describe what you did, and report results for each choice of loss function.

   NB! In normal scenarios, we perform similar ablation studies for all model choices, including augmentation, optimization tools and regularization techniques. While you will likely play around with these, for the sake of GPU load, please don't carry out full ablations.

5. Did you use ChatGPT or similar tools? If yes, please briefly describe how you used it and how they were useful.

Your process, performance evaluation and results should be documented and discussed in the project report.