Assignment Report on Polyps Image Segmentation

Name: Phan Trần Việt Bách

Student ID: 20225435

1. Abstract

The assignment requires training of a U-Net model to perform semantic segmentation task on the BKAI-IGH NeoPolyp dataset, identifying polyps from images as well as classifying them as either neoplastic or non-neoplastic.

The dataset consists of 1000 images and their corresponding segmentation masks for training, and 200 unsegmented images for testing. Several image transformation techniques are applied to allow better performance during training and inference process. Images are resized to a smaller size to reduce computational complexity of the neural network model and to ensure all images are viewed in a similar perspective. Then, the images are randomly flipped - horizontally and/or vertically, normalized to have color values reduced to a smaller range, and have their color channels’ values randomly shifted to better generalize how a polyp is visualized on an image.

Several model architectures are also considered, including a self-constructed model and a model provided by existing libraries with diverse characteristics. However, all models have the same structure of a U-Net model: an encoder block to analyze features of an image, a decoder block to construct the segmentation mask from these features, and the ability to deal with information loss from reducing feature maps’ size. Although these models differ in architecture, the differences in performance seem to come from data transformation techniques, allowing the models to generalize well between images.

1. Data Transformation

Based on investigation of the dataset, some characteristics are considered to perform preprocessing. First, all images have sizes of around 1280x960, and while the width of 1280 is consistent across them, the height is not. The approximate size is also quite large and will require huge memory usage when converting to data structures to pass to the model. So, all images, both original and mask, are resized to a smaller size; here I selected the size of 256x256 after multiple tries as it can balance well enough between memory consumption and performance.