* Training Data Construction for Segmentation
* Output = [3 x scan height x scan width]
* Unet - f**ully convolutional networks** / ViT / Both - pixel wise classification
  + Removing Transformer to simplify; compare results in report
* Vary input size problem - spatial pyramid pooling + Unet (Allow Fully Connected Layer)
* Segment Anything + classification - (1) produce masking → classify each mask. (2) Transfer learning by using Segmentat Anything network
  + <https://github.com/facebookresearch/segment-anything>
* Add pixel-wise cross-entropy loss function along with Dice Similarity Coefficient
  + Do training with and without cross-entropy training for report comparison

Train\_input.npy: (num\_samples, num\_channel = 3, height, width) num\_channel is the 2.5D construction such that 0 layer is the previous image slice and 1 is the current image slice and 2 is the next image slice.

Train\_labels.npy: (num\_samples, num\_class = 3, height, width) num\_class is the number of organs that we are interested in such that 0 is the segmentation mask for stomach, 1 is small\_bowel, 2 is large\_bowel.

Fully convolutional (Mingyu, Peter)

-Operational, masks produced are not correct

ViT (Ruide, Gabriel)

-Non-operational: Need to implement segmentation

Segment Anything (Ryan)

-Non-operational: Segments image, no classification