Homework 2

Notes you want the TAs to consider when grading.

Problem 2

2.1

The instance segmentation problem is about training a model to segment different objects. It classifies pixels to a certain class according to the color of the object that the pixel belongs to. The input of the model is RGB-D images, while the ground truth masks are target labels. The output is the accurate segmentation of the objects. This is a supervised learning problem given the training has ground truth masks as target.

2.4

1.
1x1 kernel size is used for dimension reduction to reduce the number of channels. It only need to captures the interaction of input channels in one pixel of the feature map. But the 3x3 kernels used in most of the convolutional layers need to extract useful information about each pixel by collecting its adjacent pixels.

2.
The 6 output channels is designed due to the classes of different objects. In this model, each pixel can be predicted and classified into 6 classes corresponding to the object colors: 0 background black 1 sugar box red 2 tomato soup can green 3 tuna fish can blue 4 banana yellow 5 bowl purple For 1 channel ground truth mask, it collects information about a certain class rather than multiple classes.

2.5
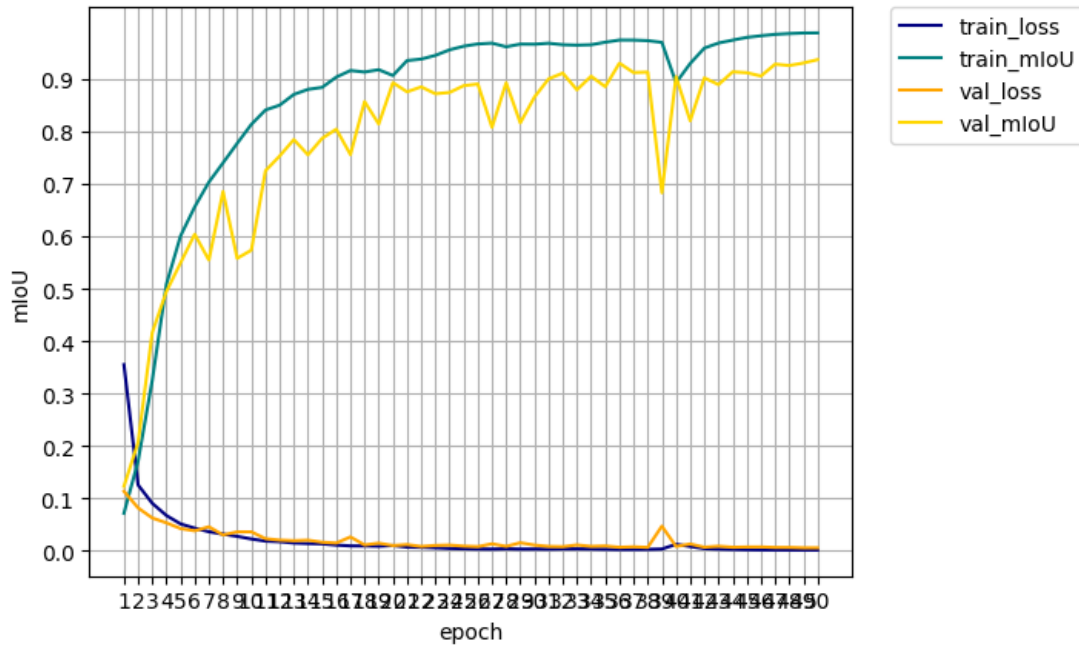
This learning curve represent using Optimizer= Adam, Epoch number=50.
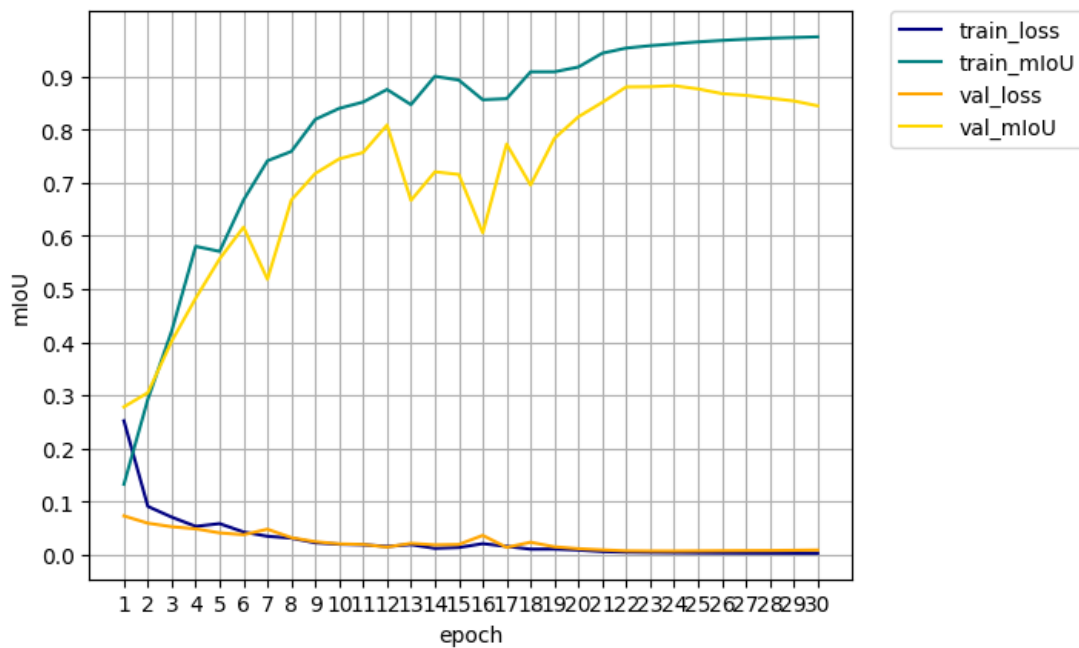
This learning curve represent using Optimizer= SGD, Epoch number=20

The Adam optimizer works better than SGD, and the larger epoch number is, the higher val_IoU is. The batch size influences the stability of the model.

Problem 3

3.2

1.
In this part, I try different max_ iteration and threshold from low to high. At first, The improvement is drastic, but beyond a certain limitation, increasing max_iterations makes negligible improvement to the align function. And it also wastes the computational resources. Qualitatively, increasing max_iterations also displayed larger variances between the predicted and gt masks in Meshlab. My final result used max_iteration = 10000, and the threshold = 1e-8. For initial transformation I used the trimesh Procrustes-method and this greatly improved the Average Closest Point Distance metric and gave closer matches in meshlab.

2.
Note: For the second RGB image, the banana object is missing. So the ACDP could not be calculated, and there is no metric for this class. In Meshlab, the red and green point cloud is not shown.
Another note: when the objects that occlude with each other, the orientation may be different from the ground truth mask. For example, the bowl probably reverse in the predicted point cloud.

Leave comments on your code for the TA to consider. Not required.