Question 3

1.

## todo:

# conv\_block2 = nn.Sequential()

# conv\_block2 is missing. However, you know the block2 is similar to block1 with deeper CNN structure.

# Can you find it out?

self.conv\_block2 = nn.Sequential(

nn.Conv2d(c1,c2,3,padding = 1),

nn.ReLU(inplace = True),

nn.Conv2d(c2,c2,3,padding=1),

nn.ReLU(inplace=True),

nn.MaxPool2d(2,stride = 2,ceil\_mode=True)

)

self.conv\_block3 = nn.Sequential(

nn.Conv2d(c2, c3, 3, padding=1),

nn.ReLU(inplace=True),

nn.Conv2d(c3, c3, 3, padding=1),

nn.ReLU(inplace=True),

nn.Conv2d(c3, c3, 3, padding=1),

nn.ReLU(inplace=True),

nn.MaxPool2d(2, stride=2, ceil\_mode=True),

)

# todo:

# conv\_block4 is missing. However, you know the block4 is similar to block3 with deeper CNN structure.

# Can you find it out?

self.conv\_block4 = nn.Sequential(

nn.Conv2d(c3, c4, 3, padding=1),

nn.ReLU(inplace=True),

nn.Conv2d(c4, c4, 3, padding=1),

nn.ReLU(inplace=True),

nn.Conv2d(c4, c4, 3, padding=1),

nn.ReLU(inplace=True),

nn.MaxPool2d(2, stride=2, ceil\_mode=True),

)

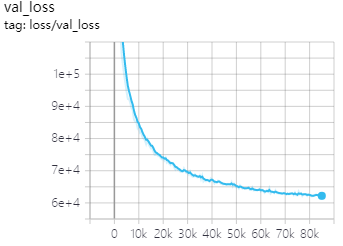
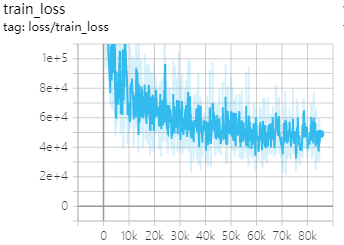
The model starts with a set of 2d convolution layers and followed by relu and pooling layers, then the output features is upsampled by a set of upsampling convolution layers to get the segmentation

2.

We do use weights from a pretrained vgg16 model

3.

The training and validation curve is shown below



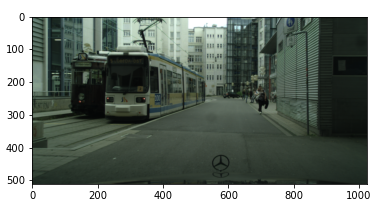
4.

The metrics used by the original paper is mean pixel intersection over union, with the mean taken over all classes, including background.

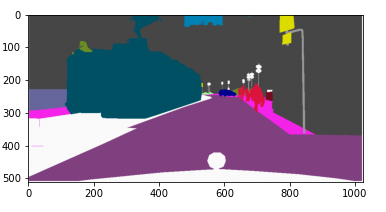
Class 1,2,3,8,10 and 13 works well with accuracy 0.96, 0.60, 0.85, 0.86, 0.89 and 0.85, class 4,5,6,7,9,11,12,14,15,16,17,18 doesn't work very well

with accuracy 0.45, 0.36, 0.25, 0.30, 0.40, 0.51, 0.54, 0.23, 0.53, 0.47, 0.34, 0.22, 0.55

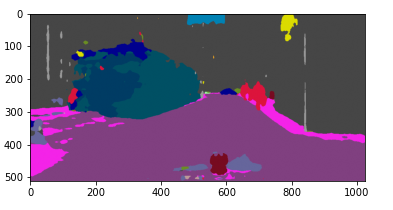
5. The pictures are shown below



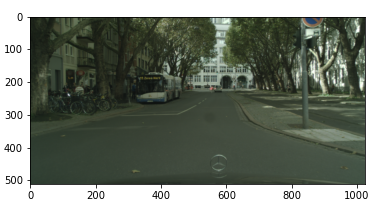
Example 1 original image



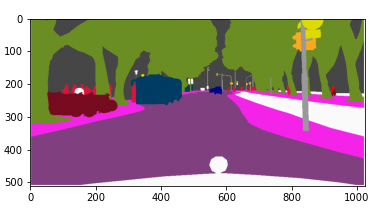
Example 1 label



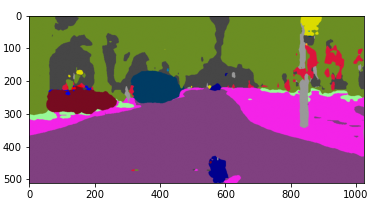
Example 1 prediction



Example 2 original image

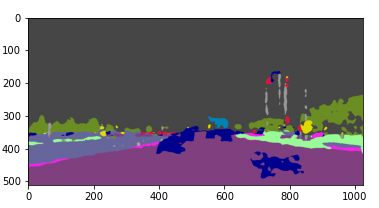


Example 2 label



Example 2 prediction

1. The photo I take and its prediction are shown below. Though some part of the image is misunderstood, the prediction looks reasonable



7.

We can use Atrous Spatial Pyramid Pooling to get mutipul scale information from the original image in order to improve the performance.

We can try different optimizer from SGD, adagrad, rmsprop and adam.

Code:

Upload in gradescope