Deep Learning Lab WS2018
Exercise 3

Ralph Lesch

December 4, 2018

The topic of this exercise is semantic segmentation. We implement several decoder modules for Fully Convolutional Networks (FCNs) in 4 configurations:

1. a simple single stage decoder module, upsampling the encoder results via transposed convolution 16x to its original spatial resolution.
2. multi-stage decoder with 2x upsamling and a skip connection to refine the results.
3. 2x refinement with 2x upsamling.
4. 3x ref., also with 2x upsampling each.

Results

Below you can see the maximum Intersection Over Union (IoU) for each configuration (Table 1) and their plots (Figure 1, 2). The difference between a single stage decoder (config 1) and the multi-stage decoders (config 2-4) is clearly visible. With three refinements / skip connections (config 4) the maximum IoU is 10x higher than without (config 1).

config	max(IoU)	epoch
1	0.03592	32 000
2	0.07798	36 000
3	0.22098	31 000
4	0.38193	40 000

Table 1: Intersection ver Union (IoU)

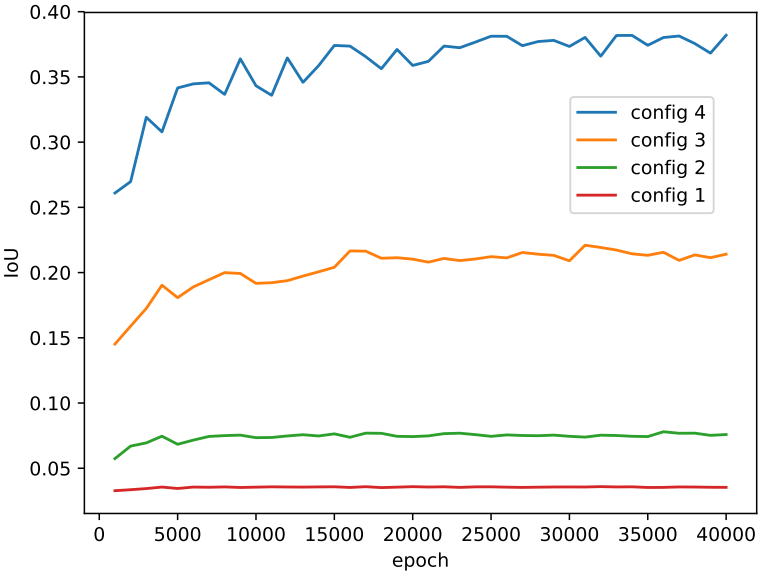

Figure 1: IoU vs epoch plot of all configurations

After a high increase of IoU at the first 5-10 thousand epochs (longer with more refinements), the curves saturate, with only a small increase in IoU. With overall higher and longer increasing IoU (until saturation), config 4, with the most refinements is clearly the best decoder here.

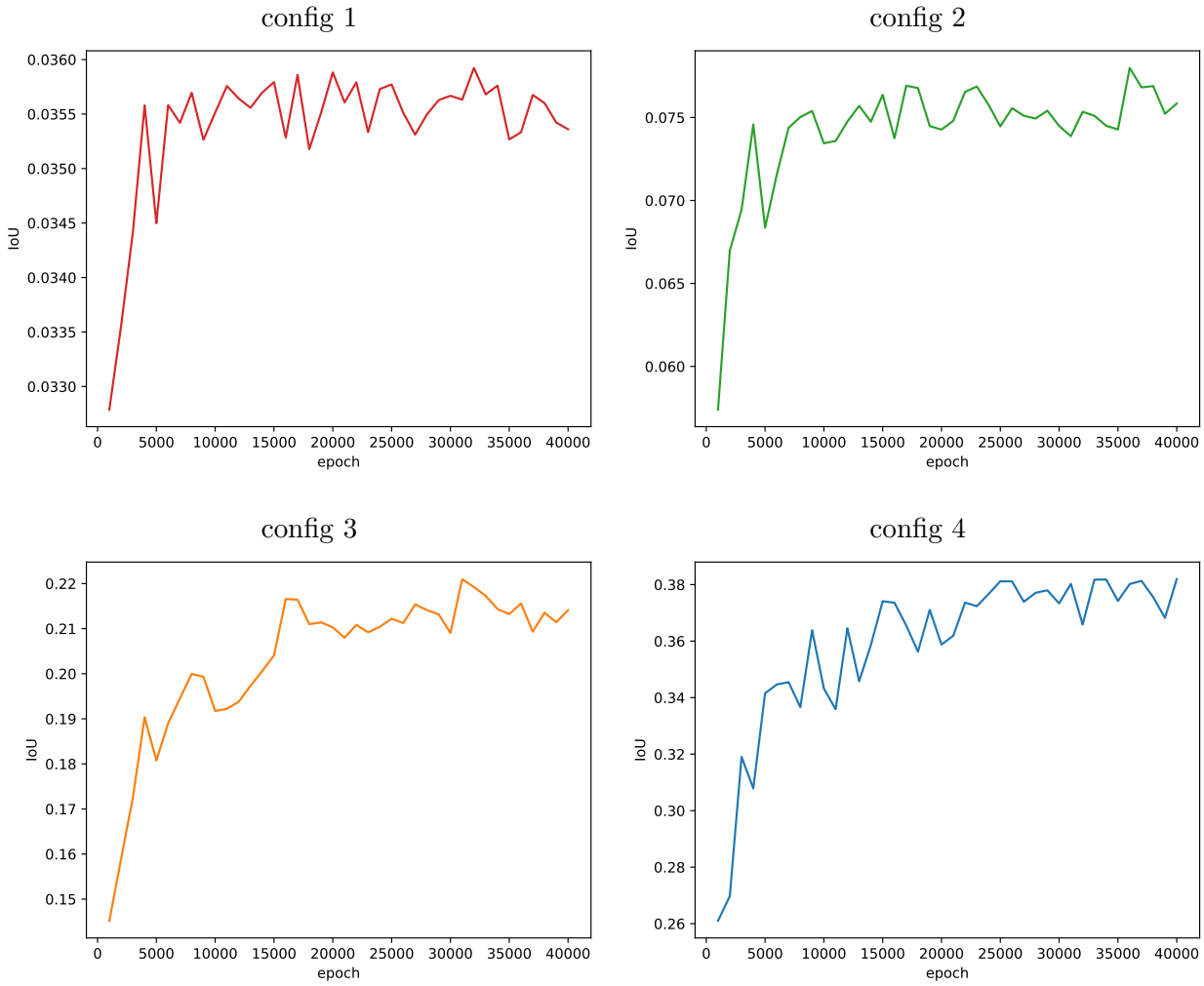

Figure 2: plot for each configuration