Deep Learning Lab Report
Assignment 3

Ravin Kohli

December 2, 2018

1 Objective

Training Encoder-Decoder network with 4 different configurations of the decoder network
We were tasked with training an encoder-decoder network for semantic segmentation. The main aim of the task is to analyse the impact of the number of upsamples performed on the performance of the network. We implemented four configurations

1. Single Upsample -16x (hereby referred to as Configuration 1)
2. Two Upsamples -2x → 8x (hereby referred to as Configuration 2)
3. Three Upsamples -2x → 2x → 4x (hereby referred to as Configuration 3)
4. Four Upsamples -2x →2x→2x → 2x (hereby referred to as Configuration4)

2 Network Architecture

Architecture of the Decoder Network

1. Configuration 1 - There is no refinement block and we directly upsample the feature map from the encoder to the size of the image.

Table 1: Configuration 1

Layer Number	Output Feature maps	Upsampling Rate	Kernel size
Upsample 1	120	16	3
Conv	No. of Classes	1	1

2. Configuration 2 - There is one refinement block with its corresponding skip connection.

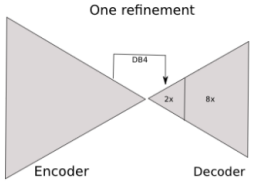

Table 2: Configuration 2

Layer Number	**Output Feature maps**	**Upsampling Rate**	**Kernel size**
Upsample 1	256	2	3
Conv 1	256	1	3
Upsample 2	120	8	3
Conv 2	No. of Classes	1	1

3. Configuration 3 - There are two refinement blocks with their corresponding skip connections.

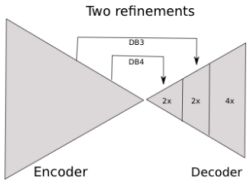

Table 3: Configuration 3

Layer Number	**Output Feature maps**	**Upsampling Rate**	**Kernel size**
Upsample 1	256	2	3
Conv 1	256	1	3
Upsample 2	160	2	3
Conv 2	160	1	1
Upsample 3	120	4	3
Conv 3	No. of Classes	1	1

4. Configuration 4 - There are three refinement blocks with their corresponding skip connections.

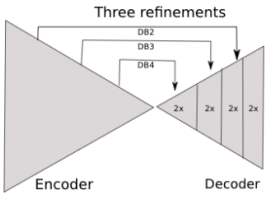

Table 4: Configuration 4

Layer Number	Output Feature maps	Upsampling Rate	Kernel size
Upsample 1	256	2	3
Conv 1	256	1	3
Upsample 2	160	2	3
Conv 2	160	1	1
Upsample 3	96	2	3
Conv 3	96	1	3
Upsample 4	120	4	3
Conv 4	No. of Classes	1	1

3 Results

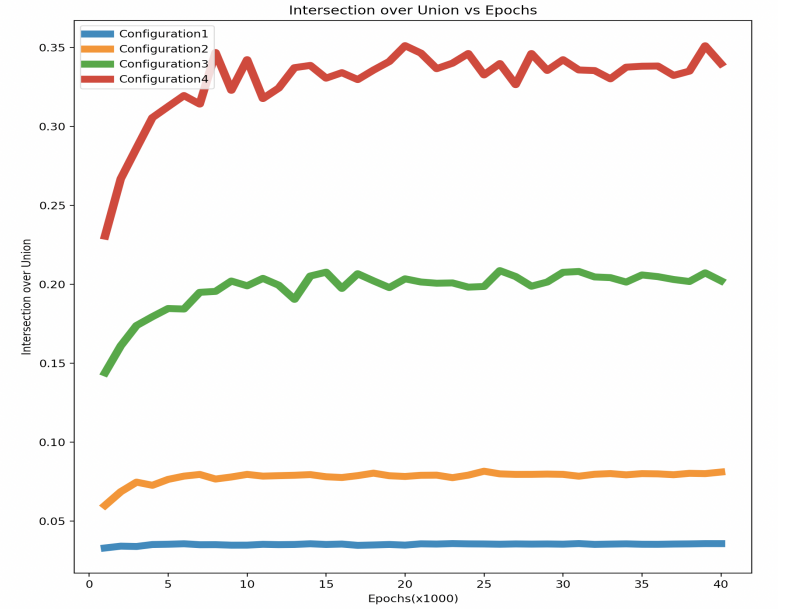

Figure 1: IoU vs Epochs

Table 5: Results

Configuration	IoU
Configuration 1	0.035
Configuration 2	0.081
Configuration 3	0.202
Configuration 4	0.339

4 Conclusion

We have tested four configurations according to the number of refinement blocks present in them. The best performing configuration is with using three refinement blocks and upsampling it 2x in every step. This can be understood as increasing the number of upsampling increases the number of parameters to be learnt. Also, this configuration has 3 skip connections. That means it is using the output from the convolutional layers in the encoder part of the network and feeding it into the output of the upsampling which in turn helps the network to estimate the image well.