

AI6126: Homework 2

Deadline: 21 September 2020 11:59PM

[Total marks: 20, will be normalized to 5 marks]

[2 marks] Question 1: Explain why is scaling and shifting often applied after batch normalization? (*Hint: Check the original batch normalization paper*)

[3 marks] Question 2: Describe the main changes, e.g., architecture change or new losses, made in Faster R-CNN in comparison to Fast R-CNN.

[3 marks] Question 3: Describe the differences between the operation of RoIPool and RoIAlign. Explain why RoIAlign is preferred over RoIPool (*Hint: Check the paper K. He, Mask R-CNN, ICCV 2017*)

[4 marks] Question 4: 1) Explain why the encoder-decoder architecture is widely used in semantic segmentation tasks. 2) Does the plain encoder-decoder architecture have potential drawbacks? If so, how can we fix them?

[4 marks] Question 5: When we apply consecutive 1-dilated, 2-dilated, 4-dilated and 8-dilated 3x3 convolution, what is the final receptive field?

(A question to challenge you!)

[4 marks] Question 6: 1) Even though dilated convolution improves upon standard convolution, what are the potential hard cases for dilated convolution? 2) How will you further improve upon dilated convolution? (*Hint: Is there a more flexible form of convolution?*)