EE5934/6934 Q&A Assignment

Note: Please submit this assignment with name "Yourname_StudentID.pdf"via luminus before April 3rd 13:00. For drawing, you can choose hand drawing or PowerPoint. You will be deducted 5 points for every 12 hours of delay.

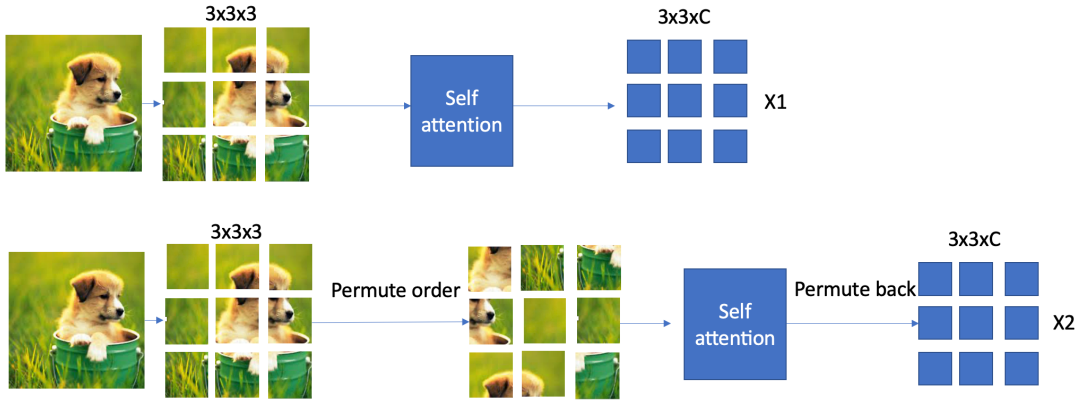

Figure 1

1. As shown in figure 1, assuming that two self-attention layer share the same parameter, is the feature X_1 equal to X_2 ? Please illustrate how to make your point. (15 points)

2. What is RNN_____(full spells)? If computational graph in RNN is "many to many", do you need to use different parameter sets for different input vectors? (15 points)

3. Please make comparisons among R-CNN, Fast R-CNN, Faster R-CNN and explain what do "fast" and "faster" mean. (15 points)

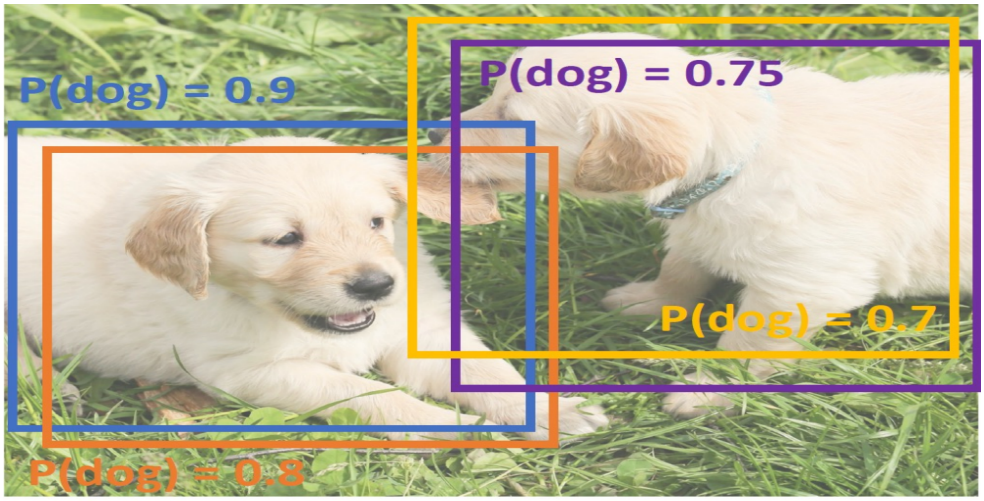

Figure 2

4. As shown in Figure 2, object detectors often output many overlapping detections. How to

solve this problem? (15 points)

5. Please illustrate inputs and outputs for the task of semantic segmentation with specific dimension i.e. $H \times W \times C$. Upsampling is the key part in fully convolution network for semantic segmentation. How to do the upsampling? (15 poins)

6. Please compare the 3D convolution and R(2+1)D convolution for spatiotemporal modeling and list key differences. You can choose to describe in words or draw diagrams. (15 points)

7. Please draw the diagram of the basic inception block with specific feature sizes. (10 points)