AI6126: Homework 2

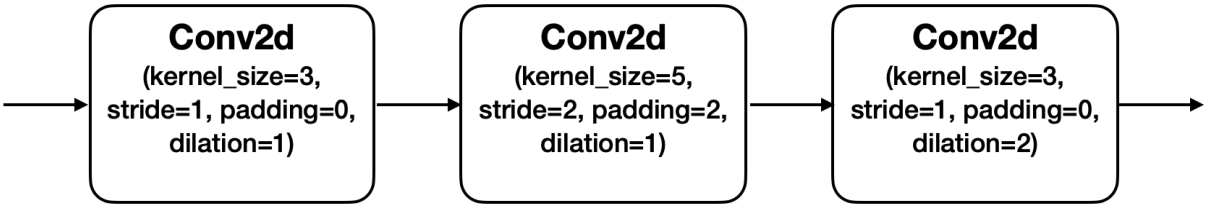
Deadline: 11 April 2022 11:59PM

Question 1: The following questions are related to image segmentation.

i) What is the difference between semantic segmentation and instance segmentation?

(1 mark)

ii) Given the following network, calculate the receptive field.



(2 marks)

iii) Spatial context is particularly important for segmentation tasks. List at least four techniques that improve semantic segmentation in terms of spatial context.

(2 marks)

iv) Given a transposed convolution kernel as follow, whose stride=1, padding=0, dilation=1,

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{bmatrix}$$

and input,

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

What is the output matrix after the transposed convolution?

(Hint: the output should be a 4x4 matrix.)

(2 marks)

Question 2: What is the purpose of the reparameterization trick in VAE?

(3 marks)

Question 3: What is the difference between GANs and Conditional GANs?

(2 marks)

Question 4: The Transformer architecture proposed by Vaswani et al. in "Attention is All You Need" NIPS 2017 is shown in Figure Q1.

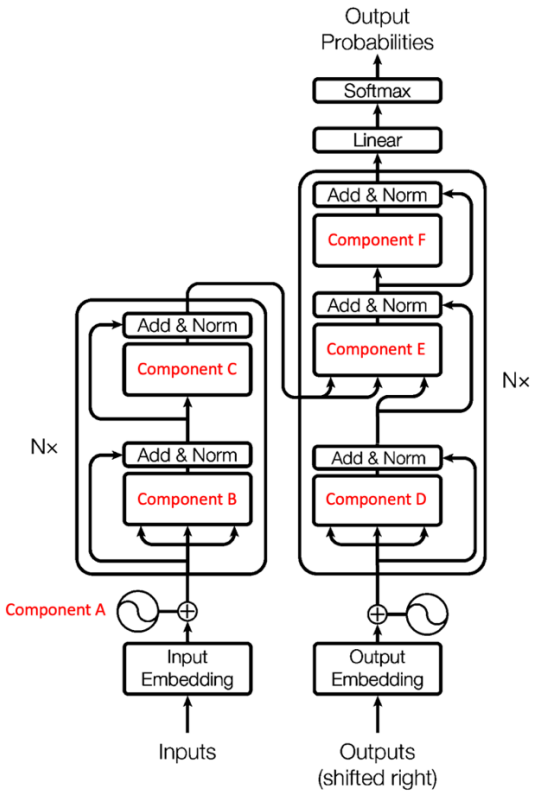


Figure Q1

Answer the following questions:

a) Write the name of components A, B, C, D, E, and F.

(3 marks)

b) Explain the role of Component A and suggest a way to generate the output encoding of this component.

(1 mark)

c) What are the inputs of Component E?

(1 mark)

d) Explain why masked self-attention is needed in the decoder of the Transformer. Suggest a way to achieve masked self-attention in practice.

(3 marks)

[END]