Deep Learning

Assignment- Week 11

TYPE OF QUESTION: MCQ/MSQ

Number of questions: 10   **Total mark: 10 X 1 = 10**

QUESTION 1:

Which of following can be a target output of semantic segmentation problem with 4 class?

a.

0	1	0
0	1	0
1	0	0

I

1	1	0
0	0	0
0	0	0

II

0	0	1
1	0	0
0	0	0

III

0	0	0
1	0	0
0	0	0

IV

b.

0	1	0
0	1	0
1	0	0

I

1	0	0
0	1	0
0	0	0

II

0	0	1
1	0	0
0	0	0

III

0	1	0
0	0	0
0	1	1

IV

c.

0	1	0
0	1	0
1	0	0

I

1	0	0
0	0	1
0	0	0

II

0	0	1
1	0	0
0	0	0

III

0	0	0
0	0	0
0	1	1

IV

d.

0	1	0
0	1	0
1	1	0

I

1	0	0
1	0	0
0	1	0

II

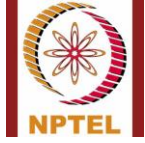| 0	0	1
1	0	0
0	0	1

III

0	0	0
0	0	0
0	1	1

IV

Correct Answer: c

Detailed Solution:

Target output should be one hot encoded vector at every pixel location. It should one if the pixel belongs to that particular class otherwise 0.

QUESTION 2:

Suppose you have a $1D$ signal $x = [1,2,3,4,5]$ and a filter $f = [1,2,3,4]$, and you perform stride 2 transpose convolution on the signal x by the filter f to get the signal y. What will be the signal y if we don't perform cropping?

        a. $y = [1,2,5,8,9,14,13,20,19,26,3,4]$
        b. $y = [1,2,3,4,5,4,3,2,1]$
        c. $y = [1,2,5,8,9,14,13,20,17,26,15,20]$
        d. $y = [0,0,5,8,9,14,13,20,19,26,0,0]$

Correct Answer: c

Detailed Solution:

	1	2	3	4	5
1	1				
2	2				
1*3+2*1=5	3	1			
4*1+2*2=8	4	2			
3*2+1*3=9		3	1		
4*2+2*3=14		4	2		
3*3+1*4=13			3	1	
4*3+2*4=20			4	2	
3*4+1*5=17				3	1
4*4+2*5=26				4	2
15					3
20					4

QUESTION 3:

What are the different challenges one face while creating a facial recognition system?

       a. Different illumination condition
       b. Different pose and orientation of face images
       c. Limited dataset for training
       d. All of the above

Correct Answer: d

Detailed Solution:

Please refer to the lecture of week 11.

_____

QUESTION 4:

Fully Connected Convolutional network or FCN became one of the major successful network architectures. Can you identify what are the advantages of FCN which makes it a successful architecture for semantic segmentation?

       a. Larger Receptive Field
       b. Mixing of global feature
       c. Lesser computation required
       d. All of the above

Correct Answer: d

Detailed Solution:

Please refer to the lecture of week 11. It has a larger receptive field by using strided Convolution layer, also it mixes global feature and number of computations are reduced as we are down sampling the image resolution.

_____

QUESTION 5:

In a Deep CNN architecture, the feature map before applying a max pool layer with (2x2) kernel is given bellow.

12	6	15	9
19	2	7	18
14	2	17	6
3	5	19	2

After few successive convolution layers, the feature map is again up-sampled using Max Un-pooling. If the following feature map is present before Max-Unpooling layer, what will be the output of the Max-Unpooling layer?

5	6
8	13

a.

0	0	0	0
5	0	0	6
8	0	0	0
0	0	13	0

b.

5	5	6	6
5	5	6	6
8	8	13	13
8	8	13	13

c.

5	0	6	0
0	0	0	0
8	0	13	0
0	0	0	0

d. None of the above

Correct Answer: a

Detailed Solution:

Please refer to lectures of week 11.

QUESTION 6:

What could be thought as disadvantage of Fully Convolutional neural network for semantic segmentation addressed by other researchers?

    a. It has a fixed receptive field, so the object with lesser size than the receptive filed will be missed by the network.
    b. Down Sampling the image dimension over the depth makes the feature map sparse.
    c. It requires lot of computation.
    d. None of the above

Correct Answer: a

Detailed Solution:

The fixed receptive filed is the disadvantage of Fully convolution network for semantic segmentation addressed by "Learning Deconvolution Network for Semantic Segmentation" paper.

QUESTION 7:

What will be the dice coefficient of following two one hot encoded vector? (|A|=no of 1 bit)

A	1	0	1	0	0	0	1	1	1	0	0	1	0	1
B	1	0	0	0	0	1	1	1	0	0	0	1	0	0

    a. 0.83
    b. 0.41
    c. 0.67
    d. 0.90

Correct Answer: c

Detailed Solution:

No of 1 bit in A=7
No of 1 bit in B =5

Overlapping 1 bit =4

Dice Coefficient $= \frac{2*|A \cap B|}{|A|+|B|} = \frac{2*4}{5+7} = 0.67$

QUESTION 8:

What will be the value of dice coefficient between A and B?

0.01	0.03	0.02	0.02
0.05	0.12	0.09	0.07
0.89	0.85	0.88	0.91
0.99	0.97	0.95	0.97

A

0	0	0	0
0	0	0	0
1	1	1	1
1	1	1	1

B

(Consider, |A|= sum of all elements.)

    a. 0.23
    b. 0.77
    c. 0.11
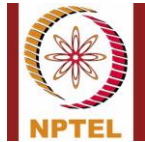   d. 0.89

Correct Answer: a

Detailed Solution:

|A|=7.82

|B|=8

A∩B=

0	0	0	0
0	0	0	0
0.89	0.85	0.88	0.91
0.99	0.97	0.95	0.97

|A∩B|=7.42

Dice Coefficient = 2*|A∩B|/(|A|+|B|)=2*7.42/ (8+7.82) =0.23

QUESTION 9:

In FaceNet, why the L2 normalization layer is used?

    a. To constrain the embedding function in a d-dimensional hyper-sphere.
    b. For regularization of weight vector, i.e. L2 regularization.
    c. For getting a sparse embedding function.
    d. None of the above.

Correct Answer: a

Detailed Solution:

Using the L2 normalization layer we impose the constrain that $\|f(x)\|_2^2 = 1$. This will constrain the embedding function to live on the d-dimensional hyper-sphere.

QUESTION 10:

What is the use of Skip Connection in image denoising networks?

    a. Helping de-convolution layer to recover an improved clean version of image.
    b. Back propagating the gradient to bottom layers, which makes the training easy.
    c. To create the direct path between convolution layer and the corresponding mirror de-convolution layer.
    d. All of the above.

Correct Answer: d

Detailed Solution:

Please refer to lecture of week 11.

***********END*******