CS 7150, in-class quiz 1   Your Name:

Note 1: There may be one or multiple correct choices for the first 4 questions.
Note 2: Briefly answer the questions 5 to 7.

1. What do you think of the pace of the course so far? (8 points)
   (a) Too slow.   (b) Just right.   (c) Too fast.

2. If you input image's shape is $16 \times 32 \times 32$ (16 is the number of channels and 32×32 is the spatial dimension), how many parameters are there in a single 3×3 convolution filter, including bias? (10 points)

   (a) 9
   (b) 10
   (c) 145
   (d) 10,240

3. Which of the following are true about Batch Normalization (BN)? (10 points)

   (a) BN is another way of doing convolution
   (b) BN makes training converge faster
   (c) BN can only be used in the training phase
   (d) BN is a non-linear transformation that centers the output around the origin

4. Which of the following are true about a Convolution layer? (10 points)

   (a) The number of weights depends on the number of channels (depth) of the input volume
   (b) The number of biases is equal to the number of filters
   (c) The total number of parameters is dependent on the stride
   (d) The total number of parameters is dependent on the padding

5. Bob made the following statement: a fully-connected layer is essentially equivalent to a convolution layer. Is it correct? Why or why not? (10 points)

6. Name two reasons why downsampling is essential in CNNs. (10 points)

7. Alice is designing a convolutional neural network. But somehow she forgets to add non-linear activation functions in-between two convolution layers. Why may happen to the neural network? (10 points)

8. Consider the convolutional neural network defined by the layers in the left column below. Fill in the shape of the output volume and the number of parameters at each layer. You can write the shapes in the format of $3 \times 128 \times 64$ (3 being the channel dimension, 128 being the height, and 64 being the width). (32 points)

   Notation:

   - CONV-x-y-z-N denotes a convolutional layer with N filters. Kernel height and width are x (a square kernel). Padding is y, and stride is z.
   - FLATTEN flattens its input.
   - FC-N denotes a fully-connected layer with N neurons/output.

   Note: For the number of parameters, it is okay to leave an equation there. But for the output volume dimension, due to the notation definition, $1 \times 2 \times 3$ means differently from 6.

Layer	Output Volume Dimension	Number of Parameters
Input	$3 \times 24 \times 24$	0
CONV-$5 - 0 - 2 - 4$	4x10x10	4x(5x5x3 + 1)=304
Batch Normalization	4x10x10	8
ReLU	4x10x10	0
CONV-$3 - 1 - 2 - 8$	8x5x5	8x(3x3x4+1) =
Batch Normalization	8x5x5	16
ReLU	8x5x5	0
FLATTEN	200	0
FC-10	10	10x(200 + 1) = 2010