1. What do you think applying this filter to a grayscale image will do? 1 point

$$\begin{bmatrix} -1 & -1 & 2 \\ -1 & 2 & 1 \\ 2 & 1 & 1 \end{bmatrix}$$

- (•) Detect 45-degree edges.
- ( ) Detect vertical edges.
- ( ) Detecting image contrast.
- ( ) Detect horizontal edges.

2. Suppose your input is a 300 by 300 color (RGB) image, and you are not using a convolutional network. If the first hidden layer has 100 neurons, each one fully connected to the input, how many parameters does this hidden layer have (including the bias parameters)? 1 point

- ( ) 9,000,100
- ( ) 27,000,001
- (•) 27,000,100
- ( ) 9,000,001

3. Suppose your input is a 256 by 256 grayscale image, and you use a convolutional layer with 128 filters that are each 3×3. How many parameters does this hidden layer have (including the bias parameters)? 1 point

- ( ) 75497600
- (•) 1280
- ( ) 1152
- ( ) 3584

4. You have an input volume that is $127 \times 127 \times 16$, and convolve it with 32 filters of 5×5, using a stride of 2 and no padding. What is the output volume? 1 point

- ( ) $123 \times 123 \times 16$
- ( ) $123 \times 123 \times 32$
- ( ) $62 \times 62 \times 16$
- (•) $62 \times 62 \times 32$

5. You have an input volume that is 31x31x32, and pad it using "pad=1". What is the dimension of the resulting volume (after padding)? 1 point

- ( ) 33x33x33
- ( ) 31x31x34
- (•) 33x33x32
- ( ) 32x32x32

6. You have an input volume that is 63x63x16, and convolve it with 32 filters that are each 7x7, and stride of 1. You want to use a "same" convolution. What is the padding? 1 point

- ( ) 7
- (•) 3
- ( ) 1
- ( ) 2

7. You have an input volume that is 66x66x21, and apply max pooling with a stride of 3 and a filter size of 3. What is the output volume? 1 point

- ( ) $21 \times 21 \times 21$
- (•) $22 \times 22 \times 21$
- ( ) $22 \times 22 \times 7$
- ( ) $66 \times 66 \times 7$

8. Which of the following are hyperparameters of the pooling layers? (Choose all that apply) 1 point

- [ ] Number of filters.
- [ ] Average weights.
- [x] Filter size.
- [x] Whether it is max or average.

9. Which of the following are true about convolutional layers? (Check all that apply) 1 point

- [x] It allows a feature detector to be used in multiple locations throughout the whole input volume.
- [ ] It speeds up the training since we don't need to compute the gradient for convolutional layers.
- [x] Convolutional layers provide sparsity of connections.

10. In lecture we talked about "sparsity of connections" as a benefit of using convolutional layers. What does this mean? 1 point

- ( ) Regularization causes gradient descent to set many of the parameters to zero.
- ( ) Each filter is connected to every channel in the previous layer.
- ( ) Each layer in a convolutional network is connected only to two other layers
- (•) Each activation in the next layer depends on only a small number of activations from the previous layer.