

1. The input to my CNN is 10×10 , and a single 5×5 filter is applied. The output of this operation is 10×10 . How is this possible? (2 points)

We can use padding. 2 layers on each side.

2. Given the following input, what is the result of the operations below?

1	1	1	1	1
1	1	1	1	1
1	1	1	1	1
1	1	1	1	1
1	1	1	1	1

- (a) 3×3 Max Pooling (2 points)

Same as 691

- (b) 5×5 Max Pooling (2 points)

Same as 691

- (c) Application of the following filter: (2 points)

1	1	1
1	1	1
1	1	1

Same as 691

3. How is the step size in a deep network different from the learning rate? (2 points)

Step size tells us when to lower learning rate. LR tells us how far to step in GD.

1. Given an input image of $25 \times 25 \times 3$, that is, a color image of size 25×25 , what size filter is needed to get a first hidden layer of size 19×19 ? Assume no padding, and stride=1. (2 points)

$$25 - 19 = 6 / 2 = 3 = \text{floor}(\frac{1}{2} \times \text{Filter size})$$

$$7 \times 7 \times 3$$

2. Given the following input, what is the result of the operations below?

1	1	1	1	1
1	1	1	1	1
1	1	1	1	1
1	1	1	1	1
1	1	1	1	1

- (a) 3×3 Max Pooling (2 points)

1	1	1
1	1	1
1	1	1

- (b) 5×5 Max Pooling (2 points)

1

- (c) Application of the following filter: (2 points)

1	1	1
1	1	1
1	1	1

9	9	9
9	9	9
9	9	9

3. Explain how dropout is applied during training and testing. (2 points)

Training: turn off activation for % of neurons (no forward or back prop)

Test: not used.