Your grade: 100%

Your latest: 100% • Your highest: 100% • To pass you need at least 80%. We keep your highest score.

Next item

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

$$\begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 3 & 3 & 1 \\ -1 & -3 & -3 & -1 \\ 0 & -1 & -1 & 0 \end{bmatrix}$$

- Detect 45-degree edges.
- Detect vertical edges.
- Detect horizontal edges.
- Detecting image contrast.

Expand

Correct
Correct. There is a high difference between the values in the top part from those in the bottom part of the matrix. When convolving this filter on a grayscale image, the horizontal edges will be detected.

2. Suppose your input is a 128 by 128 color (RGB) image, and you are not using a convolutional network. If the first hidden layer has 64 neurons, each one fully connected to the input, how many parameters does this hidden layer have (including the bias parameters)? 1 / 1 point

- 3145792
- 3145728
- 1048640
- 1048576

Expand

Correct
Correct, the number of inputs for each unit is $128 \times 128 \times 3$ since the input image is RGB, so we need $128 \times 128 \times 3 \times 64$ parameters for the weights and 64 parameters for the bias parameters, thus $128 \times 128 \times 3 \times 64 + 64 = 3145792$.

3. Suppose your input is a 300 by 300 color (RGB) image, and you use a convolutional layer with 100 filters that are each 5x5. How many parameters does this hidden layer have (including the bias parameters)? 1 / 1 point

- 2501
- 2600
- 7600
- 7500

Expand

Correct
Correct, you have $25 \times 3 = 75$ weights and 1 bias per filter. Given that you have 100 filters, you get 7,600 parameters for this layer.

4. You have an input volume that is $121 \times 121 \times 16$, and convolve it with 32 filters of 4×4, using a stride of 3 and no padding. What is the output volume? 1 / 1 point

- $118 \times 118 \times 16$
- $40 \times 40 \times 32$
- $118 \times 118 \times 32$
- $40 \times 40 \times 16$

Expand

Correct
Correct, using the formula $n_H^{[l]} = \frac{n_H^{[l-1]} + 2 \cdot p - f}{s} + 1$ with $n_H^{[l-1]} = 121, p = 0, f = 4$, and $s = 3$ we get 40

5. You have an input volume that is 61x61x32, and pad it using "pad=3". What is the dimension of the resulting volume (after padding)? 1 / 1 point

- 64x64x32
- 64x64x35
- 67x67x32
- 61x61x35

Expand

Correct
Yes, if the padding is 3 you add 6 to the height dimension and 6 to the width dimension.

6. You have a volume that is $121 \times 121 \times 32$, and convolve it with 32 filters of 5×5, and a stride of 1. You want to use a "same" convolution. What is the padding? 1 / 1 point

- 3
- 2
- 5
- 0

Expand

Correct
Yes, when using a padding of 2 the output volume has $n_H = \frac{121 - 5 + 4}{1} + 1$.

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 / 1 point

- $66 \times 66 \times 7$
- $22 \times 22 \times 7$
- $22 \times 22 \times 21$
- $21 \times 21 \times 21$

Expand

Correct
Yes, using the formula $n_H^{[l]} = \frac{n_H^{[l-1]} + 2 \cdot p - f}{s} + 1$ with $p = 0, f = 3, s = 3$ and $n_H^{[l-1]} = 66$.

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

- [x] Stride

Correct
Yes, although usually, we set $f = s$ this is one of the hyperparameters of a pooling layer.

- [x] Whether it is max or average.

Correct
Yes, these are the two types of pooling discussed in the lectures, and choosing which to use is considered a hyperparameter.

- [ ] $b^{[l]}$ bias.
- [ ] $W^{[l]}$ weights.

Expand

Correct
Great, you got all the right answers.

9. In lecture we talked about "parameter sharing" as a benefit of using convolutional networks. Which of the following statements about parameter sharing in ConvNets are true? (Check all that apply) 1 / 1 point

- [ ] It allows gradient descent to set many of the parameters to zero, thus making the connections sparse.
- [x] It allows a feature detector to be used in multiple locations throughout the whole input image/input volume.

Correct
Yes, by sliding a filter of parameters over the entire input volume, we make sure a feature detector can be used in multiple locations.

- [ ] It allows parameters learned for one task to be shared even for a different task (transfer learning).
- [x] It reduces the total number of parameters, thus reducing overfitting.

Correct
Yes, a convolutional layer uses parameter sharing and usually has a lot less parameters than a fully-connected layer.

Expand

Correct
Great, you got all the right answers.

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

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

Expand

Correct
Yes, each activation of the output volume is computed by multiplying the parameters from with a volumic slice of the input volume and then summing all these together.