Congratulations! You passed!

Grade received 100% **Latest Submission Grade** 100% **To pass** 80% or higher

Retake the assignment in **23h 50m**

Go to next item

1. Which of the following do you typically see in ConvNet? (Check all that apply.) **1 / 1 point**

- Use of multiple POOL layers followed by a CONV layer.
- Use of FC layers after flattening the volume to generate output classes.
- ConvNet makes exclusive use of CONV layers.
- Multiple FC layers followed by a CONV layer.

Expand

Correct

Yes, FC layers are typically used in the last few layers after flattening the volume to generate the output in classification.

2. In LeNet - 5 we can see that as we get into deeper networks the number of channels increases while the height and width of the volume decreases. True/False? **1 / 1 point**

- False
- True

Expand

Correct

Correct, since in its implementation only valid convolutions were used, without padding, the height and width of the volume were reduced at each convolution. These were also reduced by the POOL layers, whereas the number of channels was increased from 6 to 16.

3. Training a deeper network (for example, adding additional layers to the network) allows the network to fit more complex functions and thus almost always results in lower training error. For this question, assume we're referring to "plain" networks.

1 / 1 point

- (●) False
- ( ) True

Expand

Correct

Correct, Resnets are here to help us train very deep neural networks.

4. Which of the following equations captures the computations in a ResNet block?b

1 / 1 point

- $a^{[l+2]} = g\left(W^{[l+2]}\, g\left(W^{[l+1]}\, a^{[l]} + b^{[l+1]}\right) + b^{[l+2]} + a^{[l]}\right)$
- $a^{[l+2]} = g\left(W^{[l+2]}\, g\left(W^{[l+1]}\, a^{[l]} + b^{[l+1]}\right) + b^{[l+2]}\right)$
- $a^{[l+2]} = g\left(W^{[l+2]}\, g\left(W^{[l+1]}\, a^{[l]} + b^{[l+1]}\right) + b^{[l+2]}\right) + a^{[l]}$
- $a^{[l+2]} = g\left(W^{[l+2]}\, g\left(W^{[l+1]}\, a^{[l]} + b^{[l+1]}\right) + b^{[l+2]} + a^{[l]}\right) + a^{[l+1]}$

Expand

Correct

Correct. This expresses the computations of a ResNet block, where the last term $a^{[l]}$ is the shortcut connection.

5. In the best scenario when adding a ResNet block it will learn to approximate the identity function after a lot of training, helping improve the overall performance of the network. True/False?

1 / 1 point

- True
- False

Expand

Correct

Correct. When adding a ResNet block it can easily learn to approximate the identity function, thus in a worst-case scenario, it will not affect the performance of the network at all.

6. Suppose you have an input volume of dimension n_H x n_W x n_C. Which of the following statements do you agree with? (Assume that the "1x1 convolutional layer" below always uses a stride of 1 and no padding.)

1 / 1 point

- [x] You can use a 1x1 convolutional layer to reduce n_C but not n_H and n_W.

  ✓ **Correct**

  Yes, a 1x1 convolutional layer with a small number of filters is going to reduce n_C but will keep the dimensions n_H and n_W

- [x] You can use a 2D pooling layer to reduce $n_{H'}$, $n_{W'}$, but not n_C.

  ✓ **Correct**

  This is correct.

- [ ] You can use a 1x1 convolutional layer to reduce $n_{H'}$, $n_{W'}$, and n_C.

- [ ] You can use a 2D pooling layer to reduce $n_{H'}$, $n_{W'}$, and n_C.

Expand

Correct

Great, you got all the right answers.

7. Which of the following are true about the inception Network? (Check all that apply)

1 / 1 point

- [x] One problem with simply stacking up several layers is the computational cost of it.

✓ **Correct**

Correct. That is why the bottleneck layer is used to reduce the computational cost.

- [ ] Making an inception network deeper won't hurt the training set performance.
- [ ] Inception blocks allow the use of a combination of 1x1, 3x3, 5x5 convolutions, and pooling by applying one layer after the other.
- [x] Inception blocks allow the use of a combination of 1x1, 3x3, 5x5 convolutions and pooling by stacking up all the activations resulting from each type of layer.

✓ **Correct**

Correct. The use of several different types of layers and stacking up the results to get a single volume is at the heart of the inception network.

Expand

Correct
Great, you got all the right answers.

8. Models trained for one computer vision task can't be used directly in another task. In most cases, we must change the softmax layer, or the last layers of the model and re-train for the new task. True/False?

1 / 1 point

- (•) True
- ( ) False

Expand

Correct

Yes, this is a good way to take advantage of open-source models trained more or less for the task you want to do. This may also help you save a great number of computational resources and data.

9. Which of the following are true about Depthwise-separable convolutions? (Choose all that apply)

1 / 1 point

- [x] The depthwise convolution convolves each channel in the input volume with a separate filter.

  ✓ **Correct**

  Yes, the output of this kind of convolution is the same as the input.

- [x] Depthwise-separable convolutions are composed of two different types of convolutions.

  ✓ **Correct**

  Yes, it is composed of a depthwise convolution followed by a pointwise convolution.

- [ ] The depthwise convolution convolves the input volume with 1×1 filters over the depth dimension.
- [x] The pointwise convolution convolves the output volume with 1×1 filters.

  ✓ **Correct**

  Yes, the number of filters for the output of the depthwise-separable convolution is determined by the number of 1 × 1 filters used.

Expand

Correct

Great, you got all the right answers.

10. Suppose that in a MobileNet v2 Bottleneck block we have an $n \times n \times 5$ input volume, we use 30 filters for the expansion, in the depthwise convolutions we use 3×3 filters, and 20 filters for the projection. How many parameters are used in the complete block, suppose we don't use bias?

1 / 1 point

- 1101
- 80
- 8250
- 1020

Expand

Correct

Yes, the expansion filters use 5 × 30 = 150 parameters, the depthwise convolutions need 3 × 3 × 30 = 270 parameters, and the projection part 30 × 20 = 600 parameters.