Congratulations! You passed!

Grade received 90% **Latest Submission Grade** 90% **To pass** 80% or higher

Go to next item

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

- [x] Multiple CONV layers followed by a POOL layer

  Correct
  True, as seen in the case studies.

- [ ] Multiple POOL layers followed by a CONV layer
- [x] FC layers in the last few layers

  Correct
  True, fully-connected layers are often used after flattening a volume to output a set of classes in classification.

- [ ] FC layers in the first few layers

Expand

Correct
Great, you got all the right answers.

2. In order to be able to build very deep networks, we usually only use pooling layers to downsize the height/width of the activation volumes while convolutions are used with "valid" padding. Otherwise, we would downsize the input of the model too quickly. **1 / 1 point**

- ( ) True
- (•) False

Expand

Correct
Correct!

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**

- ( ) True
- (•) False

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**

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

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
- (selected) 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. For a volume of $125 \times 125 \times 64$ which of the following can be used to reduce this to a $125 \times 125 \times 32$ volume? **1 / 1 point**

- (selected) Use a 1×1 convolutional layer with a stride of 1, and 32 filters.
- Use a POOL layer of size 2×2 but with a stride of 1.
- Use a 1×1 convolutional layer with a stride of 2, and 32 filters.
- Use a POOL layer of size 2×2 with a stride of 2.

Expand

Correct
Yes, since using 1×1 convolutions is a great way to reduce the depth dimension without affecting the other dimensions.

7. Which ones of the following statements on Inception Networks are true? (Check all that apply.) **0 / 1 point**

- [x] Inception blocks usually use 1x1 convolutions to reduce the input data volume's size before applying 3x3 and 5x5 convolutions.

✓ Correct

- [x] A single inception block allows the network to use a combination of 1x1, 3x3, 5x5 convolutions and pooling.

✓ Correct

- [ ] Making an inception network deeper (by stacking more inception blocks together) can improve performance, but can also lead to overfitting and increase in computational cost.
- [ ] Inception networks incorporate a variety of network architectures (similar to dropout, which randomly chooses a network architecture on each step) and thus has a similar regularizing effect as dropout.

Expand

✗ **Incorrect**
You didn't select all the correct answers

8. When having a small training set to construct a classification model, which of the following is a strategy of transfer learning that you would use to build the model? **1 / 1 point**

- ( ) Use an open-source network trained in a larger dataset. Use these weights as an initial point for the training of the whole network.
- (•) Use an open-source network trained in a larger dataset freezing the layers and re-train the softmax layer.
- ( ) Use an open-source network trained in a larger dataset, freeze the softmax layer, and re-train the rest of the layers.
- ( ) It is always better to train a network from a random initialization to prevent bias in our model.

Expand

✓ **Correct**
Yes, this is a strategy that can provide a good result with small data.

9. Which of the following are true about Depth wise-separable convolutions? (Choose all that apply) **1 / 1 point**

- [ ] The result has always the same number of channels n_c as the input.
- [x] They have a lower computational cost than normal convolutions.

✓ **Correct**
Yes, as seen in the lectures the use of the depthwise and pointwise convolution reduces the computational cost significantly.

- [x] They combine depthwise convolutions with pointwise convolutions.

✓ **Correct**
Correct, this combination is what we call depth wise separable convolutions.

- [ ] They are just a combination of a normal convolution and a bottleneck layer.

Expand

✓ **Correct**
Great, you got all the right answers.

10. Suppose that in a MobileNet v2 Bottleneck block the input volume has shape $64 \times 64 \times 16$. If we use 32 filters for the expansion and 16 filters for the projection. What is the size of the input and output volume of the depthwise convolution, assuming a pad='same'?

1 / 1 point

- $64 \times 64 \times 32$ $64 \times 64 \times 16$
- $32 \times 32 \times 32$ $32 \times 32 \times 32$
- $64 \times 64 \times 16$ $64 \times 64 \times 32$
- $64 \times 64 \times 32$ $64 \times 64 \times 32$

Expand

Correct

Correct, the size of the input and output volume of the depthwise convolution is determined by the number of filters in the expansion.