1. When building a ConvNet, typically you start with some POOL layers followed by some CONV layers. True/False? 1 / 1 point

- (●) False
- ( ) True

Correct
Correct. It is typical for ConvNets to use a POOL layer after some Conv layers; sometimes even one POOL layer after each CONV layer; but is not common to start with POOL layers.

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

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

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

4. The following equation captures the computation in a ResNet block. What goes into the two blanks above? 1 / 1 point

$$a^{[l+2]} = g(W^{[l+2]}g(W^{[l+1]}a^{[l]} + b^{[l+1]}) + b^{l+2} + \_\_\_\_\_) + \_\_\_\_\_$$

- (●) $a^{[l]}$ and 0, respectively
- ( ) 0 and $z^{[l+1]}$, respectively

- ○ 0 and $z^{[l+1]}$, respectively
- ○ 0 and $a^{[l]}$, respectively
- ○ $z^{[l]}$ and $a^{[l]}$, respectively

Correct
Correct

5. Adding a ResNet block to the end of a network makes it deeper. Which of the following is true? 1 / 1 point

- ○ It shifts the behavior of the network to be more like the identity function.
- ○ The performance of the networks is hurt since we make the network harder to train.
- ◉ The performance of the networks doesn't get hurt since the ResNet block can easily approximate the identity function.
- ○ The number of parameters will decrease due to the shortcut connections.

Correct
Yes, as noted in the lectures in a ResNet block the computations are given by $a^{[l+2]} = g(W^{[l+2]}a^{[l+1]} + b^{[l+2]} + a^{[l]})$ thus if $W^{[l+2]}$ and $b^{[l+2]}$ are zero then we get the identity function.

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

- ◉ Use a 1×1 convolutional layer with a stride of 1, and 32 filters.
- ○ Use a POOL layer of size 2×2 with a stride of 2.
- ○ Use a 1×1 convolutional layer with a stride of 2, and 32 filters.
- ○ Use a POOL layer of size 2×2 but with a stride of 1.

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

7. Which of the following are true about bottleneck layers? (Check all that apply) 1 / 1 point

- ☑ The use of bottlenecks doesn't seem to hurt the performance of the network.

Correct

7. Which of the following are true about bottleneck layers? (Check all that apply) 1 / 1 point

- [x] The use of bottlenecks doesn't seem to hurt the performance of the network.

Correct
Yes, although it reduces the computational cost significantly.

- [x] By adding these layers we can reduce the computational cost in the inception modules.

Correct
Yes, by using the 1×1 convolutional layers we can reduce the depth of the volume and help reduce the computational cost of applying other convolutional layers with different filter sizes.

- [ ] The bottleneck layer has a more powerful regularization effect than Dropout layers.
- [ ] Bottleneck layers help to compress the 1x1, 3x3, 5x5 convolutional layers in the inception network.

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

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] 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.

- [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] 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.

○ False

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

☑ 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 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.

☑ 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.

☐ The depthwise convolution convolves the input volume with 1×1 filters over the depth dimension.

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

◉ 1020

○ 8250

○ 1101

○ 80

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