Congratulations! You passed!

Grade received 80% **To pass** 80% or higher

Go to next item

Deep Convolutional Models

Latest Submission Grade 80%

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

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 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. The motivation of Residual Networks is that very deep networks are so good at fitting complex functions that when training them we almost always overfit the training data. True/False? **0 / 1 point**

- True ✓
- False

Expand

Incorrect
No, very deep neural networks are hard to train and a deeper network does not always imply lower training error. Residual Networks allow us to 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**

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

- $z^{[l]}$ and $a^{[l]}$, respectively
- 0 and $z^{[l+2]}$, respectively
- $a^{[l]}$ and 0, respectively ✓
- 0 and $a^{[l]}$, respectively

Expand

Correct
Correct

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

- Using a skip-connection helps the gradient to backpropagate and thus helps you to train deeper networks ✓
  - **Correct** This is true.
- The skip-connections compute a complex non-linear function of the input to pass to a deeper layer in the network.
- The skip-connection makes it easy for the network to learn an identity mapping between the input and the output within the ResNet block. ✓
  - **Correct** This is true.
- A ResNet with L layers would have on the order of L^2 skip connections in total.

Expand

Correct
Great, you got all the right answers.

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

Expand

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 the inception Network? (Check all that apply) **1 / 1 point**

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

Expand

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

9. In Depthwise Separable Convolution you: **0 / 1 point**

- Perform two steps of convolution. ✓
  - **Correct**
- For the "Depthwise" computations each filter convolves with only one corresponding color channel of the input image. ✓
  - **Correct**
- The final output is of the dimension $n_{out} \times n_{out} \times n_{c}$ (where n_{c} is the number of color channels of the input image). ✓
  - **This should not be selected** Incorrect
- The final output is of the dimension $n_{out} \times n_{out} \times n_{c'}$ (where $n_{c'}$ is the number of filters used in the pointwise convolution step).
- For the "Depthwise" computations each filter convolves with all of the color channels of the input image.
- Perform one step of convolution.
- You convolve the input image with n_{c} number of $n_{f} \times n_{f}$ filters (n_{c} is the number of color channels of the input image).
- You convolve the input image with a filter of $n_{f} \times n_{f} \times n_{c}$ where n_{c} acts as the depth of the filter (n_{c} is the number of color channels of the input image). ✓
  - **This should not be selected** Incorrect

Expand

Incorrect
You didn't select all the correct 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**

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

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.