Congratulations! You passed!

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

Retake the assignment in 23h 51m

Go to next item

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

- ○ True
- ◉ False

Expand

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.

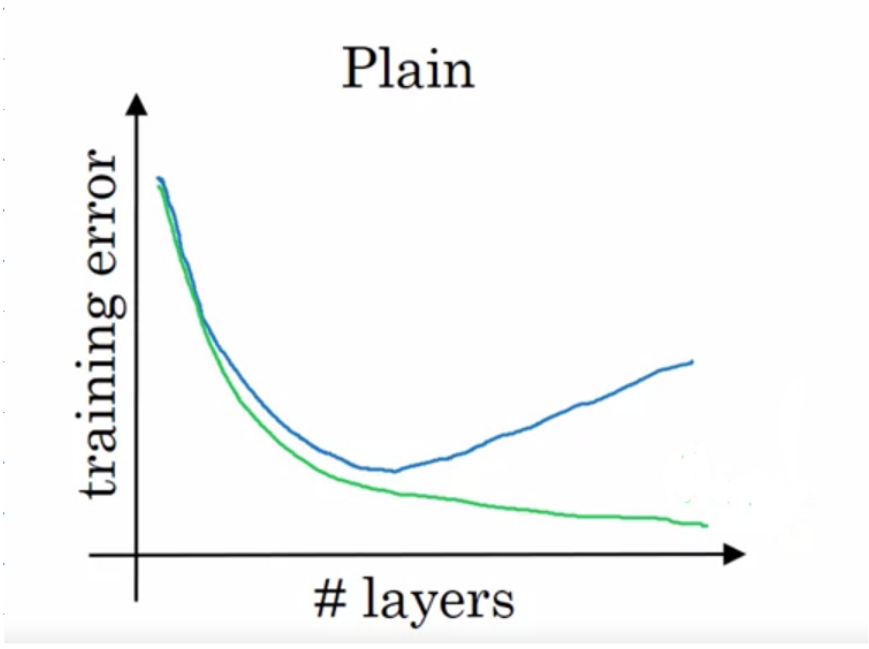
2. LeNet - 5 made extensive use of padding to create valid convolutions, to avoid increasing the number of channels after every convolutional layer. True/False? 1 / 1 point

- ○ True
- ◉ False

Expand

Correct
Yes, back in 1998 when the corresponding paper of LeNet - 5 was written padding wasn't used.

3. Based on the lectures, in the following picture, which curve corresponds to the expected behavior in theory, and which one corresponds to the behavior we get in practice? This when using plain neural networks. 0 / 1 point



- The blue one depicts the results in theory, and also in practice.
- The green one depicts the results in theory, and the blue one the reality.
- The blue one depicts the theory, and the green one the reality.
- The green one depicts the results in theory, and also in practice.

Expand

Incorrect
No, in practice the training error goes up when we add too many layers to our network.

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
- $z^{[l]}$ and $a^{[l]}$, 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**

- [ ] A ResNet with L layers would have on the order of L^2 skip connections in total.
- [ ] The skip-connections compute a complex non-linear function of the input to pass to a deeper layer in the network.
- [x] 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.

- [x] Using a skip-connection helps the gradient to backpropagate and thus helps you to train deeper networks

Correct
This is true.

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 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 with a stride of 2.
- Use a POOL layer of size 2×2 but with a stride of 1.

Expand

Correct
Yes, since using $$1 \times 1$$ 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**

- [ ] Bottleneck layers help to compress the 1x1, 3x3, 5x5 convolutional layers in the inception network.
- [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.
- [x] The use of bottlenecks doesn't seem to hurt the performance of the network.

Correct
Yes, although it reduces the computational cost significantly.

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**

- 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, freeze the softmax layer, and re-train the rest of the layers.
- 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.

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**

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

- [ ] The result has always the same number of channels n_c as the input.

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

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

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'? **0 / 1 point**

- ○ $32 \times 32 \times 32$ $32 \times 32 \times 32$
- ○ $64 \times 64 \times 16$ $64 \times 64 \times 32$
- ◉ $64 \times 64 \times 32$ $64 \times 64 \times 16$
- ○ $64 \times 64 \times 32$ $64 \times 64 \times 32$

Expand

Incorrect
Incorrect, the input and output volume of the depthwise convolution are the same.