Deep Convolutional Models

1.
Question 1

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

0 / 1 point

Expand

Incorrect

No, this is not a common practice.

2.
Question 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

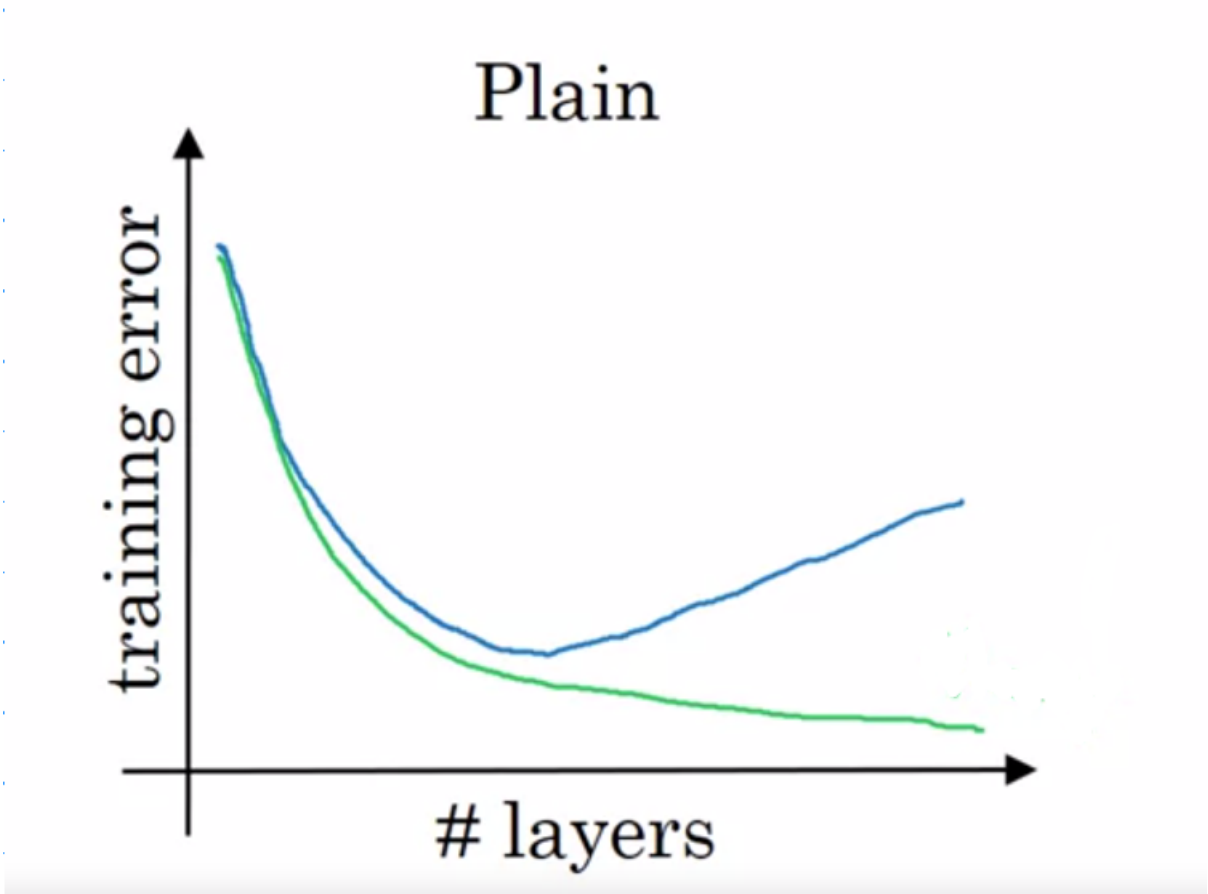
Expand

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.
Question 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.



1 / 1 point

Expand

Correct

Yes, in theory, we expect that as we increase the number of layers the training error decreases; but in practice after a certain number of layers the error increases.

4.
Question 4

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

1 / 1 point

Expand

Correct

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

5.
Question 5

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

1 / 1 point

Expand

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.
Question 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.)

0 / 1 point

Expand

Incorrect

You didn't select all the correct answers

7.
Question 7

Which of the following are true about bottleneck layers? (Check all that apply)

0 / 1 point

Expand

Incorrect

You chose the extra incorrect answers.

8.
Question 8

Which of the following are common reasons for using open-source implementations of ConvNets (both the model and/or weights)? Check all that apply.

1 / 1 point

Expand

Correct

Great, you got all the right answers.

9.
Question 9

Which of the following are true about Depth wise-separable convolutions? (Choose all that apply)

1 / 1 point

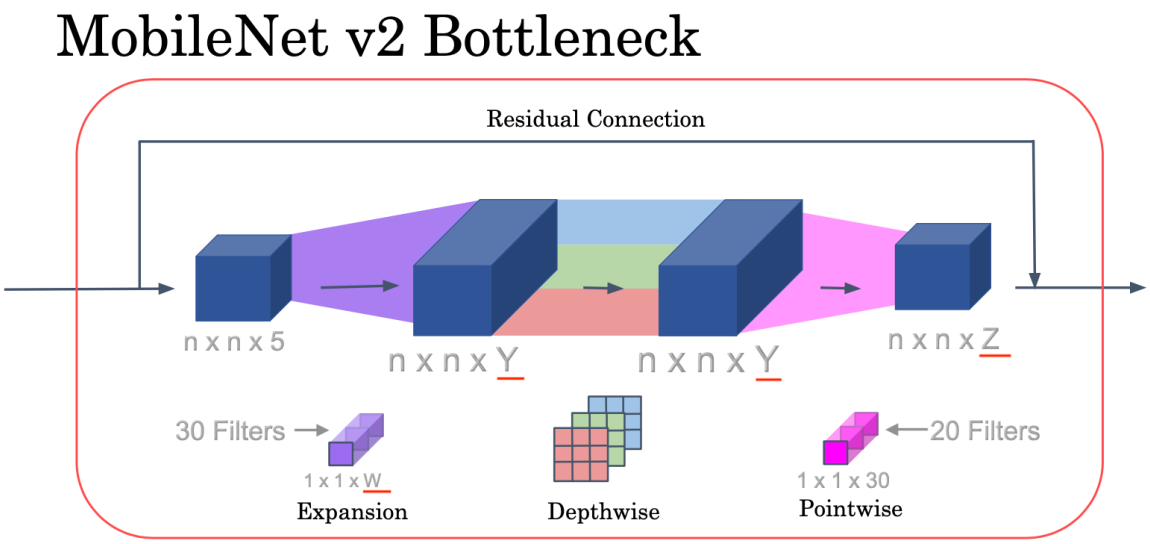
Expand

Correct

Great, you got all the right answers.

10.
Question 10

Fill in the missing dimensions shown in the image below (marked W, Y, Z).



0 / 1 point

Expand

Incorrect

Incorrect! To improve your understanding, watch the lecture MobileNet Architecture.

1.
Question 1

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

0 / 1 point

Expand

Incorrect

Incorrect. 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.
Question 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

Expand

Correct

Correct!

3.
Question 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

Expand

Correct

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

4.
Question 4

The computation of a ResNet block is expressed in the equation:

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

C A B

Which part corresponds to the skip connection?

0 / 1 point

Expand

Incorrect

No, this corresponds to the bias parameter of the $l+1$ layer.

5.
Question 5

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

1 / 1 point

Expand

Correct

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

6.
Question 6

For a volume of $125\times125\times64125\times125\times64$ which of the following can be used to reduce this to a $125\times125\times32125\times125\times32$ volume?

1 / 1 point

Expand

Correct

Yes, since using $1\times11\times1$ convolutions is a great way to reduce the depth dimension without affecting the other dimensions.

7.
Question 7

Which of the following are true about bottleneck layers? (Check all that apply)

1 / 1 point

Expand

Correct

Great, you got all the right answers.

8.
Question 8

Which of the following are common reasons for using open-source implementations of ConvNets (both the model and/or weights)? Check all that apply.

1 / 1 point

Expand

Correct

Great, you got all the right answers.

9.
Question 9

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

0 / 1 point

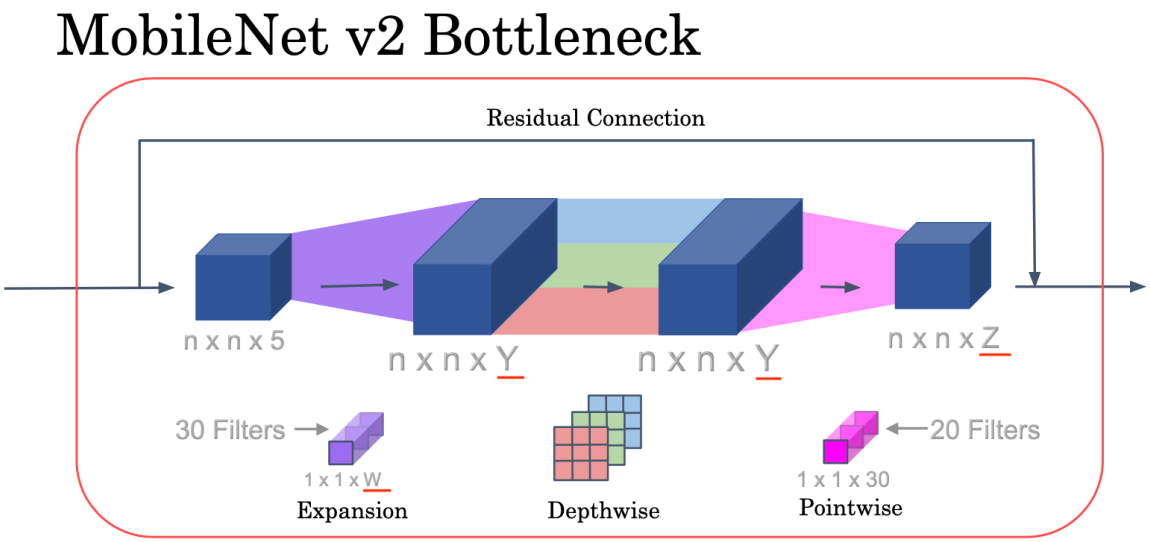
Expand

Incorrect

You didn't select all the correct answers

10.
Question 10

Fill in the missing dimensions shown in the image below (marked W, Y, Z).



1 / 1 point

Expand

Correct

1.
Question 1

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

0 / 1 point

Expand

Incorrect

You didn't select all the correct answers

2.
Question 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

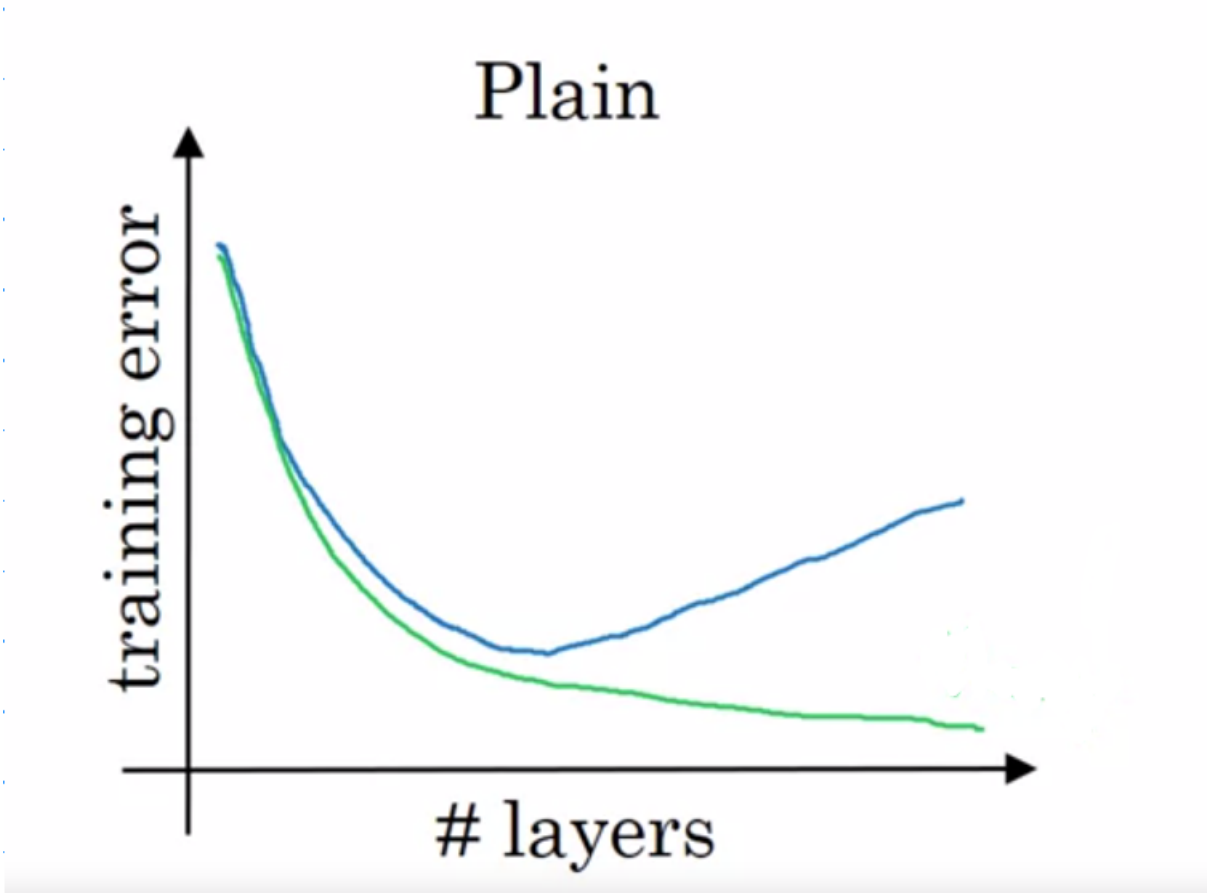
Expand

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.
Question 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.



1 / 1 point

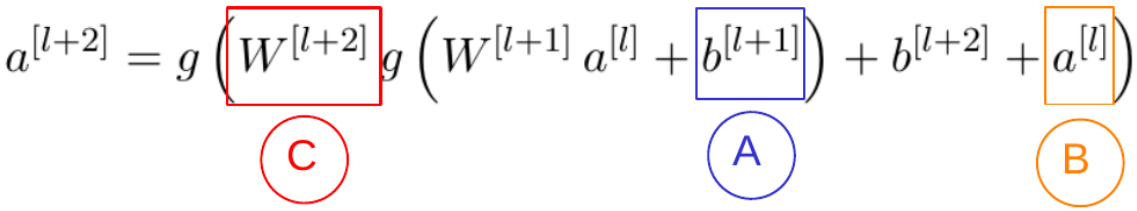
Expand

Correct

Yes, in theory, we expect that as we increase the number of layers the training error decreases; but in practice after a certain number of layers the error increases.

4.
Question 4

The computation of a ResNet block is expressed in the equation:



Which part corresponds to the skip connection?

0 / 1 point

Expand

Incorrect

No, this equation represents the computations of a ResNet as presented in the lectures.

5.
Question 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

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.
Question 6

For a volume of $125\times125\times64125\times125\times64$ which of the following can be used to reduce this to a $125\times125\times32125\times125\times32$ volume?

1 / 1 point

Expand

Correct

Yes, since using $1\times11\times1$ convolutions is a great way to reduce the depth dimension without affecting the other dimensions.

7.
Question 7

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

0 / 1 point

Expand

Incorrect

You didn't select all the correct answers

8.
Question 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

Expand

Correct

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

9.
Question 9

In Depthwise Separable Convolution you:

0 / 1 point

Expand

Incorrect

You didn't select all the correct answers

10.
Question 10

Suppose that in a MobileNet v2 Bottleneck block the input volume has shape $64 \times 64 \times 1664 \times 64 \times 16$. If we use 3232 filters for the expansion and 1616 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

Expand

Incorrect

Incorrect, the expansion phase doesn't change the width or height of the input volume.

1. Which of the following do you typically see in ConvNet? (Check all that apply.) **0 / 1 point**

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

Expand

Incorrect
No, this is not a common practice.

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**

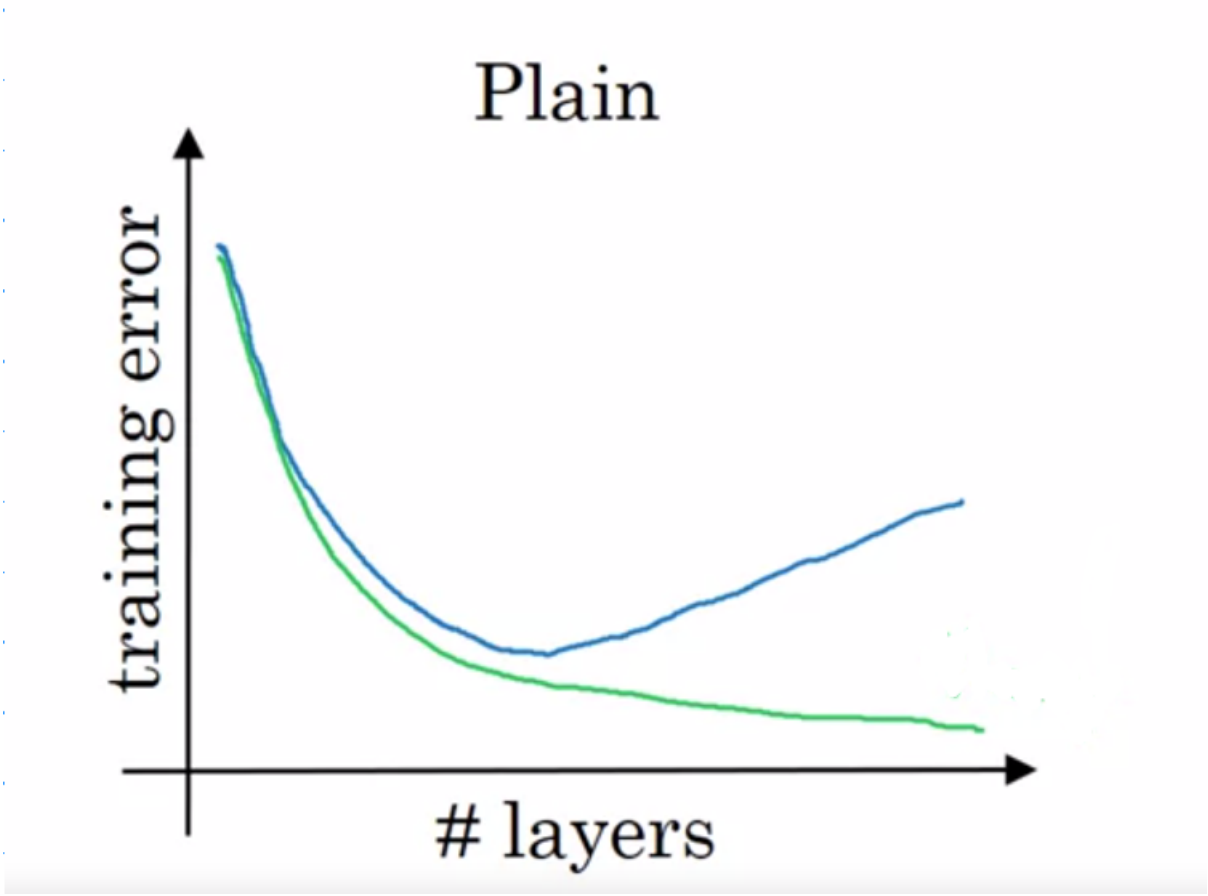
- ○ False
- ◉ True

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.
Question 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.



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

Expand

Correct

Yes, in theory, we expect that as we increase the number of layers the training error decreases; but in practice after a certain number of layers the error increases.

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

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

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

Expand

Correct
Correct

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 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.
- [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

Expand

Correct
Great, you got all the right answers.

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

- [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.

- [x] Inception blocks allow the use of a combination of 1x1, 3x3, 5x5 convolutions, and pooling by applying one layer after the other.

  This should not be selected
  Incorrect. An inception block stacks up the result of applying the different size convolutions and the pooling in a single volume.

- [x] Making an inception network deeper won't hurt the training set performance.

  This should not be selected
  Incorrect. As seen in the lectures in practice when stacking more layers the training performance might start increasing instead of decreasing.

- [ ] 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.

8. Parameters 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? **0 / 1 point**

- ○ True
- ◉ False

Expand

Incorrect
No, 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)

- ☑ 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.
- ☑ 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.

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
- ( ) 80
- ( ) 1101
- ( ) 8250

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.

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

0 / 1 point

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

Expand

Incorrect
No, ConvNet makes use of other types of layers.

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

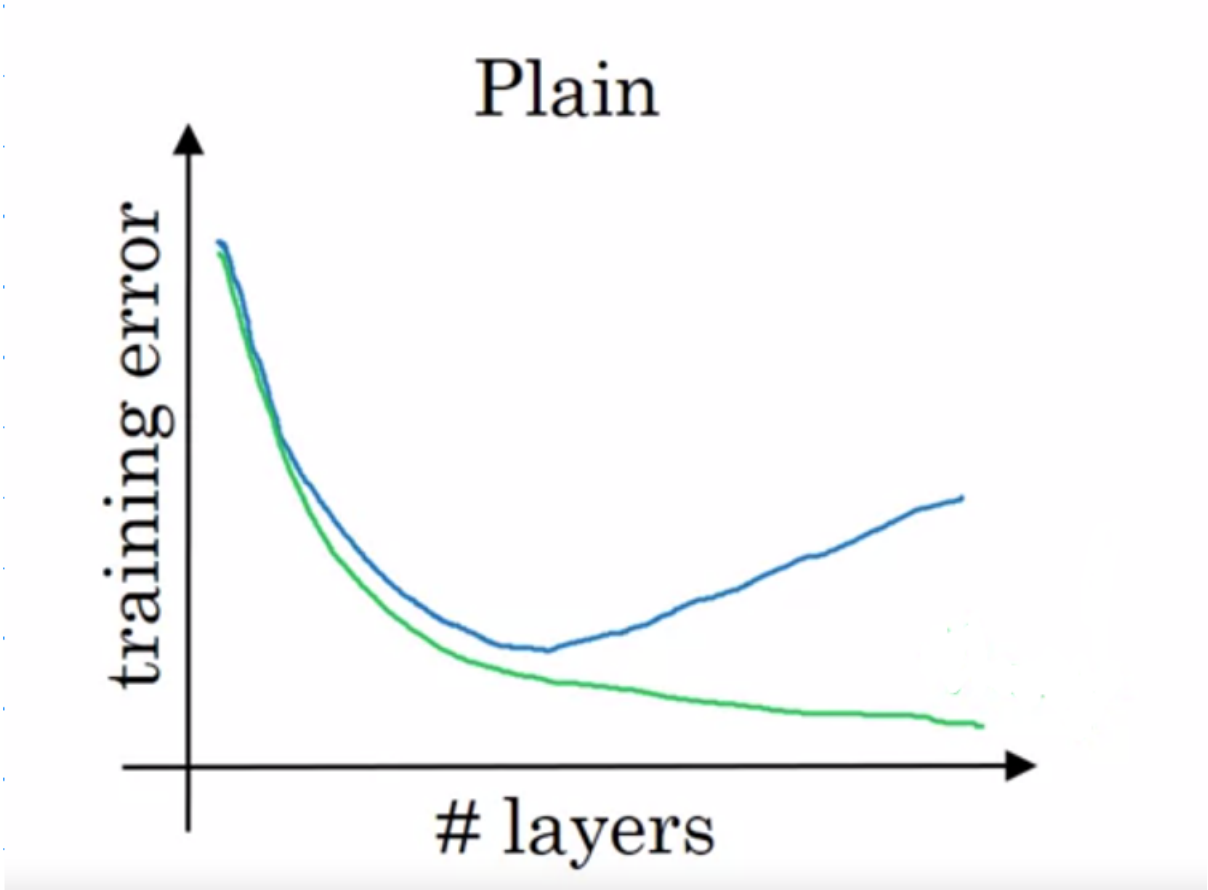
False

Expand

Correct
Correct!

3. 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. 1 / 1 point

Question 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.



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

Expand

Correct
Yes, in theory, we expect that as we increase the number of layers the training error decreases; but in practice after a certain number of layers the error increases.

4. Which of the following equations captures the computations in a ResNet block?b **1 / 1 point**

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

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**

- (•) False
- ( ) True

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. 1×1 convolutions are the same as multiplying by a single number. True/False? **0 / 1 point**

- (•) True
- ( ) False

Expand

Incorrect
No, a 1×1 layer doesn't act as a single number because it makes a sum over the depth of the volume.

7. Which of the following are true about the inception Network? (Check all that apply) **1 / 1 point**

- [ ] Inception blocks allow the use of a combination of 1x1, 3x3, 5x5 convolutions, and pooling by applying one layer after the other.
- [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.

- [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.

- [ ] Making an inception network deeper won't hurt the training set performance.

Expand

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

8. Which of the following are common reasons for using open-source implementations of ConvNets (both the model and/or weights)? Check all that apply. **1 / 1 point**

- [ ] A model trained for one computer vision task can usually be used to perform data augmentation for a different computer vision task.
- [x] It is a convenient way to get working with an implementation of a complex ConvNet architecture.

  ✓ **Correct**
  True

- [x] Parameters trained for one computer vision task are often useful as pre-training for other computer vision tasks.

  ✓ **Correct**
  True

- [ ] The same techniques for winning computer vision competitions, such as using multiple crops at test time, are widely used in practical deployments (or production system deployments) of ConvNets.

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

- [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.

- [ ] They are just a combination of a normal convolution and a bottleneck layer.
- [ ] The result has always the same number of channels n_c as the input.
- [x] They combine depthwise convolutions with pointwise convolutions.

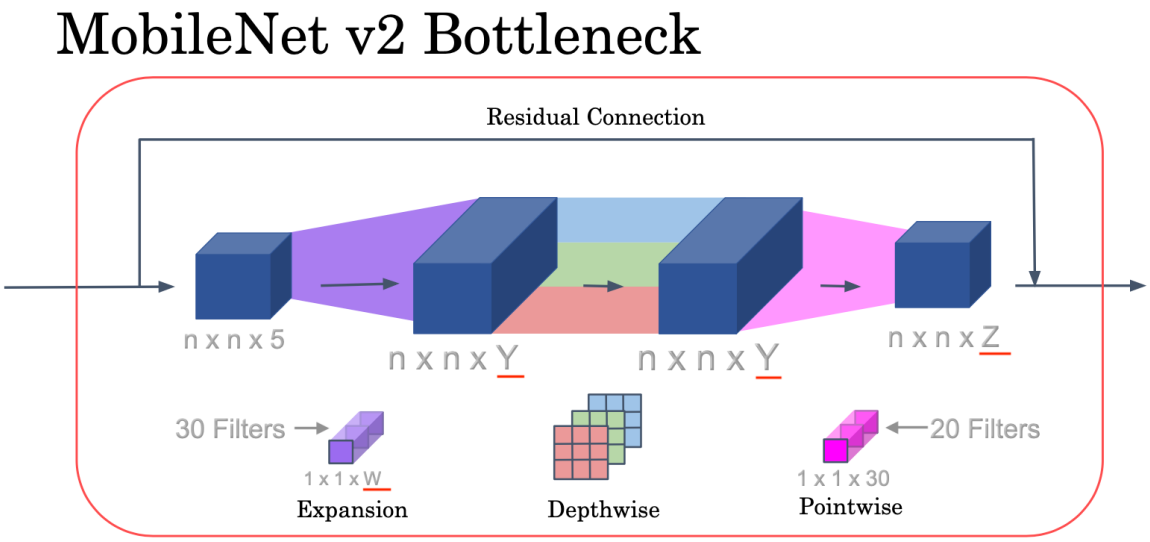
  Correct
  Correct, this combination is what we call depth wise separable convolutions.

Expand

Correct
Great, you got all the right answers.

10.
Question 10

Fill in the missing dimensions shown in the image below (marked W, Y, Z).



- (●) W = 5, Y = 30, Z = 20
- ( ) W = 30, Y = 20, Z =20
- ( ) W = 5, Y = 20, Z = 5
- ( ) W = 30, Y = 30, Z = 5

Expand

Correct