Week 2: Deep convolutional models

1. Which of the following do you typically see as you move to deeper layers in a CNN?

   Ans: n_H and n_W decrease, while n_C increases.

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

   Ans:

   - Multiple `CONV` layers followed by a `POOL` layer.
   - `FC` layers in last few layers.

3. 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.

   Ans: False

4. 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. **Ans:** False; there could be over-fitting and memorization too.

5. The following equation captures the computation in a ResNet block. What goes into the two places marked X and Y? $a^{[l+2]} = g(W^{[l+2]} * g(W^{[l+1]} a^{[l]} + b^{[l+1]}) + b^{[l+2]} + X) + Y$

   Ans: X = $a^{[l]}$, Y = 0 respectively. The skip term is added before the non-linear computation at $[l+2]^{th}$ stage.

6. Which ones of the following statements on Residual Networks are true?

   Ans:

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

- The skip-connection makes it easy to learn an identity mapping between the input and the output within the ResNet block

7. Suppose you have an input volume of dimension $64 \times 64 \times 16$. How many parameters would a single 1×1 convolutional filter have?

   Ans: The filter would be $1 \times 1 \times 16$ with a bias. Hence, 17 parameters.

8. Suppose you have an input volume of dimension $n_H x n_W x n_C$. Which of the following statements you agree with? (Assume that 1×1 `CONV` uses a `stride` $= 1$ and no padding.)

   Ans:

   - You can use a $\times 1$ convolutional layer to reduce n_C but not n_H, n_W.
   - You can use a pooling layer to reduce n_H, n_W, but not n_C.

9. Which ones of the following statements on Inception Networks are true?

   Ans:

   - A single inception block allows the network to use a combination of 1x1, 3x3, 5x5 convolutions and pooling.
   - Making an inception network deeper (by stacking more inception blocks together) should not hurt training set performance.
   - Inception blocks usually use a 1×1 filter to reduce input data, before applying 3×3, 5×5 filters

10. Which of the following are common reasons for using open-source implementations of CNNs?

    Ans:

    - A convenient way to get working an implementation of a complex CNN architecture.
    - Parameters trained for one computer vision task are often useful as pretraining for other computer vision tasks.