Your grade: 100%

Your latest: **100%** • Your highest: **100%** • To pass you need at least 80%. We keep your highest score.

Next item →

1. Face verification and face recognition are the two most common names given to the task of comparing a new picture against one person's face. True/False? **1 / 1 point**

- (•) False
- ( ) True

Correct
Correct. This is the description of face verification, but not of face recognition.

2. Why is the face verification problem considered a one-shot learning problem? Choose the best answer. **1 / 1 point**

- ( ) Because we have only have to forward pass the image one time through our neural network for verification.
- (•) Because we might have only one example of the person we want to verify.
- ( ) Because of the sensitive nature of the problem, we won't have a chance to correct it if the network makes a mistake.
- ( ) Because we are trying to compare to one specific person only.

Correct
Correct. One-shot learning refers to the amount of data we have to solve a task.

3. You want to build a system that receives a person's face picture and determines if the person is inside a workgroup. You have pictures of all the faces of the people currently in the workgroup, but some members might leave, and some new members might be added. To train a system to solve this problem using the triplet loss you must collect pictures of different faces from only the current members of the team. True/False? **1 / 1 point**

- (•) False
- ( ) True

Correct
Correct. Although it is necessary to have several pictures of the same person, it is not absolutely necessary that all the pictures only come from current members of the team.

4. In the triplet loss: **1 / 1 point**

$$\max\left(\|f(A)-f(P)\|^2-\|f(A)-f(N)\|^2+\alpha,0\right)$$

Which of the following are true about the triplet loss? Choose all that apply.

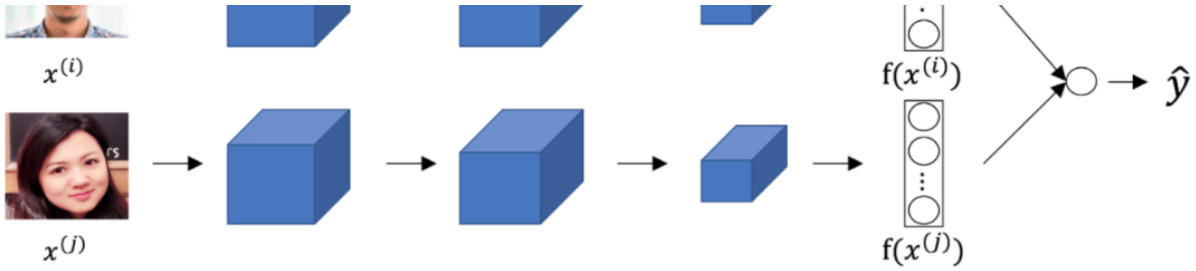
- [x] We want that $\|f(A)-f(P)\|^2 < \|f(A)-f(N)\|^2$ so the negative images are further away from the anchor than the positive images.

Correct
Correct. Being a positive image the encoding of P should be close to the encoding of A.

- [ ] α is a trainable parameter of the Siamese network.
- [ ] A the anchor image is a hyperparameter of the Siamese network.
- [x] $f(A)$ represents the encoding of the Anchor.

Correct
Correct. f represents the network that is in charge of creating the encoding of the images, and A represents the anchor image.

5. Consider the following Siamese network architecture: **1 / 1 point**



The upper and lower networks share parameters to have a consistent encoding for both images. True/False?

- (●) True
- ( ) False

Correct
Correct. Part of the idea behind the Siamese network is to compare the encoding of the images, thus they must be consistent.

6. You train a ConvNet on a dataset with 100 different classes. You wonder if you can find a hidden unit which responds strongly to pictures of cats. (I.e., a neuron so that, of all the input/training images that strongly activate that neuron, the majority are cat pictures.) You are more likely to find this unit in layer 4 of the network than in layer 1. **1 / 1 point**

- (●) True
- ( ) False

Correct
Yes, this neuron understands complex shapes (cat pictures) so it is more likely to be in a deeper layer than in the first layer.

7. In neural style transfer, we train the pixels of an image, and not the parameters of a network. **1 / 1 point**

- (●) True
- ( ) False

Correct
Correct. Neural style transfer compares the high-level features of two images and modifies the pixels of one of them in order to look artistic.

8. In the deeper layers of a ConvNet, each channel corresponds to a different feature detector. The style matrix $G^{[l]}$ measures the degree to which the activations of different feature detectors in layer l vary (or correlate) together with each other. **1 / 1 point**

- (●) True
- ( ) False

Correct
Yes, the style matrix $G^{[l]}$ can be seen as a matrix of cross-correlations between the different feature detectors.

9. In neural style transfer, which of the following better express the gradients used? **1 / 1 point**

- ( ) $\frac{\partial J}{\partial S}$
- ( ) Neural style transfer doesn't use gradient descent since there are no trainable parameters.
- ( ) $\frac{\partial J}{\partial W^{[l]}}$
- (●) $\frac{\partial J}{\partial G}$

Correct
Correct, we use the gradient of the cost function over the value of the pixels of the generated image.

10. You are working with 3D data. The input "image" has size $64 \times 64 \times 64 \times 3$, if you apply a convolutional layer with 16 filters of size $4 \times 4 \times 4$, zero padding and stride 2. What is the size of the output volume? **1 / 1 point**

- ( ) $31 \times 31 \times 31 \times 3$.
- ( ) $61 \times 61 \times 61 \times 14$.

- (●) $31 \times 31 \times 31 \times 16$.
- ( ) $64 \times 64 \times 64 \times 3$.

Correct

Correct, we can use the formula $\lfloor \frac{n^{[l-1]} - f + 2 \times p}{s} \rfloor + 1 = n^{[l]}$ to the three first dimensions.