Face verification requires comparing a new picture against one person's face, whereas face recognition requires comparing a new picture against K persons' faces. 1 point

- ◯ True
- ◯ False

2.Why do we learn a function $d(img1, img2)$ for face verification? (Select all that apply.) 1 point

- ☐ This allows us to learn to predict a person's identity using a softmax output unit, where the number of classes equals the number of persons in the database plus 1 (for the final "not in database" class).
- ☐ We need to solve a one-shot learning problem.
- ☐ This allows us to learn to recognize a new person given just a single image of that person.
- ☐ Given how few images we have per person, we need to apply transfer learning.

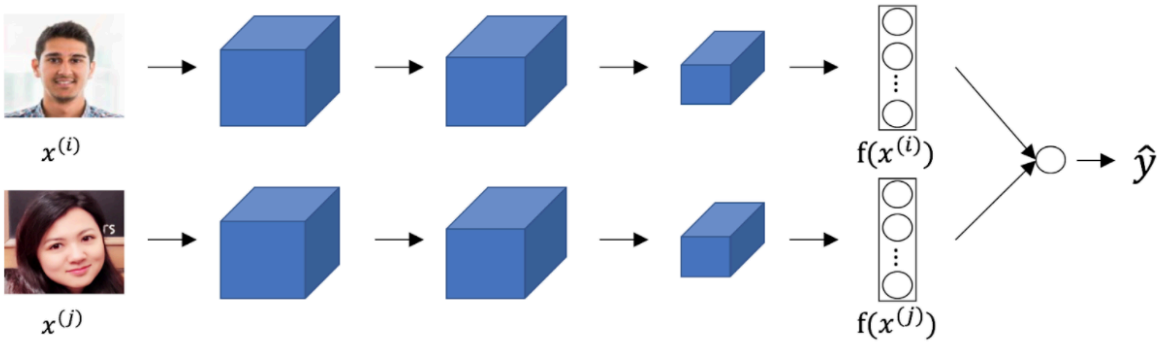
3.You want to build a system that receives a person's face picture and determines if the person is inside a workgroup. You have pictures of all the faces of the people currently in the workgroup, but some members might leave, and some new members might be added. To train a system to solve this problem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

- ◯ False
- ◯ True

4.Which of the following is a correct definition of the triplet loss? Consider that $\alpha > 0$. (We encourage you to figure out the answer from first principles, rather than just refer to the lecture.) 1 point

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

5.Consider the following Siamese network architecture: 1 point



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

- ◯ False
- ◯ True

6.You train a ConvNet on a dataset with cats, dogs, birds, and other types of animals. You try to find a filter that strongly responds to horizontal edges. You are more likely to find this filter in layer 6 of the network than in layer 1. True/False? 1 point

- False
- True

7. Neural style transfer is trained as a supervised learning task in which the goal is to input two images (x), and train a network to output a new, synthesized image (y). 1 point

- True
- False

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 point

- True
- False

9. In neural style transfer, we can't use gradient descent since there are no trainable parameters. True/False? 1 point

- False
- True

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 point

- $64 \times 64 \times 64 \times 3$.
- $31 \times 31 \times 31 \times 16$.
- $31 \times 31 \times 31 \times 3$.
- $61 \times 61 \times 61 \times 14$.