Your grade: 100%

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

Next item →

1. Which of the following do you agree with? **1 / 1 point**

- ○ Face verification requires K comparisons of a person's face.
- ○ Face recognition requires comparing pictures against one person's face.
- ◉ Face recognition requires K comparisons of a person's face.

Correct
Correct, in face recognition we compare the face of one person to K to classify the face as one of those K or not.

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

- ☐ Given how few images we have per person, we need to apply transfer learning.
- ☐ 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).
- ☑ This allows us to learn to recognize a new person given just a single image of that person.

Correct

Yes.

- [x] We need to solve a one-shot learning problem.

Correct
This is true as explained in the lecture.

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. Triplet loss: **1 / 1 point**

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

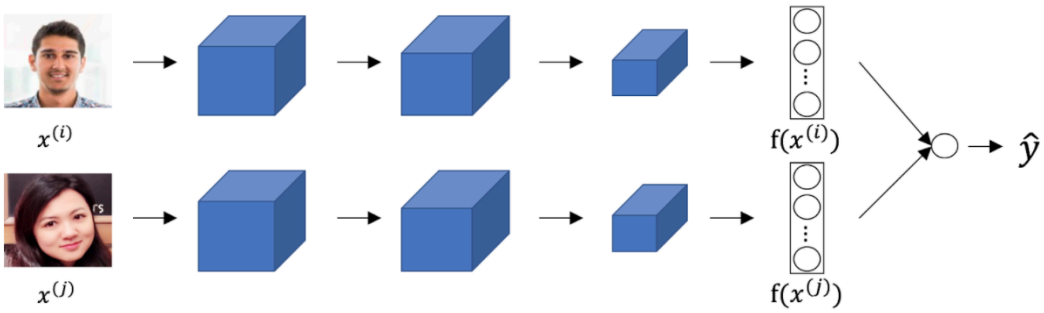
is larger in which of the following cases?

- ( ) When the encoding of A is closer to the encoding of P than to the encoding of N.
- (•) When the encoding of A is closer to the encoding of N than to the encoding of P.

- When $A = P$ and $A = N$.

Correct
Correct. In this case $\|f(A) - f(P)\|^2 - \|f(A) - f(N)\|^2$ is positive thus the triplet loss gives a positive value larger than α.

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?

- False
- True

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 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 / 1 point**

- False

- ○ True

Correct
Correct. Edges are a very low-level feature, thus it is more likely to find such a feature detector in the first layers of the network.

7. Neural style transfer uses images Content C, Style S. The loss function used to generate image G is composed of which of the following: (Choose all that apply.) **1 / 1 point**

- ☑ $J_{content}$ that compares C and G.

Correct
Correct, in neural style transfer we are interested in the similarity between S and G, and the similarity between G and C.

- ☐ J_{corr} that compares C and S.
- ☑ J_{style} that compares S and G.

Correct
Correct, in neural style transfer we are interested in the similarity between S and G, and the similarity between G and C.

- ☐ T that calculates the triplet loss between S, G, and C.

8. In neural style transfer, we define style as: **1 / 1 point**

- ○ $\|a^{[l](S)} - a^{[l](G)}\|^2$ the distance between the activation of the style image and the content image.
- ○ The correlation between the generated image G and the style image S .

- ○ The correlation between the activation of the content image C and the style image S.
- ◉ The correlation between activations across channels of an image.

Correct
Correct, this correlation is represented by $G_{kk'}^{[l](I)}$ for the image I.

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

- ○ True
- ◉ False

Correct
Correct. We use gradient descent on the cost function $J(G)$ and we update the pixel values of the generated image G.

10. You are working with 3D data. The input "image" has size $32 \times 32 \times 32 \times 3$, if you apply a convolutional layer with 16 filters of size $4 \times 4 \times 4$, zero padding and stride 1. What is the size of the output volume?

- ◉ $29 \times 29 \times 29 \times 16$.
- ○ $31 \times 31 \times 31 \times 16$.
- ○ $29 \times 29 \times 29 \times 3$.
- ○ $29 \times 29 \times 29 \times 13$.

Correct

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