Congratulations! You passed!

Grade received 100% **Latest Submission Grade** 100% **To pass** 80% or higher

Go to next item

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

- (•) Face recognition requires K comparisons of a person's face.
- ( ) Face verification requires K comparisons of a person's face.
- ( ) Face recognition requires comparing pictures against one person's face.

Expand

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**

- [ ] 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).
- [x] 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.

- [ ] Given how few images we have per person, we need to apply transfer learning.

Expand

Correct
Great, you got all the right answers.

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 get many persons and take several pictures of each one. Which of the following do you agree with? (Select the best answer.) **1 / 1 point**

- ( ) You take several pictures of the same person because this way you can get more pictures to train the network efficiently since you already have the person in place.
- ( ) It would be best to increase the number of persons in the dataset by taking only one picture of each person to have a more representative set of the population.
- (•) You take several pictures of the same person to train $d(\text{img}_1, \text{img}_2)$ using the triplet loss.
- ( ) You shouldn't use persons outside the workgroup you are interested in because that might create a high variance in your model.

Expand

Correct
Correct. To train using the triplet loss you need several pictures of the same person.

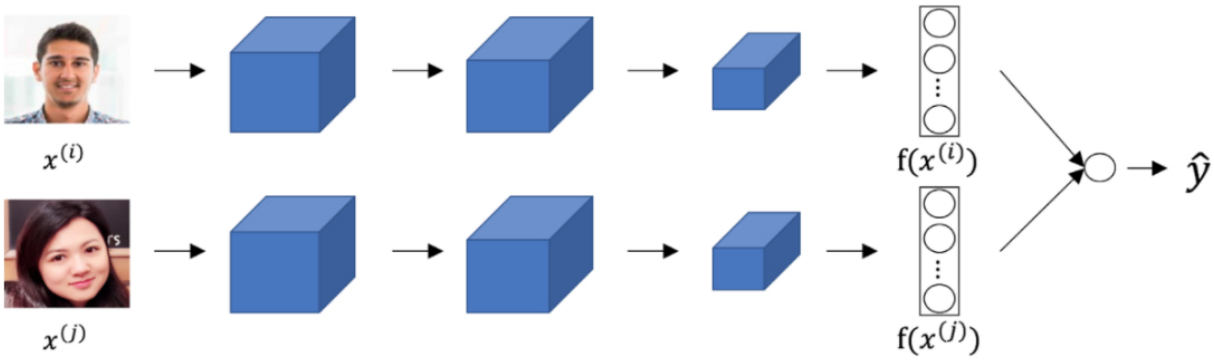
4. Which of the following is a correct definition of the triplet loss? Consider that $\alpha > 0$. (We encourage you to figure out the answer from first principles, rather than just refer to the lecture.) 1 / 1 point

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

Expand

Correct
Correct

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

Expand

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

6. Our intuition about the layers of a neural network tells us that units that respond more to complex features are more likely to be in deeper layers. True/False? 1 / 1 point

- True
- False

Expand

Correct
Correct. Neurons that understand more complex shapes are more likely to be in deeper layers of a neural network.

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 / 1 point

- True
- False

Expand

Correct
Yes, Neural style transfer is about training the pixels of an image to make it look artistic, it is not learning any parameters.

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

Expand

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

- True
- False

Expand

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 $64 \times 64 \times 64 \times 3$, 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**

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

Expand

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.