Congratulations! You passed!

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

Go to next item

1. 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 / 1 point**

- ○ False
- ◉ True

Expand

Correct
Correct.

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

- ☑ We need to solve a one-shot learning problem.

  Correct
  This is true as explained in the lecture.

- ☑ This allows us to learn to recognize a new person given just a single image of that person.

  Correct
  Yes.

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

Expand

Correct
Great, you got all the right answers.

3. In order to train the parameters of a face recognition system, it would be reasonable to use a training set comprising 100,000 pictures of 100,000 different persons. **1 / 1 point**

- ◉ False
- ○ True

Expand

Correct
Correct, to train a network using the triplet loss you need several pictures of the same person.

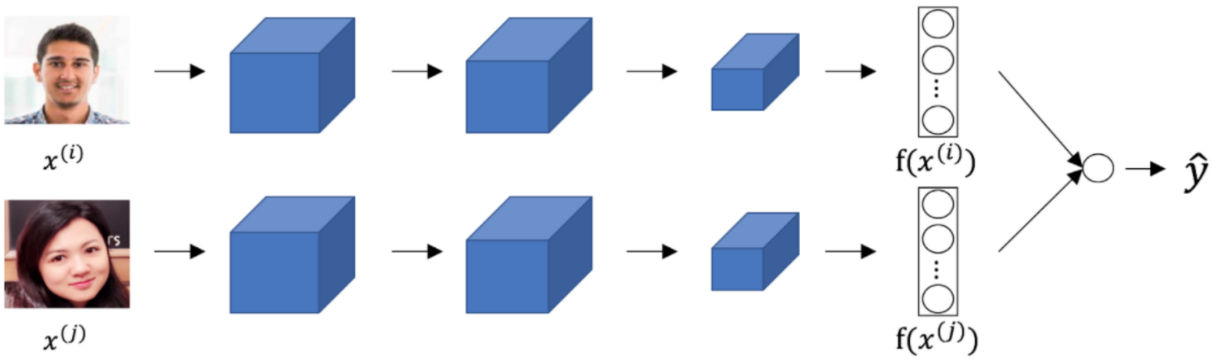
4. Which of the following is a correct definition of the triplet loss? Consider that $\alpha > 0$. (We encourage you to figure out the answer from first principles, rather than just refer to the lecture.) **1 / 1 point**

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

Expand

Correct
Correct

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



The upper and lower neural networks have different input images, but have exactly the same parameters.

- False
- True

Expand

Correct
Yes it is true, parameters are shared among these two networks.

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

Expand

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

- (•) True
- ( ) False

Expand

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 neural style transfer the content loss J_{cont} is computed as: **1 / 1 point**

$$J_{cont}(G, C) = \left|\left|a^{[l](C)} - a^{[l](G)}\right|\right|^2$$

Where $a^{[l](k)}$ is the activation of the l-th layer of a ConvNet trained for classification. We choose l to be a very high value to use compared to the more abstract activation of each image. True/False?

- ( ) True
- (•) False

Expand

Correct
Correct. We don't use a very deep layer since this will only compare if the two images belong to the same category.

9. In neural style transfer, what is updated in each iteration of the optimization algorithm? **1 / 1 point**

- ( ) The regularization parameters
- ( ) The neural network parameters
- ( ) The pixel values of the content image C
- (•) The pixel values of the generated image G

Expand

Correct
Yes, neural style transfer is different from many of the algorithms you've seen up to now, because it doesn't learn any parameters; instead it learns directly the pixels of an image.

10. You are working with 3D data. You are building a network layer whose input volume has size 32x32x32x16 (this volume has 16 channels), and applies convolutions with 32 filters of dimension 3x3x3x16 (no padding, stride 1). What is the resulting output volume? **1 / 1 point**

- (•) 30x30x30x32
- ( ) 30x30x30x16
- ( ) Undefined: This convolution step is impossible and cannot be performed because the dimensions specified don't match up.

Expand

Correct

Correct, you have used the formula $\lfloor \frac{n^{[l-1]} - f + 2 \times p}{s} \rfloor + 1 = n^{[l]}$ over the three first dimensions of the input data.