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

Expand

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 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.
- Because we might have only one example of the person we want to verify.

Expand

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 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.
- You take several pictures of the same person to train $d(\text{img}_1, \text{img}_2)$ using the triplet loss.
- 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 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. 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.

- [ ] A the anchor image is a hyperparameter of the Siamese network.
- [ ] α is a trainable parameter of the Siamese network.
- [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.

- [x] $f(A)$ represents the encoding of the Anchor.

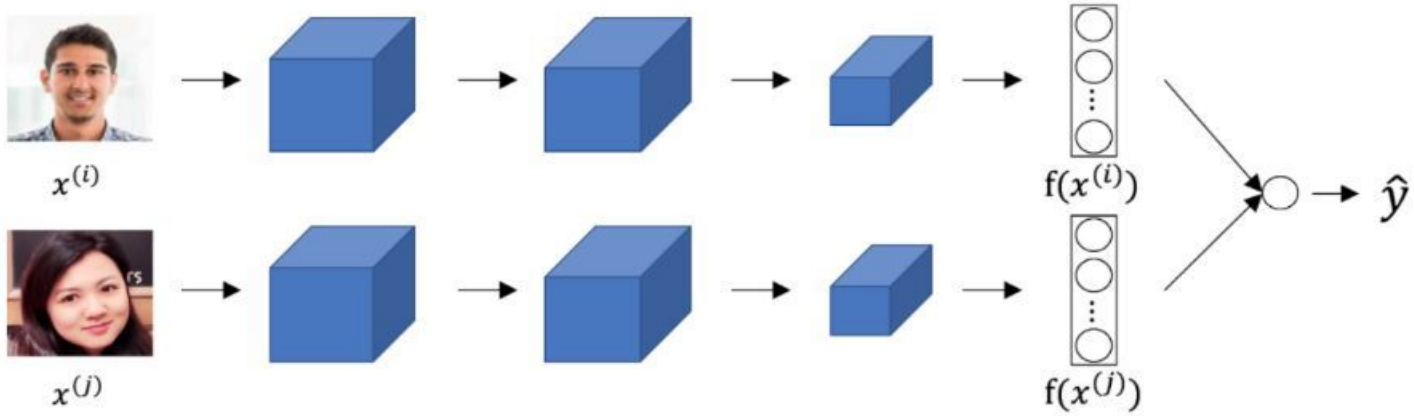
Correct

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

Expand

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

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

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

Expand

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. 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

- False
- True

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 $32 \times 32 \times 32 \times 3$, 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?

1 / 1 point

- (●) $29 \times 29 \times 29 \times 16$.
- ( ) $29 \times 29 \times 29 \times 3$.
- ( ) $31 \times 31 \times 31 \times 16$.
- ( ) $29 \times 29 \times 29 \times 13$.

Expand

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.