Congratulations! You passed!

Grade received 90% **Latest Submission Grade** 90% **To pass** 80% or higher

Go to 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. 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. Which of the following do you agree with? **0 / 1 point**

- [x] It will be more efficient to learn a function $d(\text{img}_1, \text{img}_2)$ for this task.

Correct
Correct. Since this is a one-shot learning task this function will allow us to compare two images to verify identity.

- [ ] This can't be considered a one-shot learning task since there might be many members in the workgroup.
- [ ] It is best to build a convolutional neural network with a softmax output with as many outputs as members of the group.
- [ ] This can be considered a one-shot learning task.

Expand

Incorrect
You didn't select all the correct 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.
- (•) 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. Triplet loss: **1 / 1 point**

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

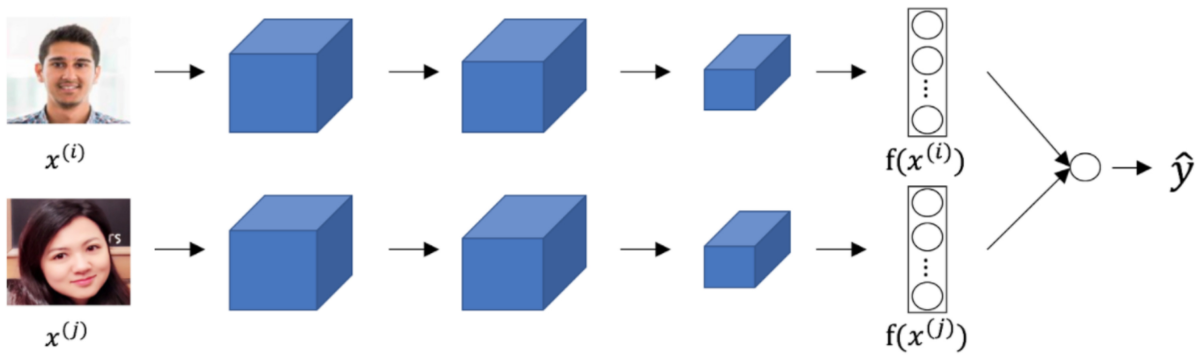
is larger in which of the following cases?

- 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 $A=N$.

Expand

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

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. 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 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**

- [ ] T that calculates the triplet loss between S, G, and C.
- [ ] J_{corr} that compares C and S.
- [x] $J_{content}$ 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.

- [x] 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.

Expand

Correct
Great, you got all the right answers.

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**

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

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.