Congratulations! You passed!

Grade received 100% | Latest Submission Grade 100% | 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? **1 / 1 point**

- [ ] 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.
- [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.

- [x] This can be considered a one-shot learning task.

Correct
Correct. Since we might have only one example of the person we want to recognize.

Loading [MathJax]/jax/output/CommonHTML/jax.js

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**

- ( ) 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 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 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**

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

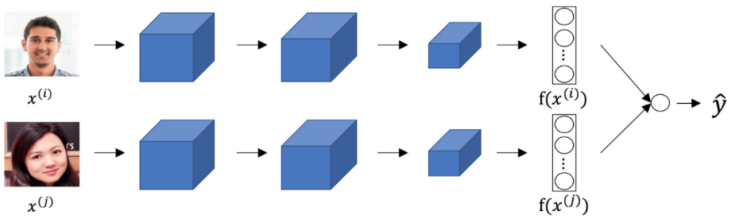
is larger in which of the following cases?

- When the encoding of A is closer to the encoding of N than to the encoding of P.
- When $A=P$ and $A=N$.
- When the encoding of A is closer to the encoding of P than to the encoding of 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?

- True
- False

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 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, we define style as: **1 / 1 point**

- ( ) The correlation between the activation of the content image C and the style image S.
- ( ) The correlation between the generated image G and the style image S.
- ( ) $\left\|a^{[l](S)} - a^{[l](G)}\right\|^2$ the distance between the activation of the style image and the content image.
- (●) The correlation between activations across channels of an image.

Expand

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

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

- ( ) The neural network parameters
- (●) The pixel values of the generated image

G

G

Loading [MathJax]/jax/output/CommonHTML/jax.js

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**

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

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.