Try again once you are ready

Grade received 70% | Latest Submission Grade 70% | To pass 80% or higher

Try again

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**

- [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.

- [ ] 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

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**

- ( ) You shouldn't use persons outside the workgroup you are interested in because that might create a high variance in your model.
- ( ) 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.

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?

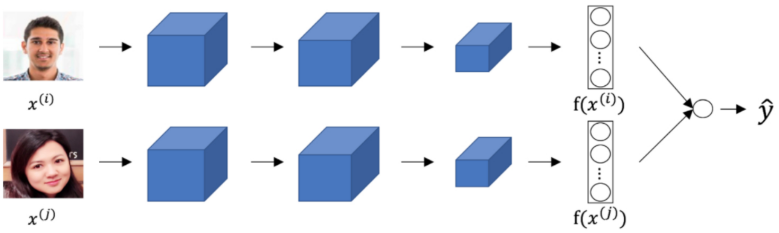
- When the encoding of A is closer to the encoding of P than to the encoding of N.
- When $A=P$ and $A=N$.
- When the encoding of A is closer to the encoding of N than to the encoding of P.

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

Expand

Correct
Correct. In this case $$\left | f(A) - f(P) \right |^2 - \left | f(A) - f(N) \right |^2$$ is positive thus the triplet loss gives a positive value larger than $$\alpha$$.

5. Consider the following Siamese network architecture: **0 / 1 point**



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

- True
- False

Expand

Incorrect
If these networks don't have the same parameters then the encoding of the two faces are not consistent with each other.

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. 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: **0 / 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

Incorrect
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 pixel values of the generated image G
- ( ) The regularization parameters
- ( ) The pixel values of the content image

  C
- ( ) The neural network parameters

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. 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? **0 / 1 point**

- ( ) $31 \times 31 \times 31 \times 16$.
- ( ) $64 \times 64 \times 64 \times 3$.
- (●) $61 \times 61 \times 61 \times 14$.
- ( ) $31 \times 31 \times 31 \times 3$.

Expand

Incorrect
No, use 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.