Special Applications: Face Recognition & Neural Style Transfer

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. Why is the face verification problem considered a one-shot learning problem? Choose the best answer. **1 / 1 point**

- ( ) 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 might have only one example of the person we want to verify.
- ( ) Because we are trying to compare to one specific person only.
- ( ) Because we have only have to forward pass the image one time through our neural network for verification.

Correct
Correct. One-shot learning refers to the amount of data we have to solve a task.

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. 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.

- [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.

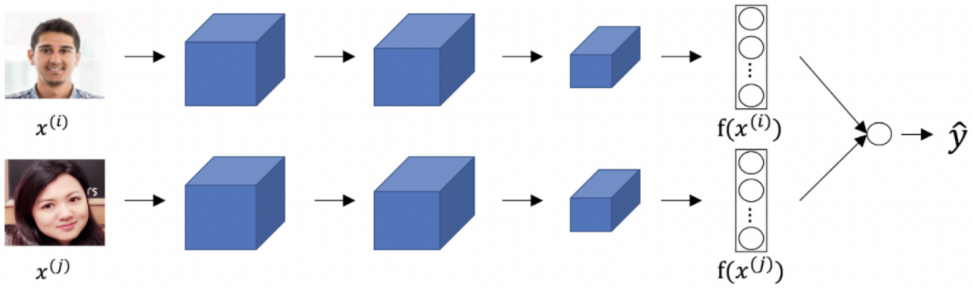
Correct
Correct. f represents the network that is in charge of creating the encoding of the images, and A represents the anchor image.

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

Expand

Correct
Great, you got all the right answers.

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



Which of the following do you agree with the most?

- The two neural networks depicted in the image have the same architecture, but they might have different parameters.
- The upper and lower neural networks depicted have exactly the same parameters, but the outputs are computed independently for each image.
- Although we depict two neural networks and two images, the two images are combined in a single volume and pass through a single neural network.
- This depicts two *different* neural networks with different architectures, although we use the same drawing.

Expand

Correct
Correct. Both neural networks share the same weights, and each image passes through the neural network in an independent manner.

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**

- ○ False
- ◉ True

Expand

Correct
Correct. Neurons that understand more complex shapes are more likely to be in deeper layers of a neural network.

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**

- ☐ J_{corr} that compares C and S.
- ☑ 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.

- ☑ $J_{content}$ 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.

- [ ] T that calculates the triplet loss between S, 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

- (•) True
- ( ) False

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, which of the following better express the gradients used? **1 / 1 point**

- $\frac{\partial J}{\partial W^{[l]}}$
- $\frac{\partial J}{\partial S}$
- $\frac{\partial J}{\partial G}$
- Neural style transfer doesn't use gradient descent since there are no trainable parameters.

Expand

Correct
Correct, we use the gradient of the cost function over the value of the pixels of the generated 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? **1 / 1 point**

- $31 \times 31 \times 31 \times 16$.
- $31 \times 31 \times 31 \times 3$.
- $64 \times 64 \times 64 \times 3$.
- $61 \times 61 \times 61 \times 14$.

Expand

Correct
Correct, we can use the formula $\lfloor \frac{n^{[l-1]} - f + 2 \times p}{s} \rfloor + 1 = n^{[l]}$ to the three first dimensions.