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

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

This should not be selected
This might be highly inefficient given that the members can change, requiring them to re-train the network each time.

Expand

Incorrect
You didn't select all the correct answers

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

- ( ) True

False

Expand

Correct
Correct, to train a network using the triplet loss you need several pictures of the same person.

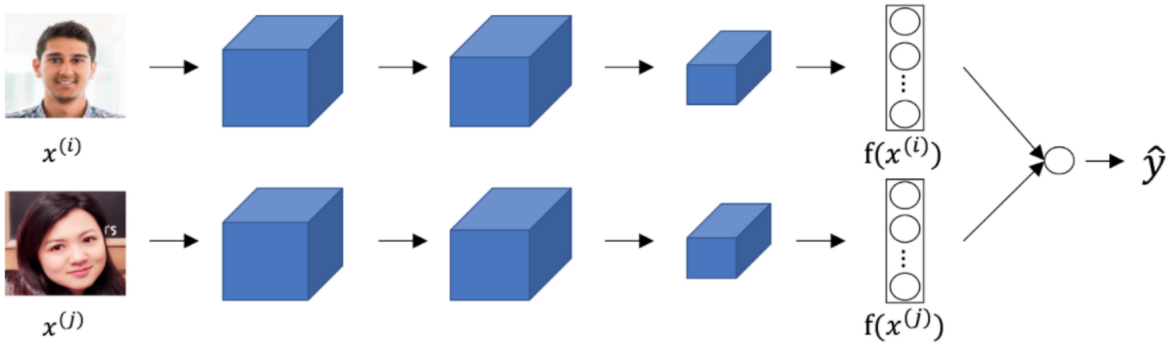
4. Which of the following is a correct definition of the triplet loss? Consider that $\alpha > 0$. (We encourage you to figure out the answer from first principles, rather than just refer to the lecture.) **1 / 1 point**

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

Expand

Correct
Correct

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



Which of the following do you agree with the most?

- This depicts two *different* neural networks with different architectures, although we use the same drawing.
- 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.

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 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**

- 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**

- Neural style transfer doesn't use gradient descent since there are no trainable parameters.

- $\frac{\partial J}{\partial S}$
- $\frac{\partial J}{\partial G}$
- ∂J

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. 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**

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

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.