1. Which of the following do you agree with?

- ○ Face verification requires K comparisons of a person's face.
- ◉ Face recognition requires K comparisons of a person's face.
- ○ Face recognition requires comparing pictures against one person's face.

Correct
Correct, in face recognition we compare the face of one person to K to classify the face as one of those K or not.

2. Why is the face verification problem considered a one-shot learning problem? Choose the best answer.

- ○ 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 have only have to forward pass the image one time through our neural network for verification.
- ○ Because we are trying to compare to one specific person only.

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

4. In the triplet loss:

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

Which of the following are true about the triplet loss? Choose all that apply.

- ☑ $f(A)$ represents the encoding of the Anchor.

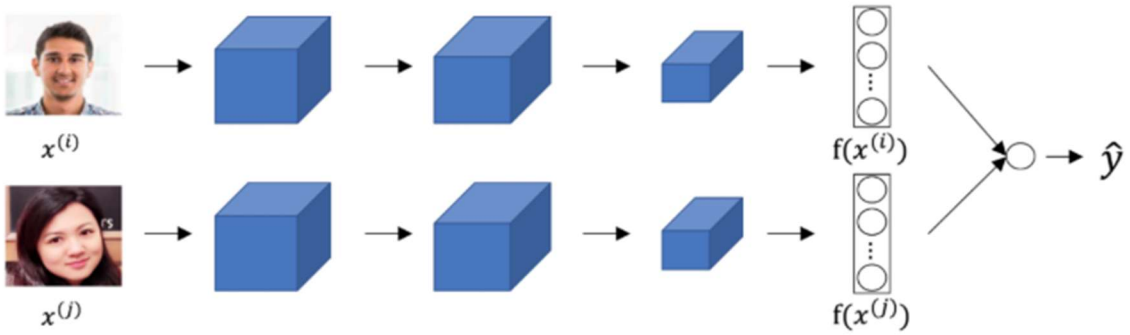
Correct
Correct. f represents the network that is in charge of creating the encoding of the images, and A represents the anchor image.

- ☑ 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.

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

5. Consider the following Siamese network architecture:



The upper and lower networks share parameters to have a consistent encoding for both images. True/False?

- (●) True
- ( ) False

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?

- ( ) True
- (●) False

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. 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.)

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

- [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.

8. In neural style transfer, we define style as:

- [x] The correlation between activations across channels of an image.
- [ ] 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.

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

9. In neural style transfer, which of the following better express the gradients used?

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

Correct
Correct, we use the gradient of the cost function over the value of the pixels of the generated image.