1. Face verification requires comparing a new picture against one person's face, whereas face recognition requires comparing a new picture against K persons' faces.

1 / 1 point

- (●) True
- ( ) False

Expand

Correct

Correct.

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

- [ ] 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 $d(\mathrm{img}_1, \mathrm{img}_2)$ 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.

- [ ] This can't be considered a one-shot learning task since there might be many members in the workgroup.
- [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.

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

- (selected) You take several pictures of the same person to train $d(\mathrm{img}_1, \mathrm{img}_2)$ 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.
- 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. 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.

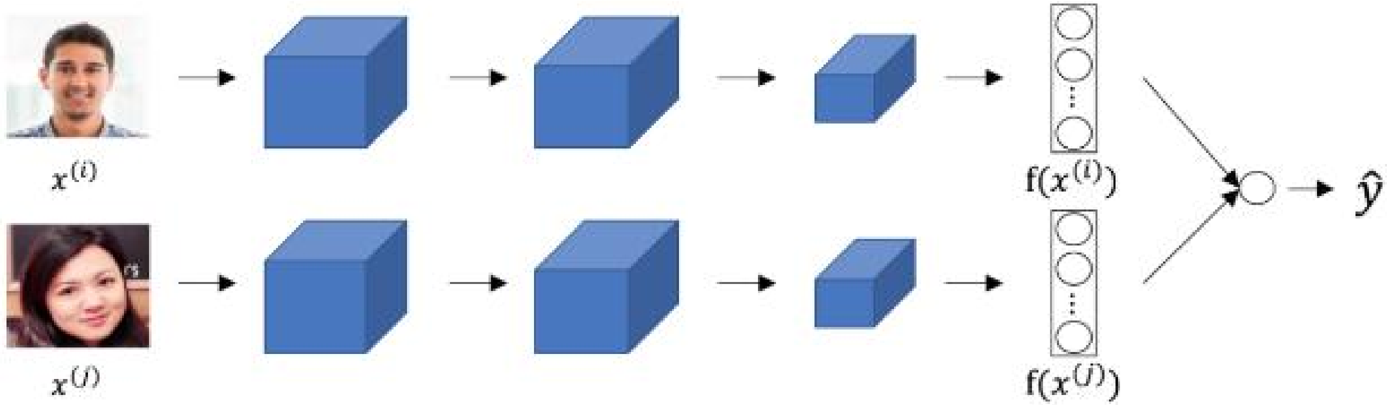
Expand

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.
- ○ This depicts two "different" neural networks with different architectures, although we use the same drawing.
- ○ 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. You train a ConvNet on a dataset with 100 different classes. You wonder if you can find a hidden unit which responds strongly to pictures of cats. (I.e., a neuron so that, of all the input/training images that strongly activate that neuron, the majority are cat pictures.) You are more likely to find this unit in layer 4 of the network than in layer 1.

1 / 1 point

- (●) True
- ( ) False

Expand

Correct

Yes, this neuron understands complex shapes (cat pictures) so it is more likely to be in a deeper layer than in the first layer.

7. In neural style transfer, we train the pixels of an image, and not the parameters of a network.

0 / 1 point

- ( ) False
- ( ) True

Expand

Incorrect

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:

1 / 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?

- (selected) False
- True

Expand

Correct

Correct. 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, we can't use gradient descent since there are no trainable parameters. True/False?

1 / 1 point

- (selected) False
- True

Expand

Correct

Correct. We use gradient descent on the cost function $J(G)$ and we update the pixel values of the generated image G.

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

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

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.