Congratulations! You passed!

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

Go to next item

1. Which of the following do you agree with? **1 / 1 point**

- ○ 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.

Expand

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 do we learn a function $d(img1, img2)$ for face verification? (Select all that apply.) **1 / 1 point**

- ☐ Given how few images we have per person, we need to apply transfer learning.
- ☑ This allows us to learn to recognize a new person given just a single image of that person.

  Correct
  Yes.
- ☑ We need to solve a one-shot learning problem.

  Correct
  This is true as explained in the lecture.
- ☐ This allows us to learn to predict a person's identity using a softmax output unit, where the number of classes equals the number of persons in the database plus 1 (for the final "not in database" class).

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.) **0 / 1 point**

- ◉ 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.
- ○ 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

Incorrect
To train using the triplet loss you need several pictures of the same person, so you don't do this only to increase the size of the dataset.

4. In the triplet loss: **0 / 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.

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

  ! **This should not be selected**
  α is a hyperparameter that prevents the network from send $f(\text{img})$ always to zeros.
- ☐ $f(A)$ represents the encoding of the Anchor.

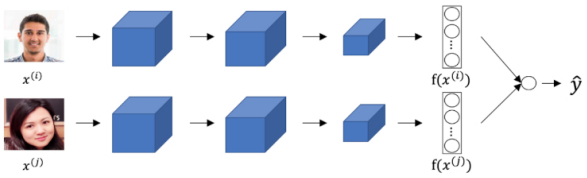
We want that $$\left\| f(A) - f(P) \right\|^2 < \left\| f(A) - f(N) \right\|^2$$ so the negative images are further away from the anchor than the positive images.

Correct
Correct. Being a positive image the encoding of $$P$$ should be close to the encoding of $$A$$.

Expand

Incorrect
You didn't select all the correct answers

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



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

- (•) True
- ( ) False

Expand

Correct
Correct. Part of the idea behind the Siamese network is to compare the encoding of the images, thus they must be consistent.

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

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

Expand

Correct
Great, you got all the right answers.

8. In neural style transfer the content loss J_{cont} is computed as: **1 / 1 point**

$$J_{cont}(G, C) = \left\| a^{[l](C)} - a^{[l](G)} \right\|^2$$

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?

- ( ) False

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, what is updated in each iteration of the optimization algorithm? **1 / 1 point**

The regularization parameters

The neural network parameters

The pixel values of the generated image G

The pixel values of the content image C

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. 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

30x30x30x16

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

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.