Quiz: Special Applications: Face Recognition & Neural Style Transfer

Congratulations! You passed!

Grade received	Latest Submission Grade	To pass
90%	90%	80% or higher

Go to next item

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

- [x] We need to solve a one-shot learning problem.

  ✓ **Correct**
  This is true as explained in the lecture.

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

  ✓ **Correct**
  Yes.

- [ ] 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.)

1 / 1 point

- 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.
- 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 take several pictures of the same person to train $d(\text{img}_1, \text{img}_2)$ using the triplet loss.
- You shouldn't use persons outside the workgroup you are interested in because that might create a high variance in your model.

Expand

Correct
Correct. To train using the triplet loss you need several pictures of the same person.

4. Triplet loss:

1 / 1 point

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

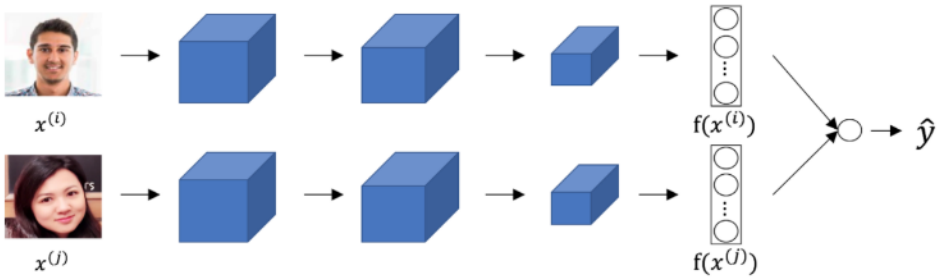
is larger in which of the following cases?

- When $A = P$ and $A = N$.
- When the encoding of A is closer to the encoding of P than to the encoding of N.
- When the encoding of A is closer to the encoding of N than to the encoding of P.

Expand

Correct
Correct. In this case $|f(A) - f(P)|^2 - |f(A) - f(N)|^2$ is positive thus the triplet loss gives a positive value larger than α.

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



The upper and lower neural networks have different input images, but have exactly the same parameters.

- False
- True

Expand

Correct
Yes it is true, parameters are shared among these two networks.

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? 1 / 1 point

- True
- False

Expand

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.) **1 / 1 point**

- [ ] 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.

Expand

Correct
Great, you got all the right answers.

8. In neural style transfer the content loss J_{cont} is computed as: **0 / 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
- (•) True

Expand

Incorrect
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 pixel values of the content image C
- The regularization parameters
- The neural network parameters
- (selected) The pixel values of the generated image G

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. 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**

- $31 \times 31 \times 31 \times 3$
- $64 \times 64 \times 64 \times 3$
- (selected) $31 \times 31 \times 31 \times 16$
- $61 \times 61 \times 61 \times 14$

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.