Congratulations! You passed!

Grade received 100% **Latest Submission Grade** 100% **To pass** 80% or higher

Go to next item

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

- (x) False
- ( ) True

Expand

Correct
Correct. This is the description of face verification, but not of face recognition.

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.
- [x] 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).
- [x] This allows us to learn to recognize a new person given just a single image of that person.

  Correct
  Yes.

Expand

Correct
Great, you got all the right 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
- (x) 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

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

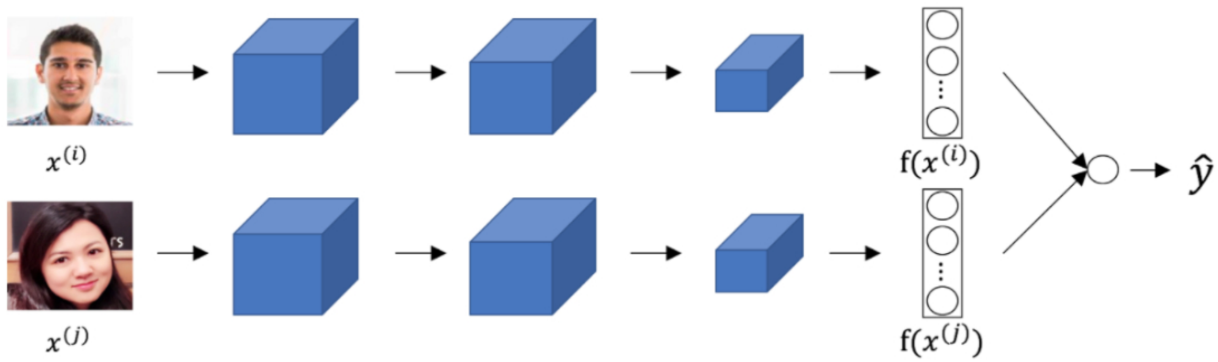
Expand

Correct

Correct

5. Consider the following Siamese network architecture:

1 / 1 point



Which of the following do you agree with the most?

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

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.

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. 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.
- ○ $\frac{\partial J}{\partial W^{[l]}}$
- ◉ $\frac{\partial J}{\partial G}$
- ○ $\frac{\partial J}{\partial S}$

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. The input "image" has size $32 \times 32 \times 32 \times 3$, if you apply a convolutional layer with 16 filters of size $4 \times 4 \times 4$, zero padding and stride 1. What is the size of the output volume?

1 / 1 point

- ○ $29 \times 29 \times 29 \times 3$.
- ○ $31 \times 31 \times 31 \times 16$.
- ◉ $29 \times 29 \times 29 \times 16$.
- ○ $29 \times 29 \times 29 \times 13$.

Expand

Correct

Correct, we can use the formula $\lfloor \frac{n^{[l-1]} - f + 2 \times p}{s} \rfloor + 1 = n^{[l]}$ on the three first dimensions.