

Your grade: 100%

Your latest: 100% • Your highest: 100% • To pass you need at least 80%. We keep your highest score.

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

- ☐ False
- ☒ True

Expand

Correct

Correct.

2. Why is the face verification problem considered a one-shot learning problem? Choose the best answer. 1 / 1 point

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

Expand

Correct

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

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

- ☐ 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.
- ☐ 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 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?

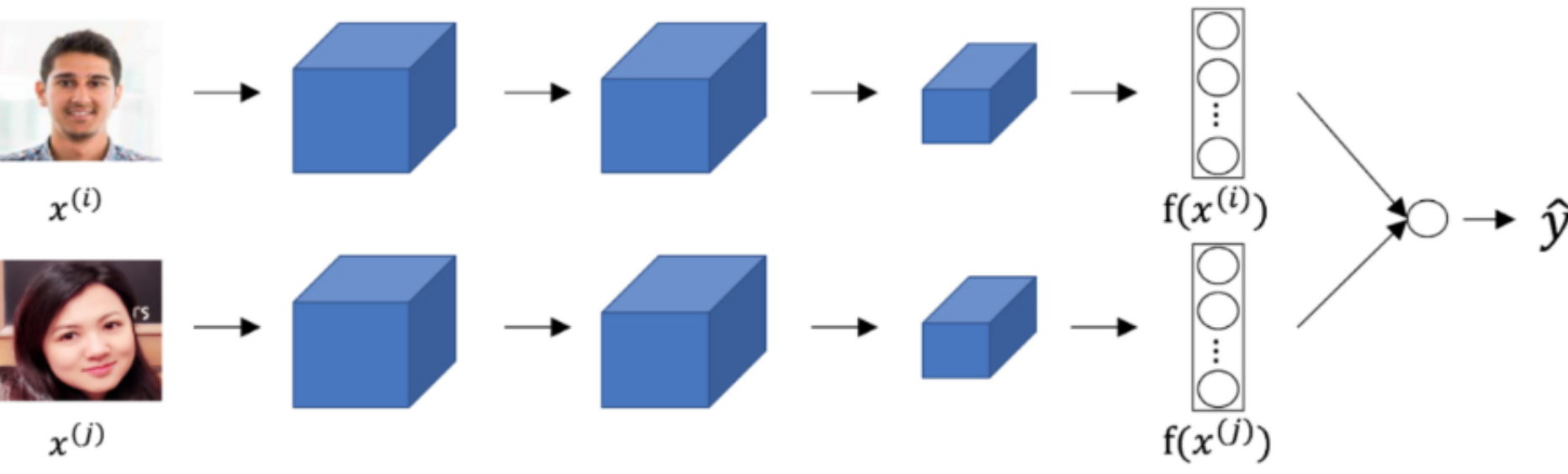
- ☒ When the encoding of A is closer to the encoding of N than to the encoding of P.
- ☐ When the encoding of A is closer to the encoding of P than to the encoding of N.
- ☐ When $A = P$ and $A = N$.

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



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.

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 neural style transfer, we define style as: 1 / 1 point

- ☐ $\|a^{[l](S)} - a^{[l](G)}\|^2$ the distance between the activation of the style image and the content image.
- ☒ 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 .

Expand

Correct

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

9. In neural style transfer, which of the following better express the gradients used? 1 / 1 point

- ☐ $\frac{\partial J}{\partial W^{[l]}}$
- ☐ $\frac{\partial J}{\partial S}$

parameters.

- ☒ $\frac{\partial J}{\partial w^2}$

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

- ☐ $29 \times 29 \times 29 \times 13$.
- ☒ $29 \times 29 \times 29 \times 16$.
- ☐ $31 \times 31 \times 31 \times 16$.
- ☐ $29 \times 29 \times 29 \times 3$.

Expand

Correct

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