Congratulations! You passed!

Grade received 80% Latest Submission Grade 80% To pass 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 is the face verification problem considered a one-shot learning problem? Choose the best answer. 1 / 1 point

- ( ) 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 are trying to compare to one specific person only.
- (•) 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

- ( ) 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 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.

Expand

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

4. In the triplet loss: 0 / 1 point

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

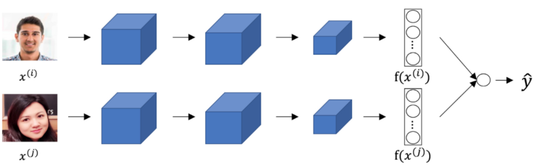
Which of the following are true about the triplet loss? Choose all that apply.

- [ ] A the anchor image is a hyperparameter of the Siamese network.
- [ ] 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.
- [ ] $f(A)$ represents the encoding of the Anchor.
- [ ] α is a trainable parameter of the Siamese network.

Expand

Incorrect
You didn't select all the correct 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.
- ( ) 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 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

- [ ] $\mathcal{T}$ that calculates the triplet loss between S, G, and C.
- [ ] J_{new} 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, we define style as: 1 / 1 point

- ( ) $\left\|a^{[l](S)} - a^{[l](G)}\right\|^2$ 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, what is updated in each iteration of the optimization algorithm? 1 / 1 point

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

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? 0 / 1 point

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

Expand

Incorrect
No, use 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.