

1. You are building a 3-class object classification and localization algorithm. The classes are: pedestrian (c=1), car (c=2), motorcycle (c=3). What would be the label for the following image? Recall $y=[pc,bx,by,bh,bw,c1,c2,c3]$

(0,0)



$y=[1, 0.3, 0.7, 0.3, 0.3, 0, 2, 0]$

2. Continuing from the previous problem, what should y be for the image below? Remember that “?” means “don’t care”, which means that the neural network loss function won’t care what the neural network gives for that component of the output. As before, $y=[pc,bx,by,bh,bw,c1,c2,c3]$.



$y=[0, ?, ?, ?, ?, ?, ?, ?]$

3. If you build a neural network that inputs a picture of a person’s face and outputs N landmarks on the face (assume the input image always contains exactly one face), how many output units will the network have?

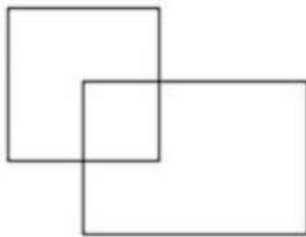
8개

4. When training one of the object detection systems described in lecture, you need a training set that contains many pictures of the object(s) you wish to detect.

However, bounding boxes do not need to be provided in the training set, since the algorithm can learn to detect the objects by itself. (True or False)

False

5. What is the IoU between these two boxes? The upper-left box is 2x2, and the lower-right box is 2x3. The overlapping region is 1x1.

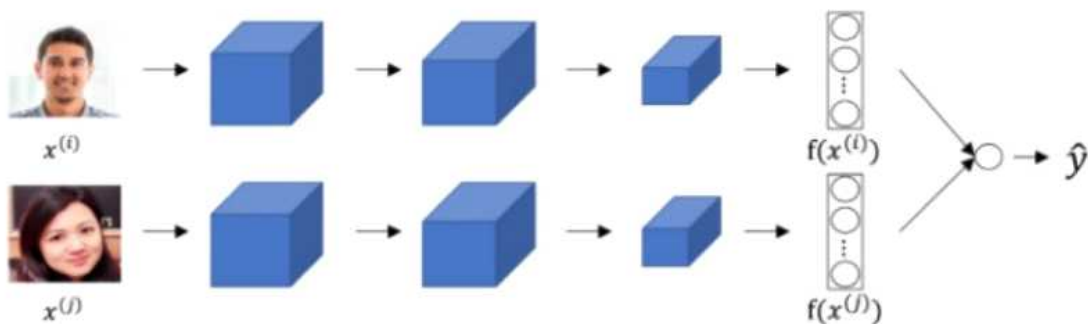


$$1/(4+6-1)=1/9$$

6. Why do we learn a function $d(\text{img1}, \text{img2})$ for face verification? - (주관식 작성)

두 사람과의 차이정도를 측정하여 어떤 값 이하이면 같은 사람이라고 판단하기 위해

7. Consider the following Siamese network architecture:



The upper and lower neural networks have different input images and the different parameters.

8. C4W3L04의 강의에서는 sliding window에 convolution을 적용하여 연산 속도를 빠르게 하는 방법을 설명하고 있습니다. 아래의 강의 슬라이드를 자신이 이해한대로 설명하는 글을 작성해주세요.

기존의 슬라이딩 윈도우 알고리즘은 합성곱 신경망을 따로따로 수행해야 하는데 이것을 합성곱 신경망으로 구현하면 반복적인 계산들을 공유할 수 있어 합성곱 신경망에 이미지를 한 번만 통과시키면, 물체의 위치를 인식할 수 있다.

Convolution implementation of sliding windows

