Congratulations! You passed!

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

Go to next item

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 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. Recall $y = [p_c, b_x, b_y, b_h, b_w, c_1, c_2, c_3]$.

1 / 1 point

https://www.pexels.com/es-es/foto/mujer-vestida-con-falda-azul-y-blanca-caminando-cerca-de-la-hierba-verde-durante-el-dia-144474/

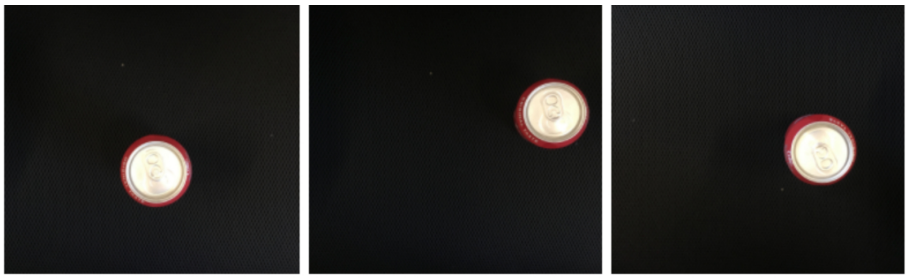
- (●) $y = [1, 0.66, 0.5, 0.75, 0.16, 1, 0, 0]$
- ( ) $$y = [1, ?, ?, ?, ?, 1, ?, ?]$$
- ( ) $$y = [1, 0.66, 0.5, 0.16, 0.75, 1, 0, 0]$$
- ( ) Loading [MathJax]/jax/output/CommonHTML/jax.js, 0, 0]$$

Expand

Correct
Correct. $p_c = 1$ since there is a pedestrian in the picture. We can see that b_x, b_y as percentages of the image are approximately correct as well b_h, b_w, and the value of $c_1 = 1$ for a pedestrian.

2. You are working on a factory automation task. Your system will see a can of soft-drink coming down a conveyor belt, and you want it to take a picture and decide whether (i) there is a soft-drink can in the image, and if so (ii) its bounding box. Since the soft-drink can is round, the bounding box is always square, and the soft-drink can always appear the same size in the image. There is at most one soft-drink can in each image. Here are some typical images in your training set:

1 / 1 point



The most adequate output for a network to do the required task is $y = [p_c, b_x, b_y, b_h, b_w, c_1]$. (Which of the following do you agree with the most?)

- False, we don't need
  b_h
  b_h,
  b_w
  ne size.
- True. since this is a localization problem.

Expand

Correct
Correct. With the position b_x, b_y we can completely characterize the position of the object if it is present. We should use only one additional logistic unit to indicate if the object is present or not.

3. When building a neural network that inputs a picture of a person's face and outputs N landmarks on the face (assume that the input image contains exactly one face), which is true about $\hat{y}^{(i)}$? **1 / 1 point**

- $\hat{y}^{(i)}$
  stores the probability that a landmark is in a given position over the face.
- $$\hat{y}^{(i)}$$ has shape (1, 2N)
- $$\hat{y}^{(i)}$$ has shape (2N, 1)

Expand

Correct
Correct. Since we have two coordinates (x,y) for each landmark we have N of them.

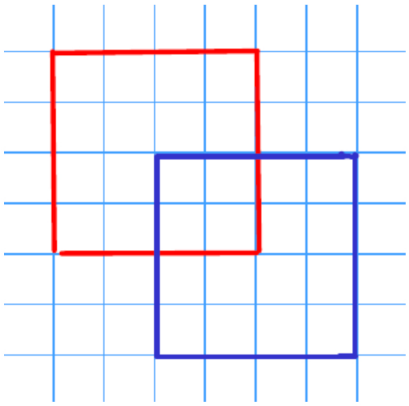
4. When training one of the object detection systems described in the lectures, each image must have zero or exactly one bounding box. True/False? **1 / 1 point**

- True
- False

Expand

Correct
Correct. In a single image, there might be more than only one instance of the object we are trying to localize, so it must have several bounding boxes.

5. What is the IoU between the red box and the blue box in the following figure? Assume that all the squares have the same measurements. **1 / 1 point**



- $\frac{1}{8}$
- $\frac{1}{7}$

$\frac{1}{7}$

Expand

Correct
Correct. IoU is calculated as the quotient of the area of the intersection (4) over the area of the union (28).

6. Suppose you run non-max suppression on the predicted boxes nelow. The parameters you use for non-max suppression are that boxes with probability ≤ 0.7 are discarded, and the IoU threshold for deciding if two boxes overlap is 0.5. 1 / 1 point

After non-max suppression, only three boxes remain. True/False?

- True
- False

Expand

Correct
Correct. After eliminating the boxes with a score less than 0.7 only three boxes remain, and they don't intersect. Thus three boxes are left.

7. Suppose you are using YOLO on a 19x19 grid, on a detection problem with 20 classes, and with 5 anchor boxes. During training, for each image you will need to construct an output volume y as the target value for the neural network; this corresponds to the last layer of the neural network. (y may include some "?", or "don't cares"). What is the dimension of this output volume?

0 / 1 point

- (●) 19x19x(5x20)
- ( ) 19x19x(5x25)
- ( ) 19x19x(20x25)
- ( ) 19x19x(25x20)

Expand

Incorrect
How is a box defined for a problem with 20 classes?

8. We are trying to build a system that assigns a value of 1 to each pixel that is part of a tumor from a medical image taken from a patient.

1 / 1 point

This is a problem of localization? True/False

- ( ) True
- (●) False

Expand

Correct
Correct. This is a problem of semantic segmentation since we need to classify each pixel from the image.

9. Using the concept of Transpose Convolution, fill in the values of **X**, **Y** and **Z** below.

1 / 1 point

(*padding = 1, stride = 2*)

Input: 2x2

1	3
2	4

Filter: 3x3

1	0	1
0	0	0
1	0	1

Result: 6x6

	0	0	0	0	
	0	X	0	7	

	0	0	0	Y	
	0	Z	0	4	

- (●) X = 10, Y = 0, Z = 6
- ( ) X = 3, Y = 0, Z = 4
- ( ) X = 10, Y = 0, Z = 0
- ( ) X = 4, Y = 3, Z = 2

Expand

Correct
Correct.

10. When using the U-Net architecture with an input $h \times w \times c$, where c denotes the number of channels, the output will always have the shape $h \times w \times c$. True/False? **1 / 1 point**

- (●) False
- ( ) True

Expand

Correct
Correct. The output of the U-Net architecture can be $h \times w \times k$ where k is the number of classes. The number of channels doesn't have to match between input and output.