Congratulations! You passed!

Grade received 100% | Latest Submission Grade 100% | 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/

- $y = [1, 0.66, 0.5, 0.16, 0.75, 1, 0, 0]$
- $y = [1, 0.66, 0.5, 0.75, 0.16, 0, 0, 0]$
- $y = [1, ?, ?, ?, ?, 1, ?, ?]$
- $y = [1, 0.66, 0.5, 0.75, 0.16, 1, 0, 0]$

Loading [MathJax]/jax/output/CommonHTML/jax.js

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

What are the most appropriate (lowest number of) output units for your neural network?

- Logistic unit (for classifying if there is a soft-drink can in the image)

Logistic unit, b_x, b_y, b_h, b_w

Logistic unit,

b_x

Loading [MathJax]/jax/output/CommonHTML/jax.js

Expand

Correct
Correct!

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), we need two coordinates for each landmark, thus we need 2N output units. True/False? **1 / 1 point**

True

False

Expand

Correct
Correct. Recall that each landmark is a specific position in the face's image, thus we need to specify two coordinates for each landmark.

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**

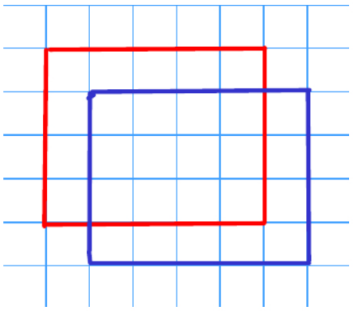
True

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**



$\frac{1}{2}$

$\frac{3}{7}$

Loading [MathJax]/jax/output/CommonHTML/jax.js

Expand

Correct
Correct. IoU is calculated as the quotient of the area of the intersection (16) over the area of the union (28).

6. Suppose you run non-max suppression on the predicted boxes below. The parameters you use for non-max suppression are that boxes with probability ≤ 0.4 are discarded, and the IoU threshold for deciding if two boxes overlap is 0.5. How many boxes will remain after non-max suppression? 1 / 1 point

- 7
- 5
- 6
- 4
- 3

Expand

Correct
Correct!

7. Which of the following do you agree with about the use of anchor boxes in YOLO? Check all that apply. 1 / 1 point

- [ ] Each object is assigned to any anchor box that contains that object's midpoint.
- [x] Each object is assigned to the grid cell that contains that object's midpoint.

Correct
Correct. This is the way we choose the corresponding cell.

- [x] Each object is assigned to an anchor box with the highest IoU inside the assigned cell.

Correct
Correct. This is the way we choose the corresponding anchor box.

- [ ] They prevent the bounding box from suffering from drifting.

Expand

Correct
Great, you got all the right answers.

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

taken from a patient.

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	2
3	4

Filter: 3x3

1	0	-1
1	0	-1
1	0	-1

Result: 6x6

	0	1	0	-2	
	0	**X**	0	**Y**	
	0	1	0	**Z**	
	0	1	0	-4	

- ◉ X = 2, Y = -6, Z = -4
- ○ X = 2, Y = -6, Z = 4
- ○ X = 2, Y = 6, Z = 4
- ○ X = -2, Y = -6, Z = -4

Expand

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 $h \times w$. True/False? **1 / 1 point**

- ○ True
- ◉ False

Expand

Correct

Correct. The output of the U-Net architecture can be $h \times w \times k$ where k is the number of classes.