Congratulations! You passed!

Grade received 80%

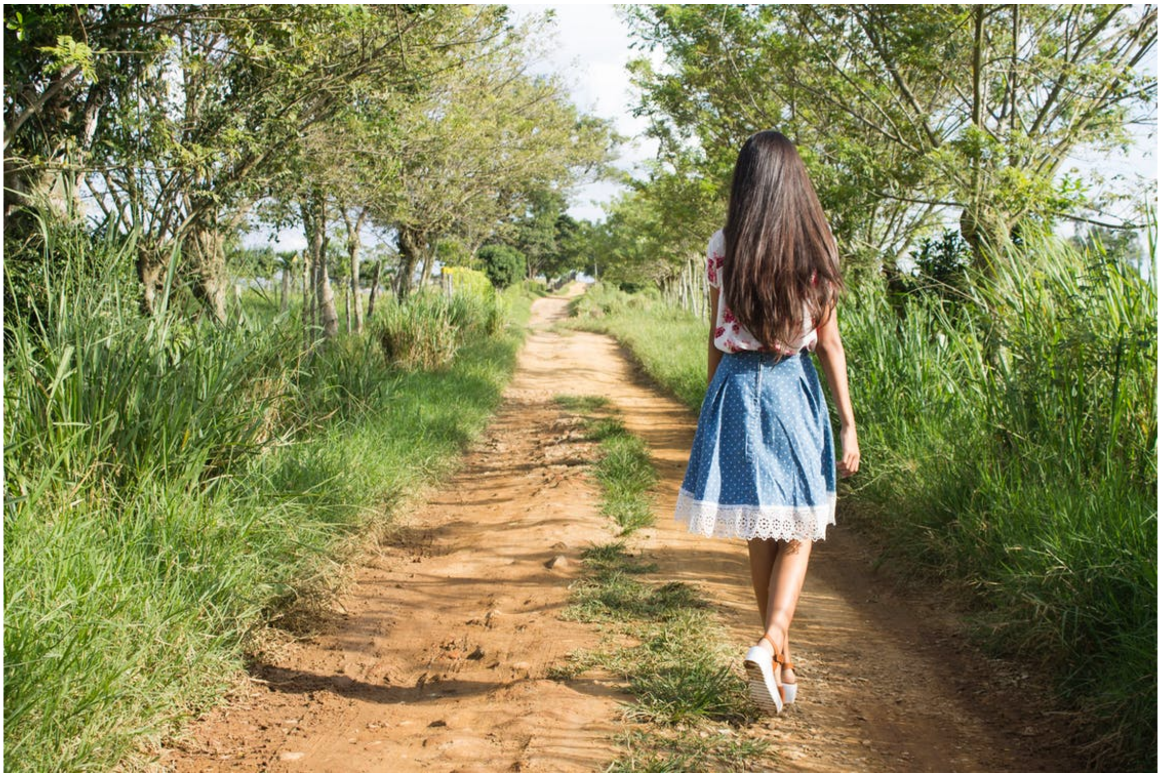
Latest Submission Grade 80%

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/

○

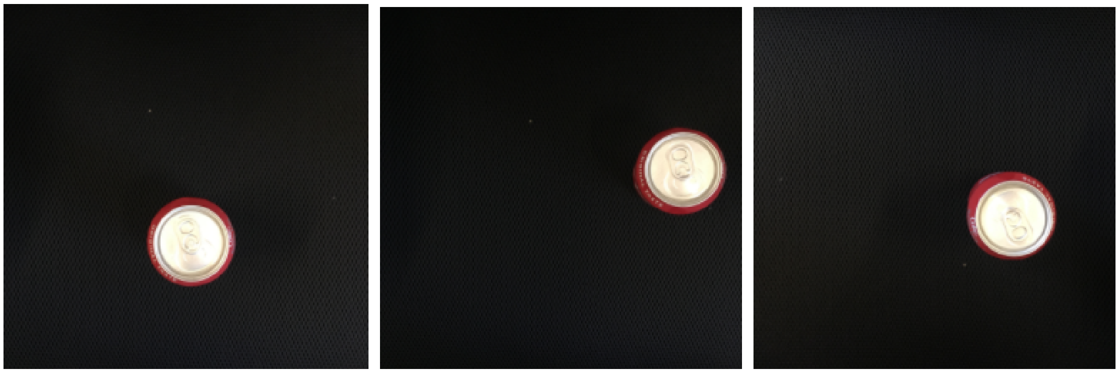
Expand

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?

Expand

✓ **Correct**
Correct!

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?

1 / 1 point

Expand

✓ **Correct**
Correct

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?

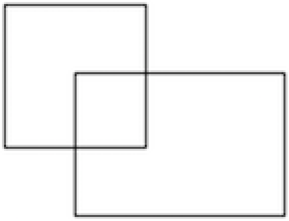
1 / 1 point

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 these two boxes? The upper-left box is 2x2, and the lower-right box is 2x3. The overlapping region is 1x1. **1 / 1 point**

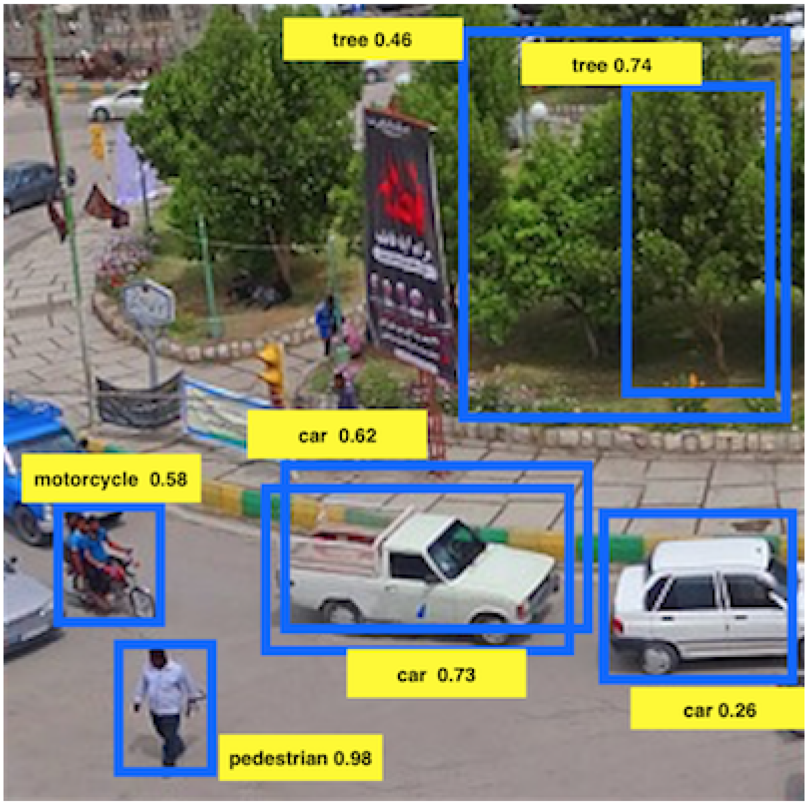


Expand

✓ **Correct**

Correct. The left box's area is 4 while the right box 's is 6. Their intersection's area is 1. So their union's area is 4 + 6 - 1 = 9 which leads to an intersection over union of 1/9.

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. **0 / 1 point**



Notice that there are three bounding boxes for cars. After running non-max suppression, only the bounding box of the car with 0.73 is kept from the three bounding boxes for cars. True/False? Choose the best answer.

Expand

ⓧ **Incorrect**

Incorrect. One of the bounding boxes for cars is eliminated because it has a lower score and an IoU higher than 0.5.

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

1 / 1 point

Expand

Correct

Great, you got all the right answers.

8. What is Semantic Segmentation?

1 / 1 point

Expand

✓ **Correct**

9. Using the concept of Transpose Convolution, fill in the values of **X**, **Y** and **Z** below. **0 / 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	

⤢ **Expand**

⊗ **Incorrect**

To revise the concepts watch the lecture .

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

⤢ **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.