Congratulations! You passed!

Grade received 90% To pass 80% or higher

Go to next item

Detection Algorithms

Latest Submission Grade 90%

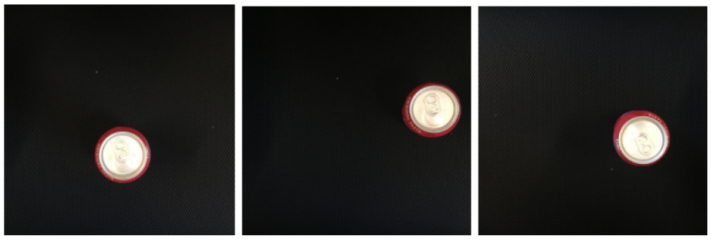
1. You are building a 3-class object classification and localization algorithm. The classes are: pedestrian (c=1), car (c=2), motorcycle (c=3). What should y be for the image below? Remember that "?" means "don't care", which means that the neural network loss function won't care what the neural network 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

- $y = [0, ?, ?, ?, ?, ?, ?, ?]$
- $y = [?, ?, ?, ?, ?, ?, ?, ?]$
- $y = [1, ?, ?, ?, ?, 0, 0, 0]$
- $y = [1, ?, ?, ?, ?, ?, ?, ?]$

Expand

Correct
Correct.

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, b_x, b_y, b_h, b_w
- Logistic unit, b_x and b_y
- Logistic unit, b_x, b_y, b_h (since $b_w = b_h$)
- Logistic unit (for classifying if there is a soft-drink can in the image)

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), 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 (2N, 1)
- $\hat{y}^{(i)}$ has shape (N, 1)
- $\hat{y}^{(i)}$ has shape (1, 2N)

Expand

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

4. You are working to create an object detection system, like the ones described in the lectures, to locate cats in a room. To have more data with which to train, you search on the internet and find a large number of cat photos. **0 / 1 point**

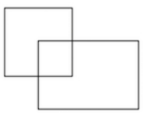
Which of the following is true about the system?

- We should add the internet images (without the presence of bounding boxes in them) to the train set.
- We can't add the internet images unless they have bounding boxes.
- We should use the internet images in the dev and test set since we don't have bounding boxes.
- We can't use internet images because it changes the distribution of the dataset.

Expand

Incorrect
It is beneficial to use extra images in the training set, and the change in distribution doesn't affect much as long as we use the images in the training set and not in the dev and test sets. In this case, the problem is that the internet images don't have 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**

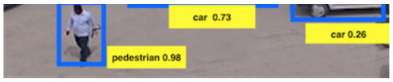


- $\frac{1}{10}$
- None of the above
- $\frac{1}{6}$
- $\frac{1}{9}$

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. How many boxes will remain after non-max suppression? **1 / 1 point**



- 4
- 6
- 3
- 5
- 7

Expand

Correct
Correct!

7. If we use anchor boxes in YOLO we no longer need the coordinates of the bounding box b_x, b_y, b_h, b_w since they are given by the cell position of the grid and the anchor box selection. True/False? **1 / 1 point**

- True
- False

Expand

Correct
Correct. We use the grid and anchor boxes to improve the capabilities of the algorithm to localize and detect objects, for example, two different objects that intersect, but we still use the bounding box coordinates.

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

- False
- True

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 = 10, Y = 0, Z = 0
- X = 4, Y = 3, Z = 2
- X = 3, Y = 0, Z = 4

Expand

Correct
Correct.

10. Suppose your input to a U-Net architecture is h x w x 3, where 3 denotes your number of channels (RGB). What will be the dimension of your output ? **1 / 1 point**

- h x w x n, where n = number of input channels
- h x w x n, where n = number of output classes
- h x w x n, where n = number of of output channels
- h x w x n, where n = number of filters used in the algorithm

Expand

Correct