축하합니다! 통과하셨습니다!

받은 학점 100% 최신 제출물 학점 90% 통과 점수: 80% 이상

다음 항목으로 이동

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 / 1점

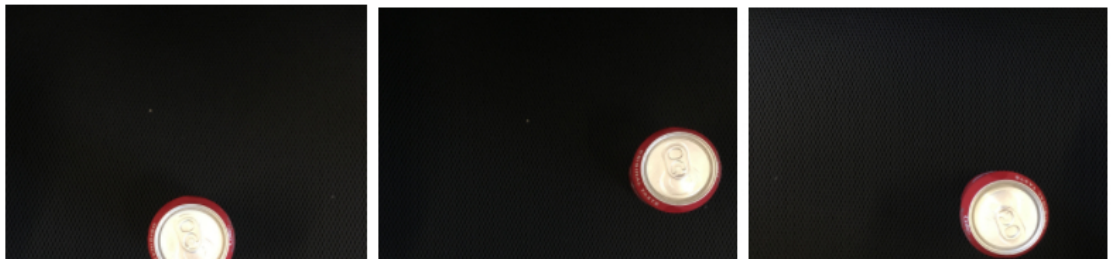
- $y = [0, ?, ?, ?, ?, ?, ?, ?]$
- $y = [?, ?, ?, ?, ?, ?, ?, ?]$
- $y = [1, ?, ?, ?, ?, ?, ?, ?]$
- $y = [1, ?, ?, ?, ?, 0, 0, 0]$

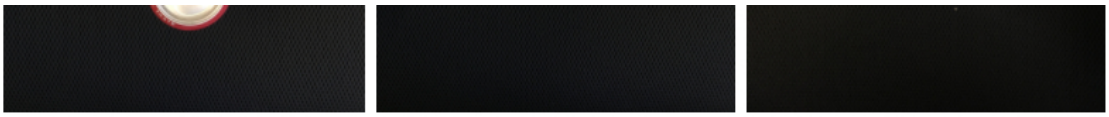
더 보기

맞습니다
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're some typical images in your training set:

1 / 1점





To solve this task it is necessary to divide the task into two: 1. Construct a system to detect if a can is present or not. 2. Construct a system that calculates the bounding box of the can when present. Which one of the following do you agree with the most?

- The two-step system is always a better option compared to an end-to-end solution.
- We can't solve the task as an image classification with a localization problem since all the bounding boxes have the same dimensions.
- An end-to-end solution is always superior to a two-step system.
- We can approach the task as an image classification with a localization problem.

더 보기

맞습니다
Correct. We can use a network to combine the two tasks similar to that described in the lectures.

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 / 1점**

- 3N
- 2N
- N^2
- N

더 보기

맞습니다
Correct

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. **1 / 1점**

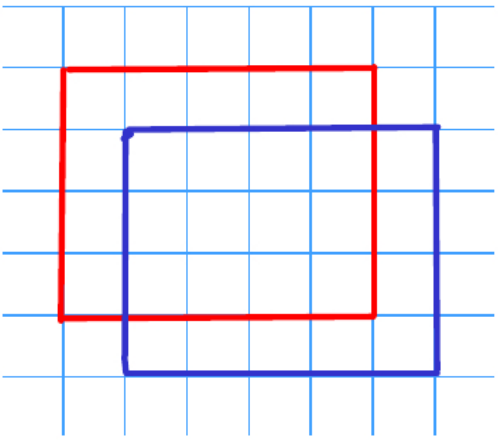
Which of the following is true about the system?

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

더 보기

맞습니다
Correct. As this is a localization model, we also need the coordinates of the bounding boxes, not just the images.

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 / 1점**



- $\frac{2}{5}$
- $\frac{4}{5}$
- $\frac{3}{7}$
- $\frac{1}{2}$

더 보기

맞습니다
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. **0 / 1점**

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.

- True. Only one of the boxes for a car is kept because the score of 0.73 is higher, and also because it has a larger bounding box.
- False. Two bounding boxes corresponding to cars are left since their IoU is zero.
- True. The non-maximum suppression eliminates the bounding boxes with scores lower than the ones of the maximum.
- False. All the cars are eliminated since there is a pedestrian with a higher score of 0.98.

더 보기

✗ 틀립니다
The algorithm doesn't act in that way; it also considers the IoU of the boxes.

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 / 1점**

- True
- False

더 보기

✓ 맞습니다
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. What is Semantic Segmentation? **1 / 1점**

- Locating an object in an image belonging to a certain class by drawing a bounding box around it.
- Locating objects in an image by predicting each pixel as to which class it belongs to.
- Locating objects in an image belonging to different classes by drawing bounding boxes around them.

더 보기

✓ 맞습니다

9. Using the concept of Transpose Convolution, fill in the values of **X**, **Y** and **Z** below. **1 / 1점**

(*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 = 0
- ○ X = 4, Y = 3, Z = 2
- ◉ X = 10, Y = 0, Z = 6
- ○ X = 3, Y = 0, Z = 4

↙↗ 더 보기

✓ 맞습니다
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 / 1점**

- h x w x n where n = number of output classes
- h x w x n where n = number of input channels
- 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

더 보기

맞습니다