

Project 5 Questions

Instructions

- 3 questions.
- Write code where appropriate; feel free to include images or equations.
- **We do NOT expect you to fill up each page with your answer.** Some answers will only be a few sentences long, and that is okay.

Questions

Q1:

- (a) Explain these common terms in machine learning in your words:
 - (i) Bias
 - (ii) Variance
- (b) Define these terms in the context of evaluating a classifier:
 - (i) Overfitting
 - (ii) Underfitting
- (c) Does bias and variance have any impact on overfitting and underfitting. Can you describe a brief real-world scenario where you can find/observe this?

Please answer overleaf.

A1: Your answer here.

- (a)
 - (i) Bias represents the ability of the algorithm to fit
 - (ii) Variance represents the robust of the algorithm
- (b)
 - (i) A classifier is particularly good for certain special cases and generalises poorly
 - (ii) A classifier is not good or can not do its job
- (c) Yes. High bias will usually led to Underfitting, and High Variance will usually led to Overfitting

Q2: Suppose you had to test the selective search algorithm on an image from pedestrian detection dataset (for example: an image taken from traffic camera as shown below), do you think that selective search algorithm will suggest bounding boxes over person(s).



- (a) If yes, what is the justification in your words for this successful behavior of the algorithm? If not, then can you think in which cases does it fail? Can you suggest one way to improve or modify the approach?

A2: Your answer here.

- (a) No, The area of the person is not visible enough compared to the surrounding area, so it may be divided into the background so that each person does not have a bounding box
- (b) We can see that people do not occupy a large area in the picture and the bounding box formed should be relatively small . So we can lower the size threshold for areas to adjust the people area size, but it may produce many false meaningless areas(false positive)

Q3: If you were to apply selective search algorithm to detect interesting regions for skin diseases, how do you think the algorithm would have performed?

- (a) Which among the four similarity (color, texture, size and shape) do you think contributes more to automatic detection of interesting regions in this case?
- (b) Combining what you understood so far, what do you think “objectness” means and do you agree that selective search algorithm inherently finds regions of interest based on “objectness”?

Please answer overleaf.

A3: Your answer here.

- (a) This depends on the different skin conditions and they may be flagged if the external manifestations of the skin condition are very different from healthy skin
- (b) color and texture
- (c) My understanding is that the properties of the object itself, which can be objectness, such as colour size and shape, are different for different objects, so we can do selective search based on this, so I think it is right

Feedback? (Optional)

Please help us make the course better. If you have any feedback for this assignment, we'd love to hear it!

Thanks to the efforts of all the tutors, I have learnt a lot from this project!