

# CAP 6412 Advanced Computer Vision - Project I

Due on 02/28/2016, 11:59pm

## Option 1: R-CNN

This project is all about object detection, especially the PASCAL VOC challenge<sup>1</sup> and R-CNN.

[R-CNN] Girshick, Ross, Jeff Donahue, Trevor Darrell, and Jagannath Malik. "Rich feature hierarchies for accurate object detection and semantic segmentation." In CVPR 2014.

You are supposed to fully understand the approach of R-CNN and be able to implement it from scratch. You can also build your code upon Ross' (cf. <https://github.com/rbgirshick/rcnn>).

You are expected to learn about how to evaluate the object detection results of an approach using the PASCAL VOC defined metrics and development toolkits.

If you need to run your code on GPU, you are encouraged to sign up the AWS Educate program to receive a free \$100 credit to access the Amazon's EC2 g2.2xlarge GPU instances running Ubuntu.

### Data

Please test the object detectors on the validation set of PASCAL VOC 2012<sup>2</sup>, object detection track.

You will earn some extra credits by training some components of the R-CNN pipeline. This should be conducted on the training set of PASCAL VOC 2012. To save some computation time, you can focus on one or a few classes, in contrast to using all the 20 classes, for the training.

### What to turn in

1. A .zip archive of all your source code excluding all dependencies (e.g., Ross' code, third-party tools, etc.)
2. A detailed report covering at least the following.
  - Detection average precision on VOC 2012 validation. Please report the results in the form of Table 1 in the [R-CNN] paper.

<sup>1</sup><http://host.robots.ox.ac.uk/pascal/VOC/>

<sup>2</sup><http://host.robots.ox.ac.uk/pascal/VOC/voc2012/index.html>

- Detection average precision on VOC 2012 validation of the results before regression. Please report the results in the form of Table 1 in the [R-CNN] paper.
- Detection average precision on VOC 2012 validation using 1000 (in contrast to about 2000) detection proposals.
- Intermediate output of more than three images randomly chosen from VOC 2012 validation: all the detection proposals on the randomly selected images, a few warped proposals, top 20–50 proposals ranked by the SVM scores, remaining proposals after non-maximum suppression, refined proposals after regression, and also the groundtruth bounding boxes for comparison. All the results should be shown as bounding boxes on the images except the few warped proposals.
- **Bonus:** You are encouraged to add new ingredients to R-CNN. Describe them in the report and refer to the source files in the .zip archive.

See the slides of 02/02 for grading policy.

## Option 2: Your own project

You could propose a computer vision topic and work on it as the course project. The candidate topics include but are not limited to the Papers and Resources on the course homepage<sup>3</sup>.

If you choose this option, please draft your proposal using a few slides and then set up a meeting to discuss the plausibility of the project. The slides should at least cover the following.

- Motivation of the project.
- Problem statement.
- Approach: outline and details.
- Data to be used in the project.
- How to evaluate the approach.
- How to report the results.
- Related work and/or baselines (optional).

<sup>3</sup><http://www.cs.ucf.edu/~bgong/CAP6412.html>