

Object Detection

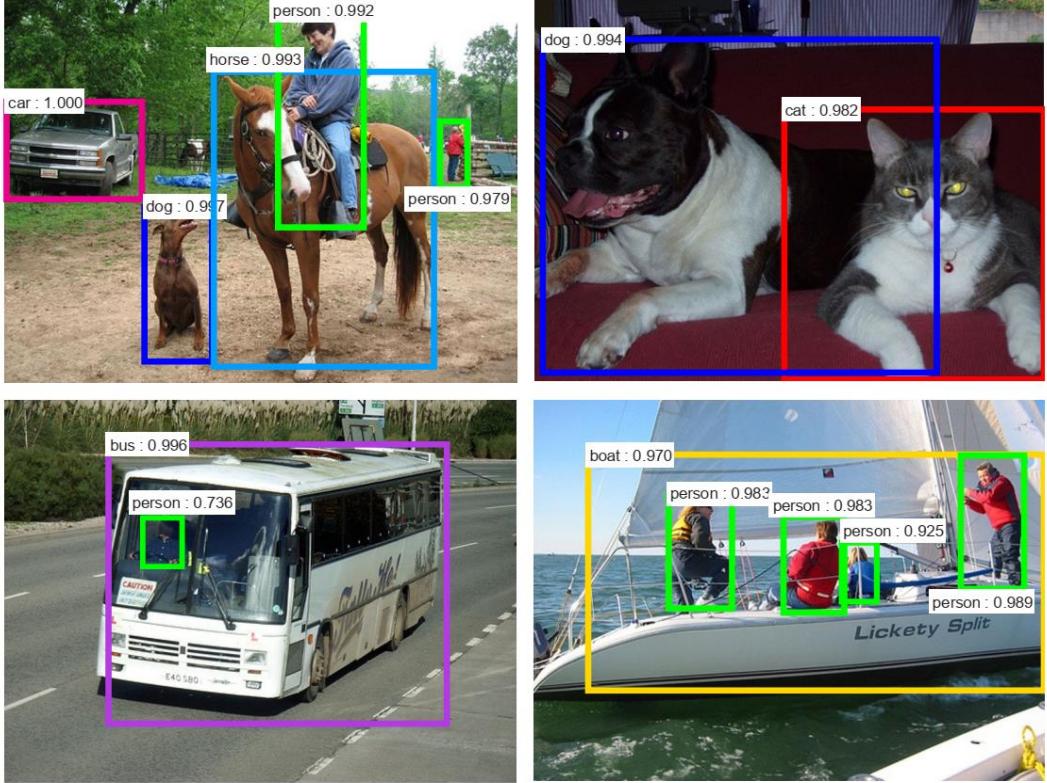


Fig.1 Object detection task: samples

■ Project description:

Given an input image, find the positions of multiple objects in the image as well as their class label.

1. **Dataset construction:** Replace one of 20 classes of objects in PASCAL VOC 2007 with another class of objects (you can replace a class from other datasets, or build a new object class data by yourself) to form a dataset with 20 classes that is slightly different from PASCAL VOC 2007 and used for training/testing. If you would like to try more, you can build a larger dataset (but you cannot directly use the existing data set, you need to replace some classes of objects.)
2. **Method implementation:** You can adapt from the existing methods in Tensorflow/Pytorch, etc., but you must retrain the network structure with the new dataset mentioned above. Moreover, any improvement for the algorithm is highly expected. Please highlight the improvement part in your report, if any.
3. **Experiment:** Refer to the papers about object detection, and list the mAP on the above dataset in the table, including the results of each object class.
4. **Visualization:** Show the detection result with a box, and annotate the object class with text on the box as shown in Fig.1.

■ Some object detection approaches:

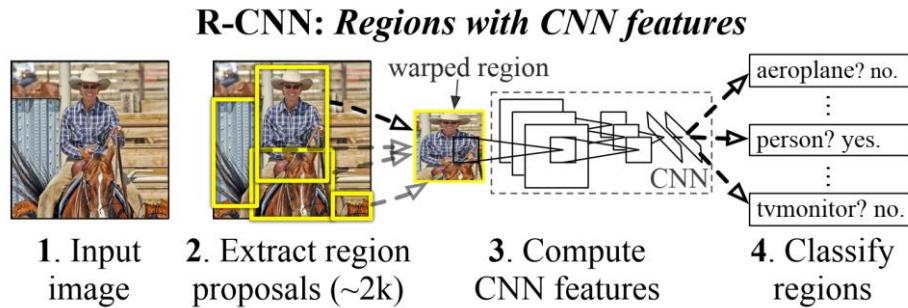


Fig.2 Overview of object detection with R-CNN

1. R-CNN
2. SPPnet
3. Fast R-CNN
4. Faster R-CNN
5. YOLO / YOLO9000 / YOLOv3 / YOLOv4
6. SSDs

■ Object detection datasets:

Dataset	Images	Boxes	Categories	Boxes/img	Fully Annotated
Pascal VOC	11.5k	27k	20	2.4	Yes
ImageNet All	477k	534k	200	1.1	Yes
ImageNet Dense	80k	186k	200	2.3	Yes
COCO	123k	896k	80	7.3	Yes
OpenImages	1,515k	14,815k	600	9.8	Partial
Objects365	638k	10,101k	365	15.8	Yes

PASCAL VOC: <https://pjreddie.com/projects/pascal-voc-dataset-mirror/>

COCO: <http://cocodataset.org/>

旷世 Objects365: <https://www.objects365.org>

References

- [1] R. Girshick, J. Donahue, T. Darrell, and J. Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In CVPR, 2014. (**R-CNN**)
- [2] K. He, X. Zhang, S. Ren, and J. Sun. Spatial pyramid pooling in deep convolutional networks for visual recognition. In ECCV, 2014. (**SPPnet**)
- [3] Ross Girshick, Fast R-CNN, ICCV 2015. (**Fast R-CNN**)
- [4] S Ren, K He, R Girshick, J Sun. Faster R-CNN, NIPS 2015. (**Faster R-CNN**)
- [5] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once:

- Unified, real-time object detection. CVPR 2016. (**YOLO**)
- [6] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C Berg. SSD: Single shot multibox detector. ECCV 2016. (**SSD**)
- [7] Shuai Shao, Zeming Li, Tianyuan Zhang, Chao Peng, Gang Yu, Xiangyu Zhang, Jing Li, Jian Sun, **Objects365**: A Large-Scale, High-Quality Dataset for Object Detection, ICCV 2019. (**Object365 Dataset**)