

Advanced Robotic Perception: Lecture 7

Lecture Notes

Victor Risager

October 3, 2023

1 Understanding neural networks

The features are getting more holistic the further down the network we go. High level features combine shapes of the previous features.

Datasets:

- PASCAL VOC
- MS COCO
- LVIS

1.1 Sliding window

Instead of taking every window in the image, you start by identifying different regions, or blobs in the image that you can do CNN on. This is called R-CNN. Originally, run the entire Conv-Net and whenever we have classified the image, then we can run the last part of the network for each class/bounding box.

RPN - Learnable region proposals

Single stage detectors are now better than multi stage. This was up to 2017.

2 Object detectors evaluation

Comparing prediction bounding box to a ground truth bounding box, we can use the union of the two bounding boxes.

mAP = Mean average precision

you can do learnable convolution to upsample the image.