Detection

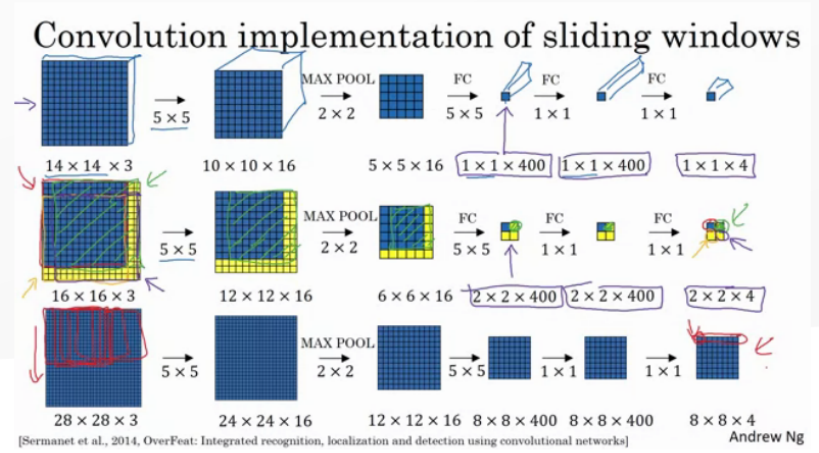
Friday, August 3, 2018 11:18 AM

Classification with localization
Have output 4 more outputs -> bounding box
Depends on whether contains an object.

Landmark detection
Find some landmark of the face -> is that a face and coordinate of face
Add a brunch of outputs

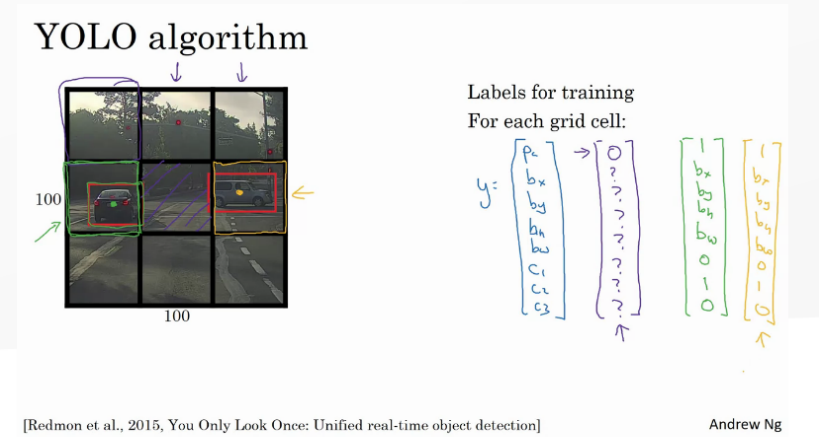
Object detection(sliding window)
Choose a sliding window through the picture -> predict that region
The choose different size of sliding window
Much more expensive and slow

Convolutional implementation of sliding window
Turning FC layer into Conv layer



Position of bounding box maybe inaccurate.

Bounding box predictions
YOLO split into grids. For each grid cell, have a prediction output

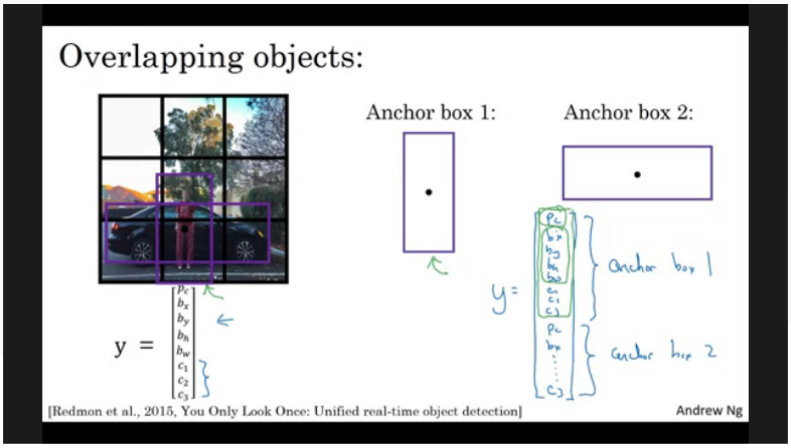


1. Like image classification and localization
2. Convolutional implementation (fast, real-time)

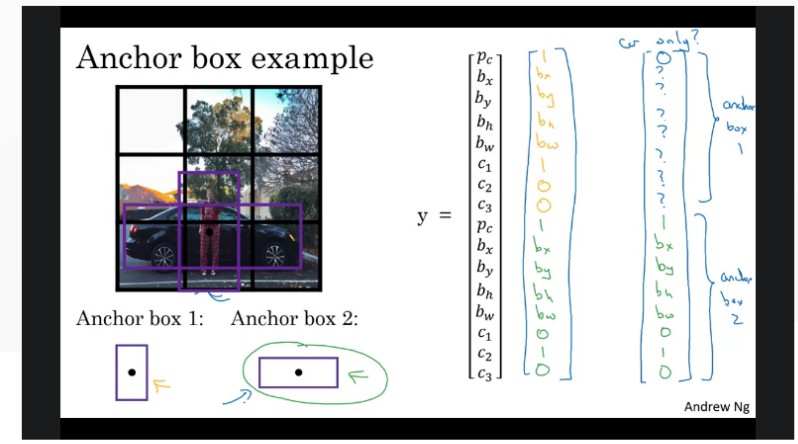
Intersection Over Union(IoU)
Compute union over bounding boxes (size of intersection / size of union)
Correct if IoU >= 0.5 (hyper-parameter)
Measure of the overlap between two bounding boxes

Non-max suppression (NMS)
Clean up the multiple detections
Suppress the higher IoU bounding boxes
Output the maximum probability box
First step, discard all boxes that have a lower probability
While loop(any remaining boxes), pick the box with highest probability, then output the prediction
Finally, Discard higher IoU boxes.

Anchor boxes
Predefine several size of anchor boxes



With two anchor boxes, each object in training image is assigned to grid cell that contains object's midpoint and anchor box for the grid cell with highest IoU.



If objects are more than anchor boxes number, this algorithm cannot handle.
Or several object share save anchor box shape.

YOLO algorithm – put together

Region Proposal