

UET-AI-Robot INT3409-21

Nguyễn Duy Võ Anh
Nguyễn Trọng Dương

Challenge 1

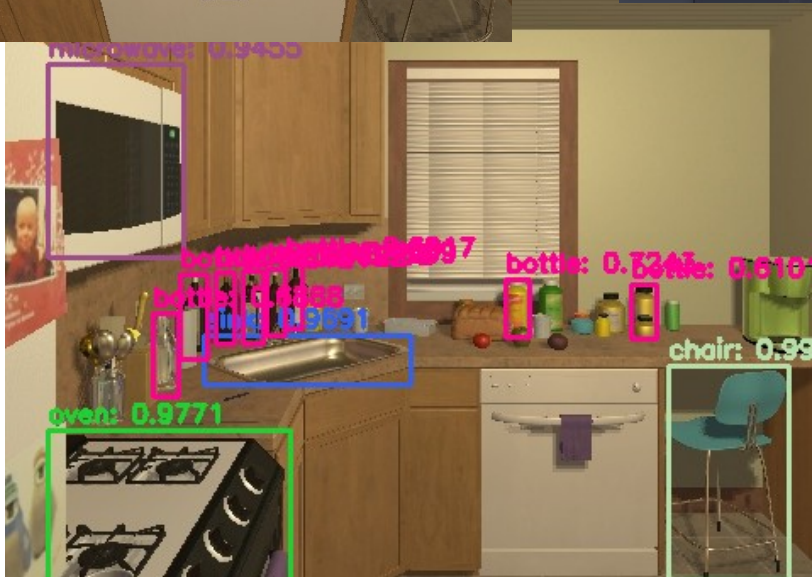
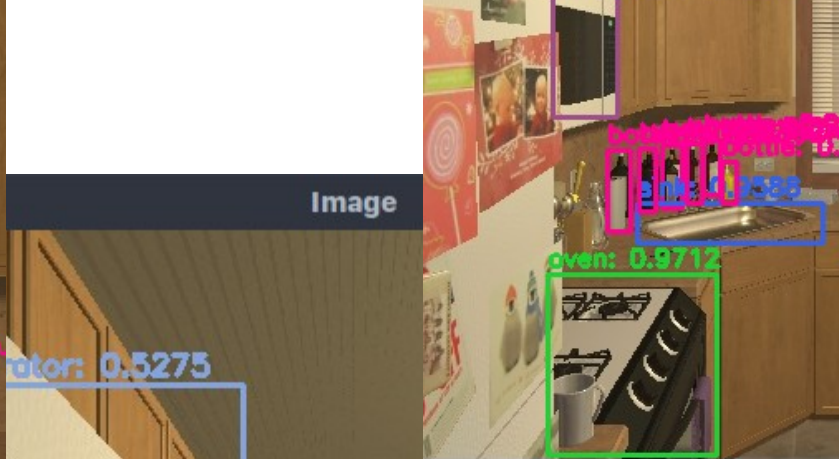
Object detection and Classification on First View Cameras

- Why do we choose this challenge?

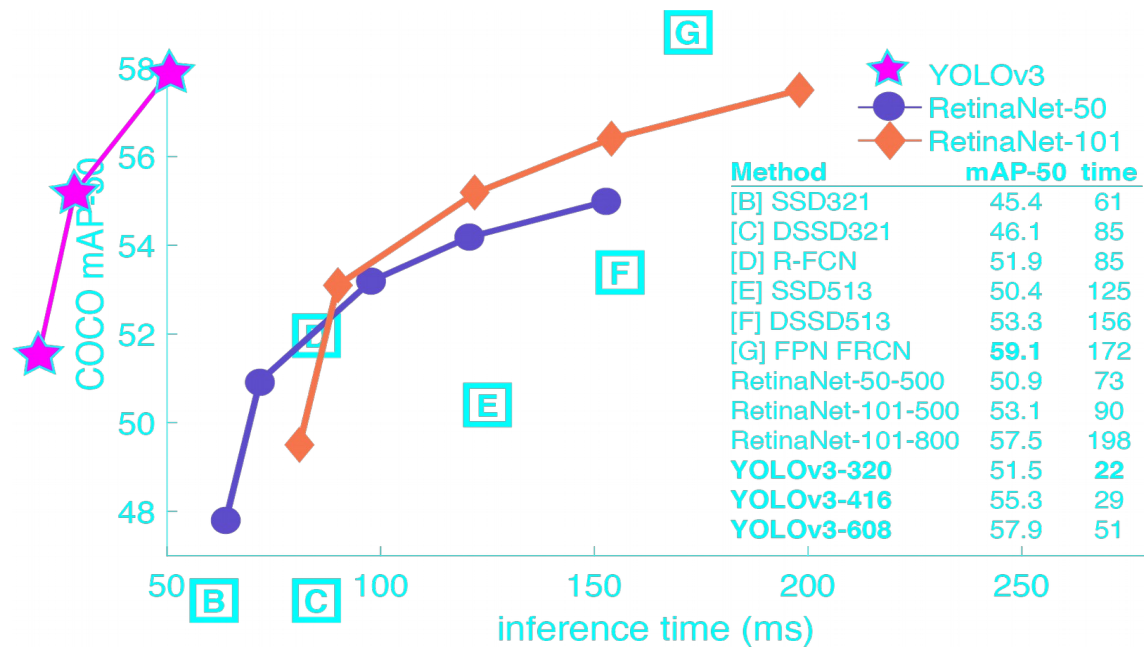
Object detection is important for Computer vision and we found this problem is fascinating

- Why yolo?

Yolo is fast so it can detect the objects in real-time.

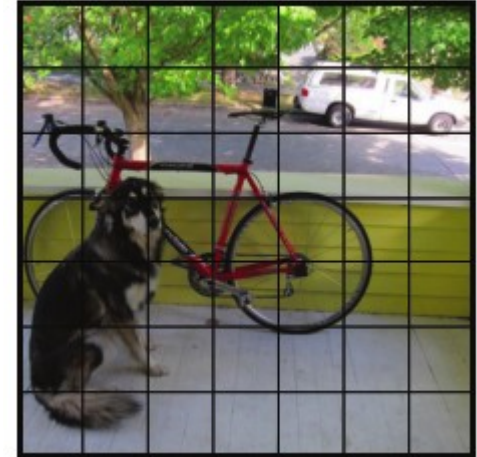
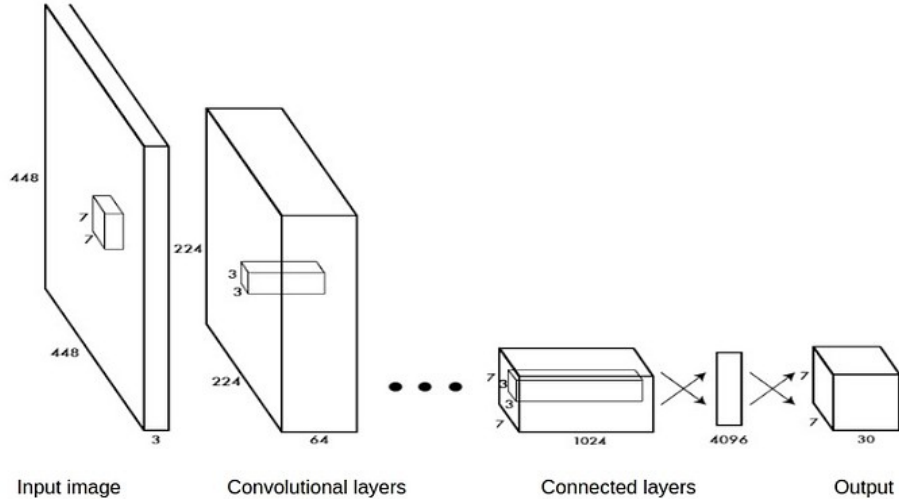


Yolo



Yolo

YOLO is a deep net that combine with convolutional layers and connected layers

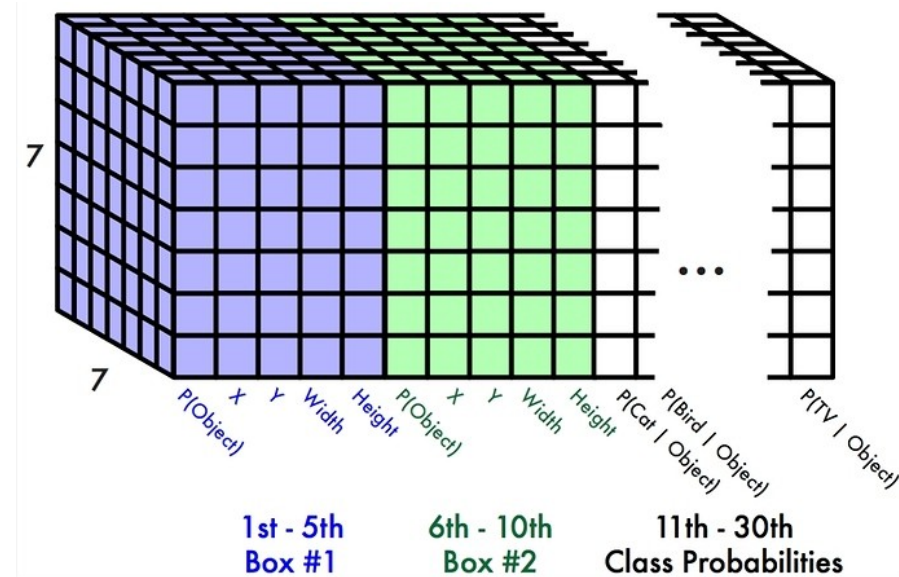


YOLO divided a image into 7x7 grids (grid_size=7x7):

Yolo

YOLO will predict for each grid that include the object, predict the object size and find the center point.

We can understand that each cell will have 30 parameters, the first parameter is to determine the cell containing 1 object, parameter 2,3,4,5 times are x_center , y_center , width, height of the box 1. Similar. Self parameter 6,7,8,9,10 is box 2. The 11th parameter is to determine the object in the cell is object1 (in 20_objects need to be identified). Similarly, parameter 12 is to determine that the object in the cell is object2 ... until parameter 30 is to determine that the object in the cell is object20.



Idea

For each step that the agent move,
we will call yolo to process
that frame with pre-trained
weight



Requirements

- OS: Mac OS X 10.9+, Ubuntu 14.04+
- Graphics Card: DX9 (shader model 3.0) or DX11 with feature level 9.3 capabilities.
- CPU: SSE2 instruction set support.
- Python 2.7 or Python 3.5+
- Linux users: X server with GLX module enabled
- CUDA

Installation and Getting Started

- Installing AI2-THOR using pip:
\$ pip install ai2thor
- Git clone the repository:
\$ git clone <https://github.com/voanh01/ai2thorxyolov3.git>
- Setup:
\$ cd ai2thor
\$ sudo python yolo_ai2thor.py

Actions

- “w” - Move the agent forward by gridSize.
- “a” - Move the agent left by gridSize (without changing view direction).
- “s” - Move the agent backward by gridSize (without changing view direction).
- “d” - Move the agent right by gridSize (without changing view direction).
- “left arrow” - Rotate the agent by 90 degrees to the left of its current facing.
- “right arrow” - Rotate the agent by 90 degrees to the right of its current facing.
- “up arrow” - Angle the agent’s view up in 30 degree increments.
- “down arrow” - Angle the agent’s view down in 30 degree increments.