OBJECT DETECTION AND CLASSIFICATION ON FIRST VIEW CAMERA USING YOLOV3

Robot – INT3409 21: Man Minh Thao - 16020282
Chu Minh Tien - 16020287
Nguyen Duy Quang - 16020272

1. Introduction:

- AI2-THOR: An Interactive 3D Environment for Visual AI.
  - Functions: 3D views, customizable, photorealistic, physics (forces, friction,..), object interaction and multi-agent.
  - Requirement:
    - 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
  - Installation:

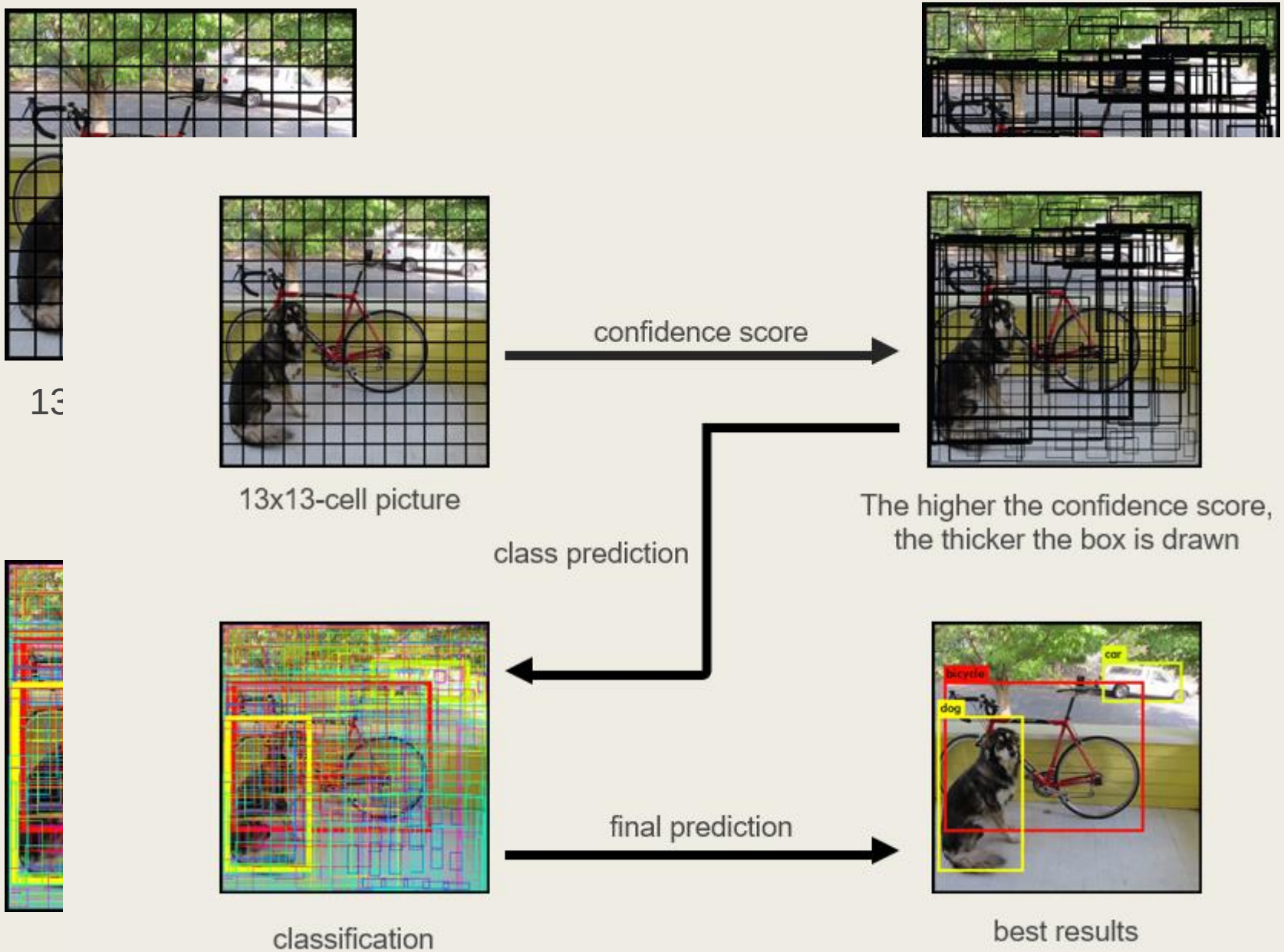
```
$ pip install ai2thor
```

1. Introduction:

- YOLO: Real-time Object Dectection system
  - Version: YOLOv3, which is improved in training and increase performance, including: multi-scale predictions, a better backbone classifier, and more.
  - Installation:

```
$ git clone https://github.com/pjreddie/darknet
$ cd darknet
$ make
```

  - How YOLO works:


13x13-cell picture
confidence score
The higher the confidence score,
the thicker the box is drawn
class prediction
classification
final prediction
best results
bicycle
car
dog

2. Project:

- Install and run: Run in Ubuntu 14.04+ using terminal

```
$ git clone https://github.com/thaomm/ai2thor-objdetect-yolov3.git
$ cd ai2thor-objdetect-yolov3/roboto
$ python3 robot.py
```

- Method:
  - Using YOLO to detect and COCO dataset to train the object detector
  - Moving the agent's view with keyboard, capturing its frames after every move, using YOLO to process the object detector and drawing a bounding box around each object in every frame.

2. Project:

- Controller:
  - “w” / move front
  - “s” / move back
  - “a” / move left
  - “d” / move right
  - “left arrow” / rotate left
  - “right arrow” / rotate right
  - “up arrow” / look up
  - “down arrow” / look down

3. Demo

Source Image

Final result