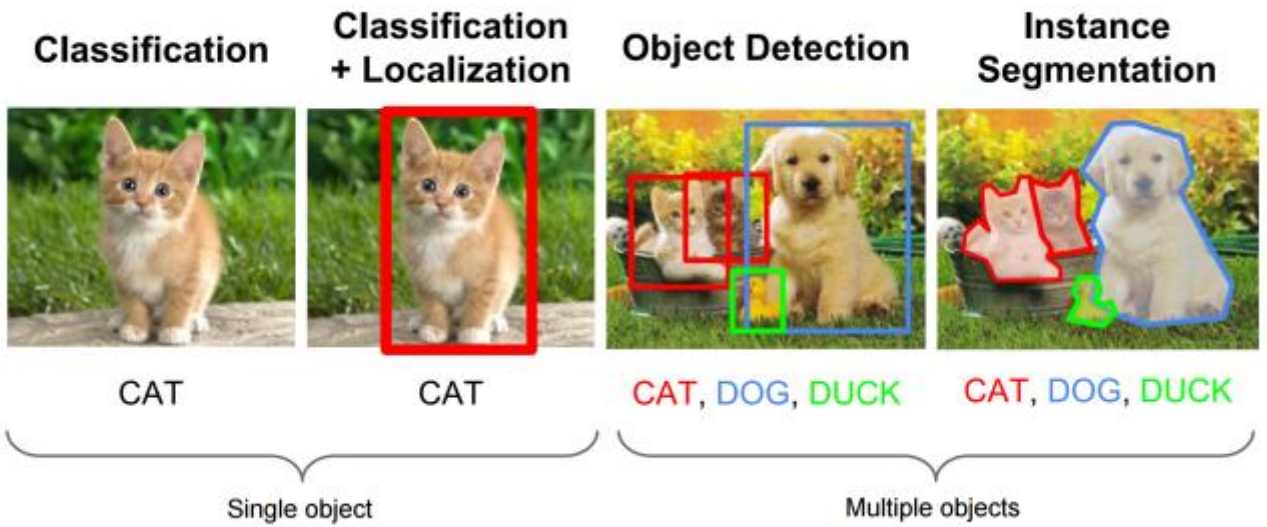
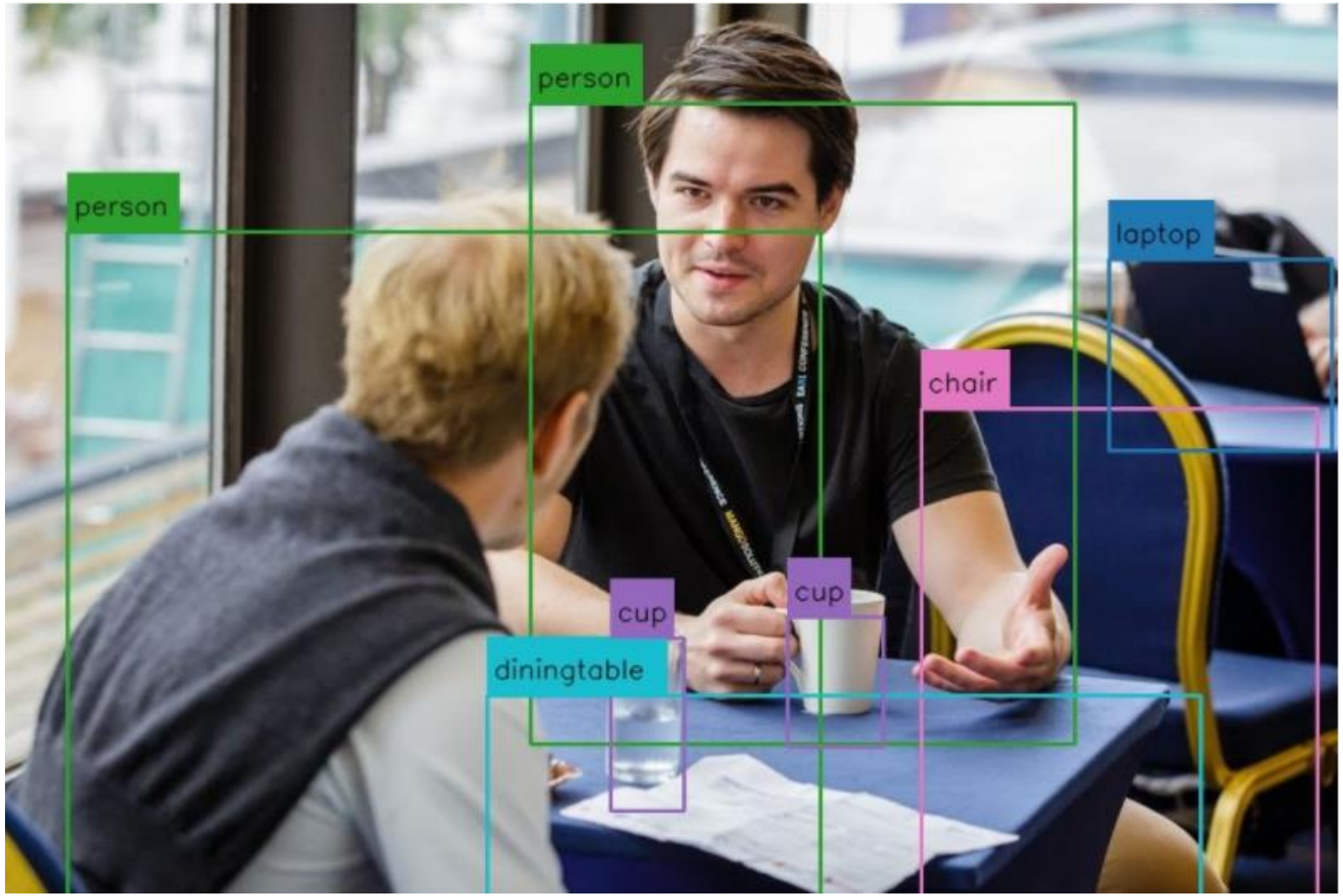
Object Detection

Related Algorithms



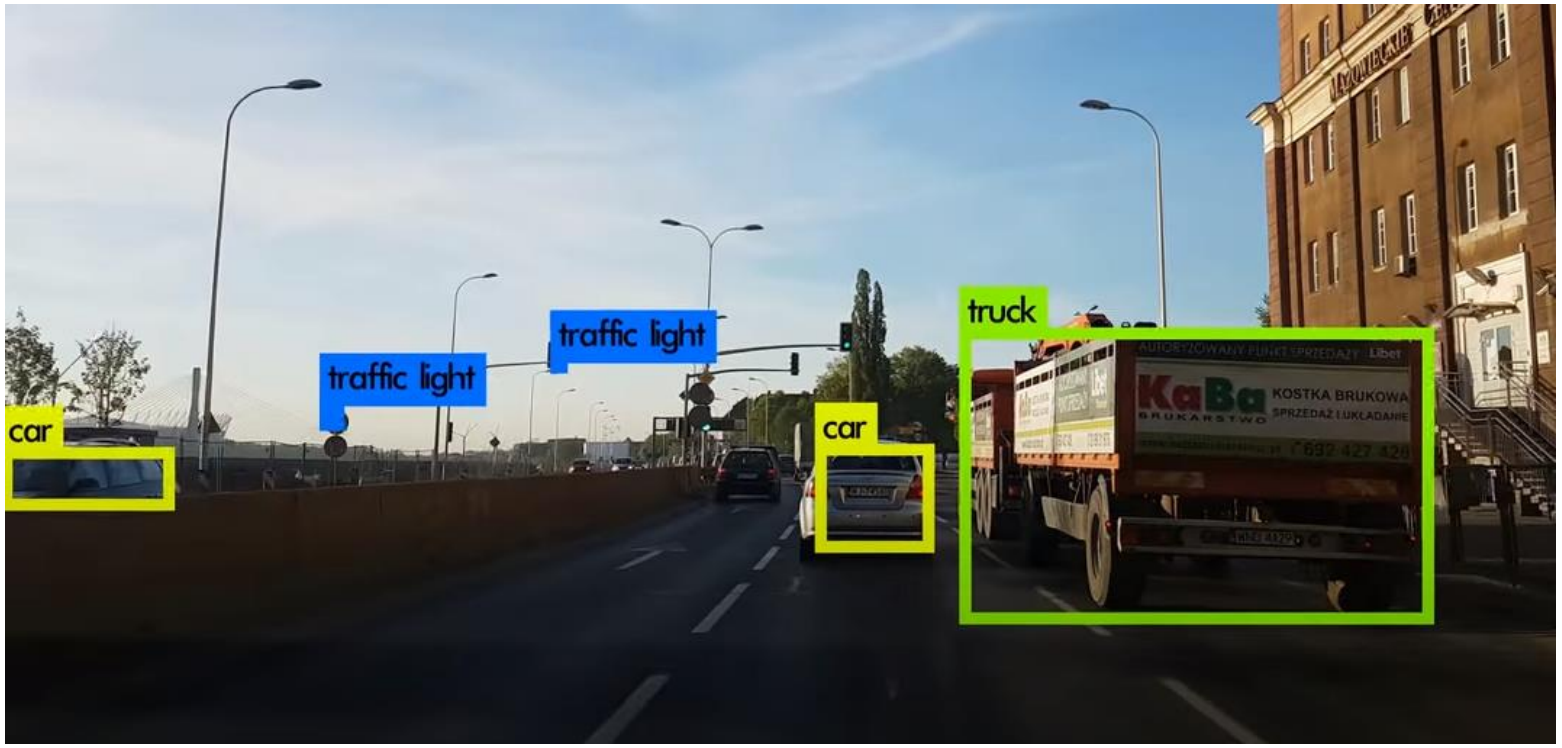
https://appsilon.com/object-detection-yolo-algorithm/



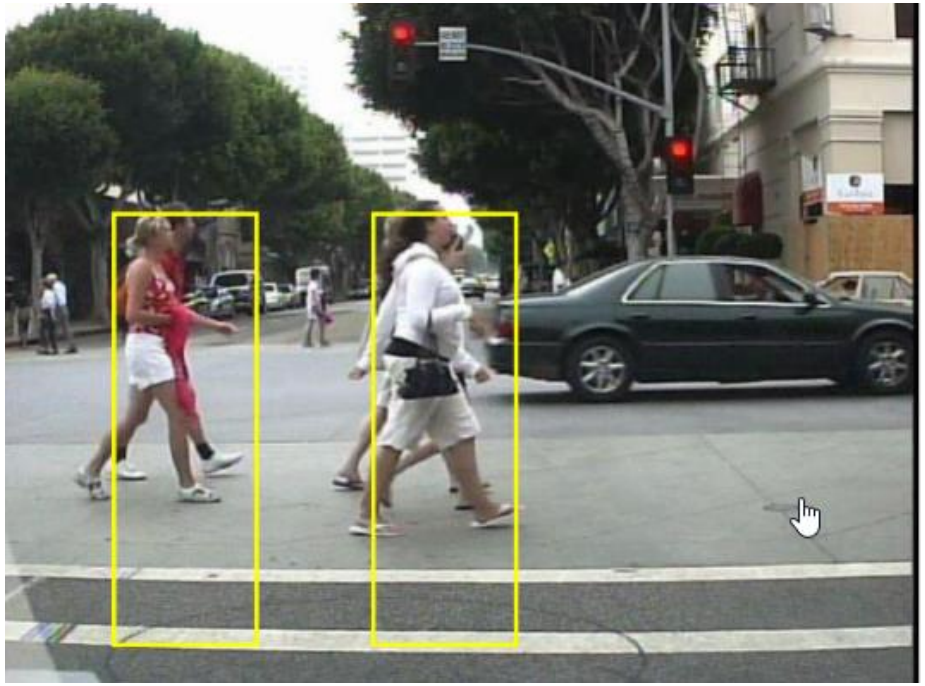
https://appsilon.com/object-detection-yolo-algorithm/

Demonstration: Detection on Video Frames

- https://www.youtube.com/watch?time_continue=525&v=yQwfDxBMtXg&feature=emb_logo



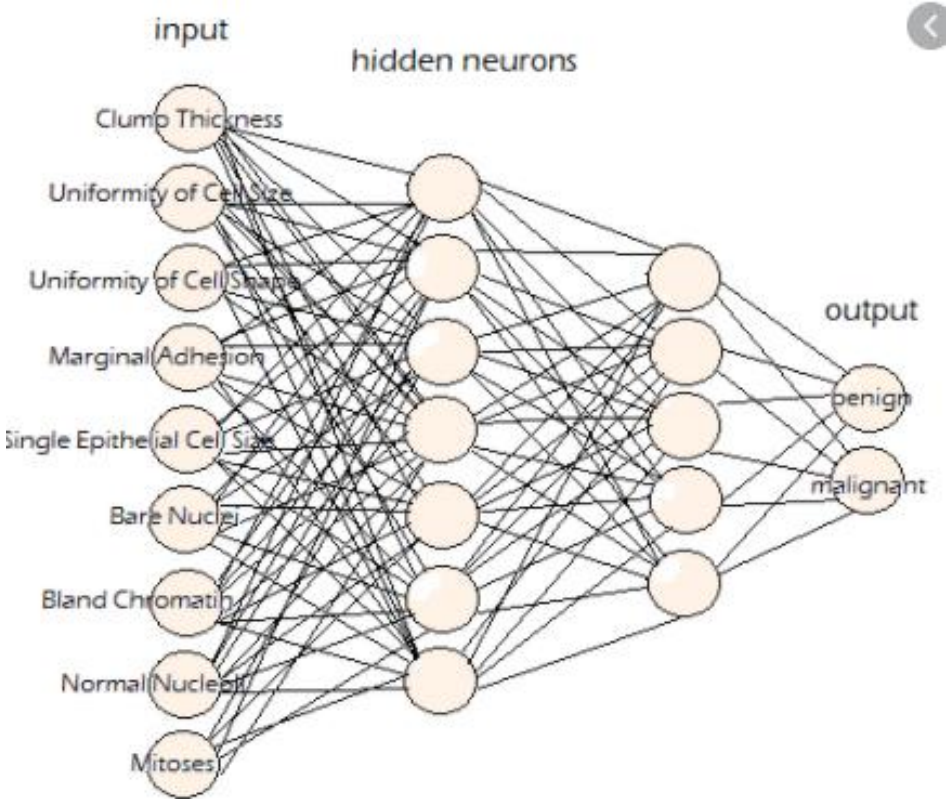
Application



- Surveillance
  - Detect and count pedestrian
  - Crowd monitoring
- Self Driving Cars
  - Detect road users and vehicles
  - Combine vision with Lidar (Light Detection and Ranging)
- Object Recognition
- Activity Recognition

Artificial Neural Network

- ANN is a brain inspired algorithm that learns from data in order to map input signal to the output signal.
- Classification
  - Output signal represents categories e.g benign (normal) and malignant (cancer cell)
- Regression
  - Output signal is a continuous value



Convolutional Neural Network (CNN)

- CNN is a special type of neural network designed to process image. It has been extended to work on other data types (video, text, audio)
- In CNN, the image is transformed by a series of filters with learnable weight
- Recall Laplacian filter used to transform an image.
  - The weights are fixed.

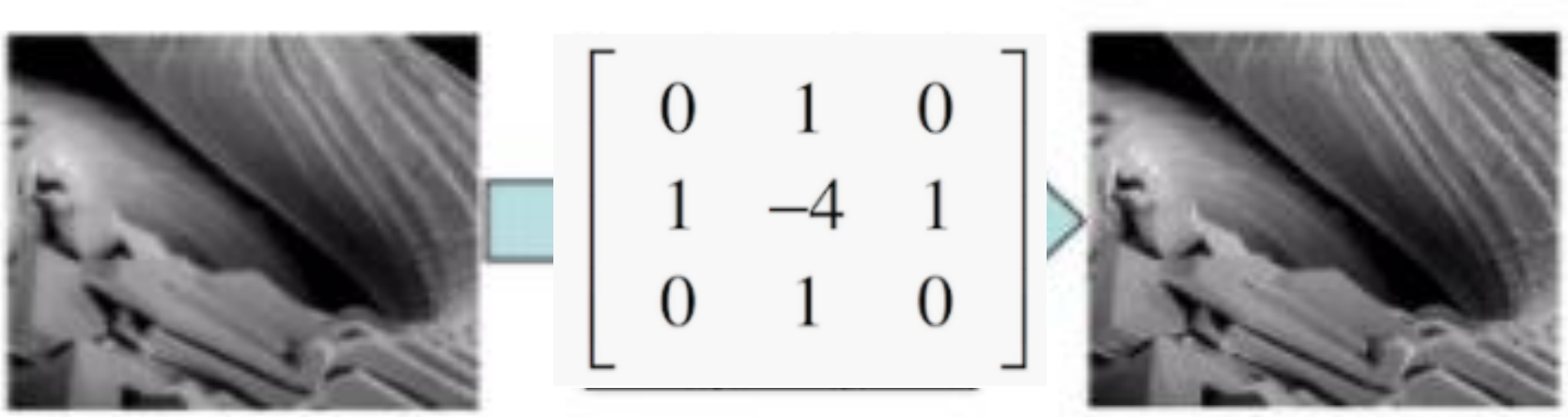
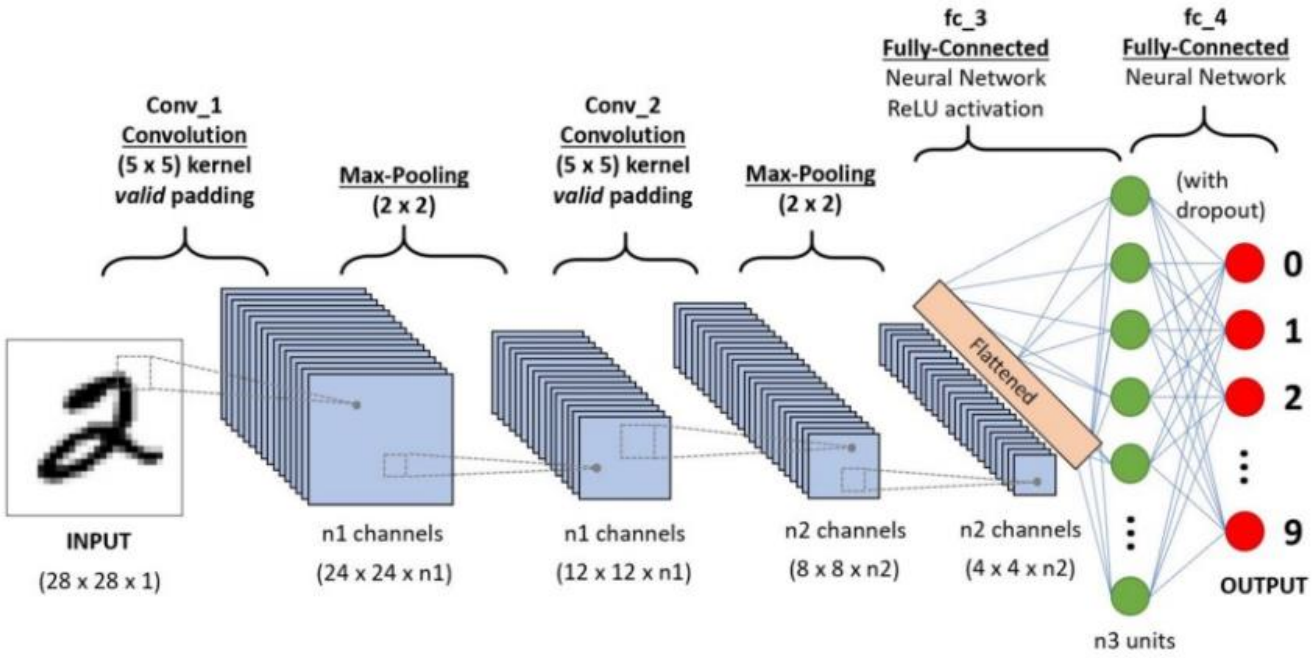


Image Classification with CNN

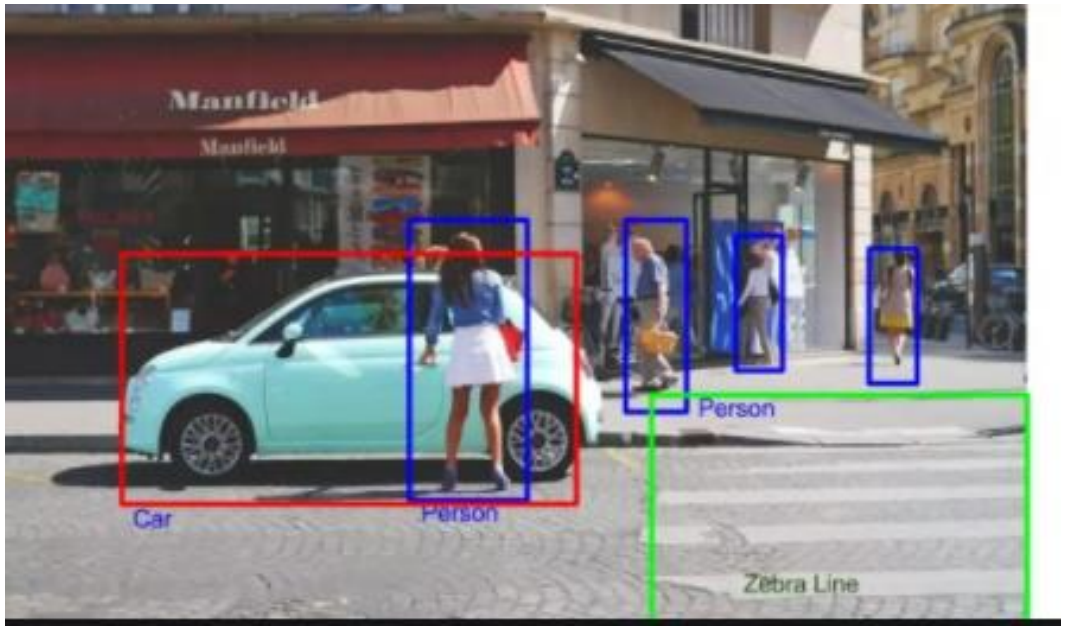
- In CNN, the image is transformed by a series of filters with learnable weight
- The final layer predict the class label of the entire image



A CNN sequence to classify handwritten digits

Object Detection with Convolutional Neural Network (CNN)

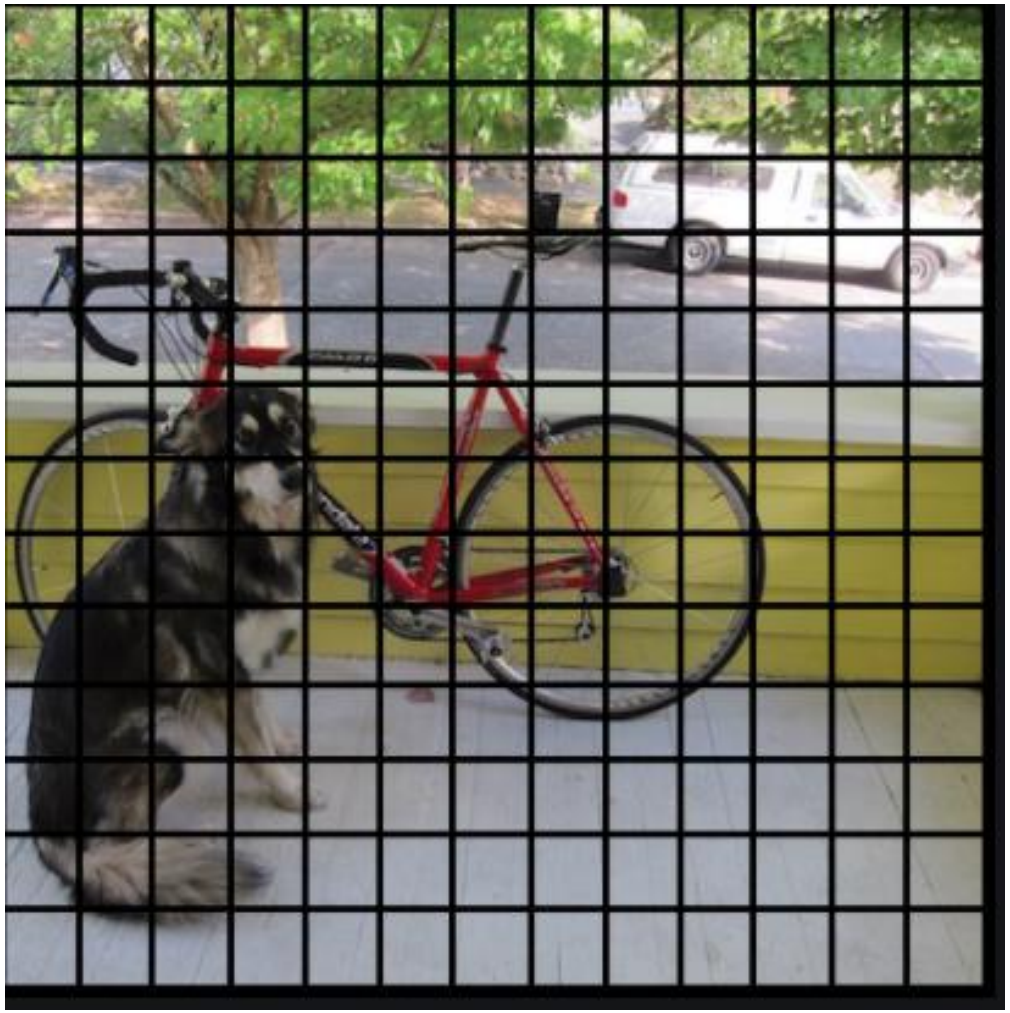
- In classification problem, CNN predict the class label for the entire image
- When applied to detection problem, CNN predict if a location contain the target object class



YOLOv3

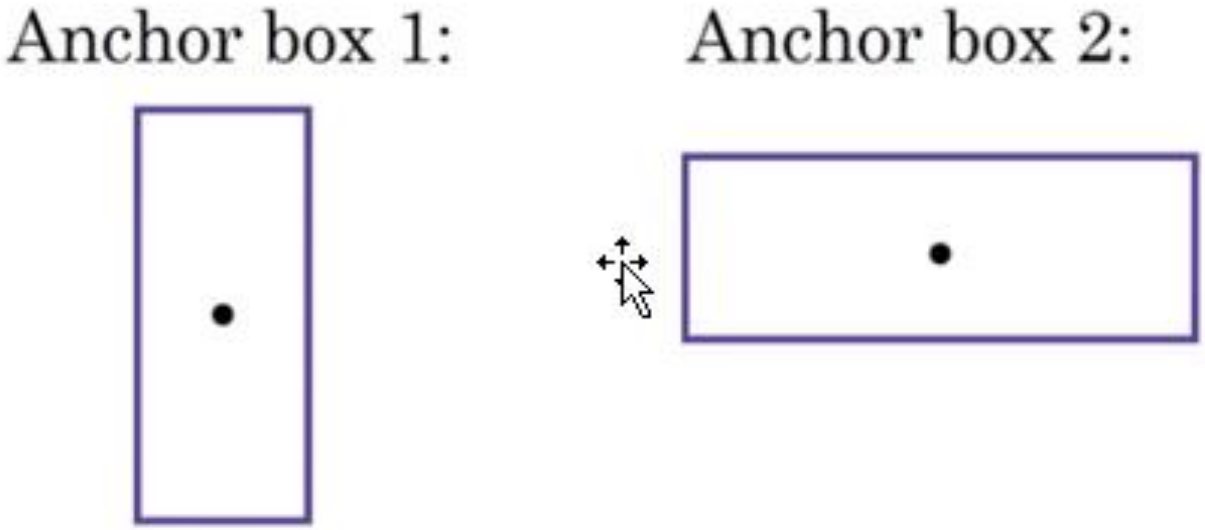
YOLO First Divides Image into Grid of Cells

- Each cell tries to predict the object

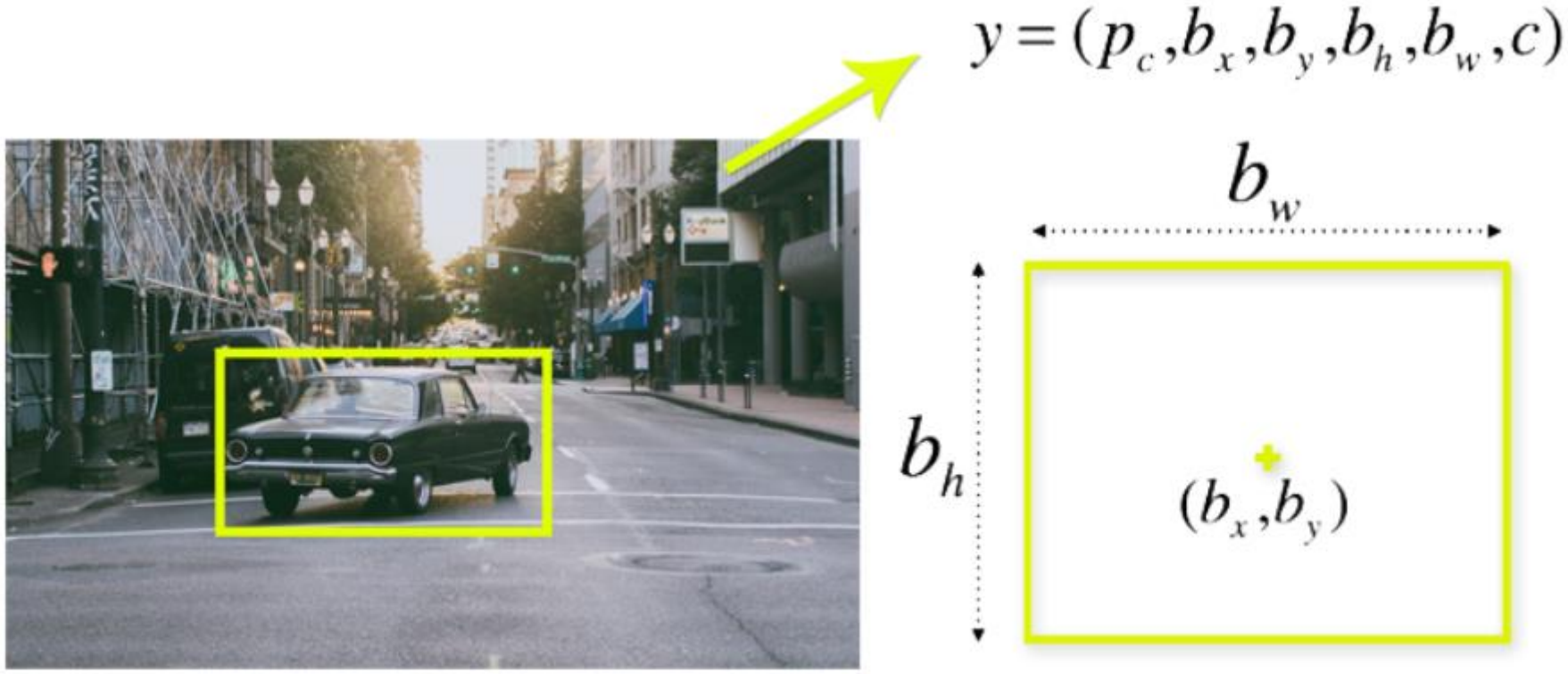


Anchor Box

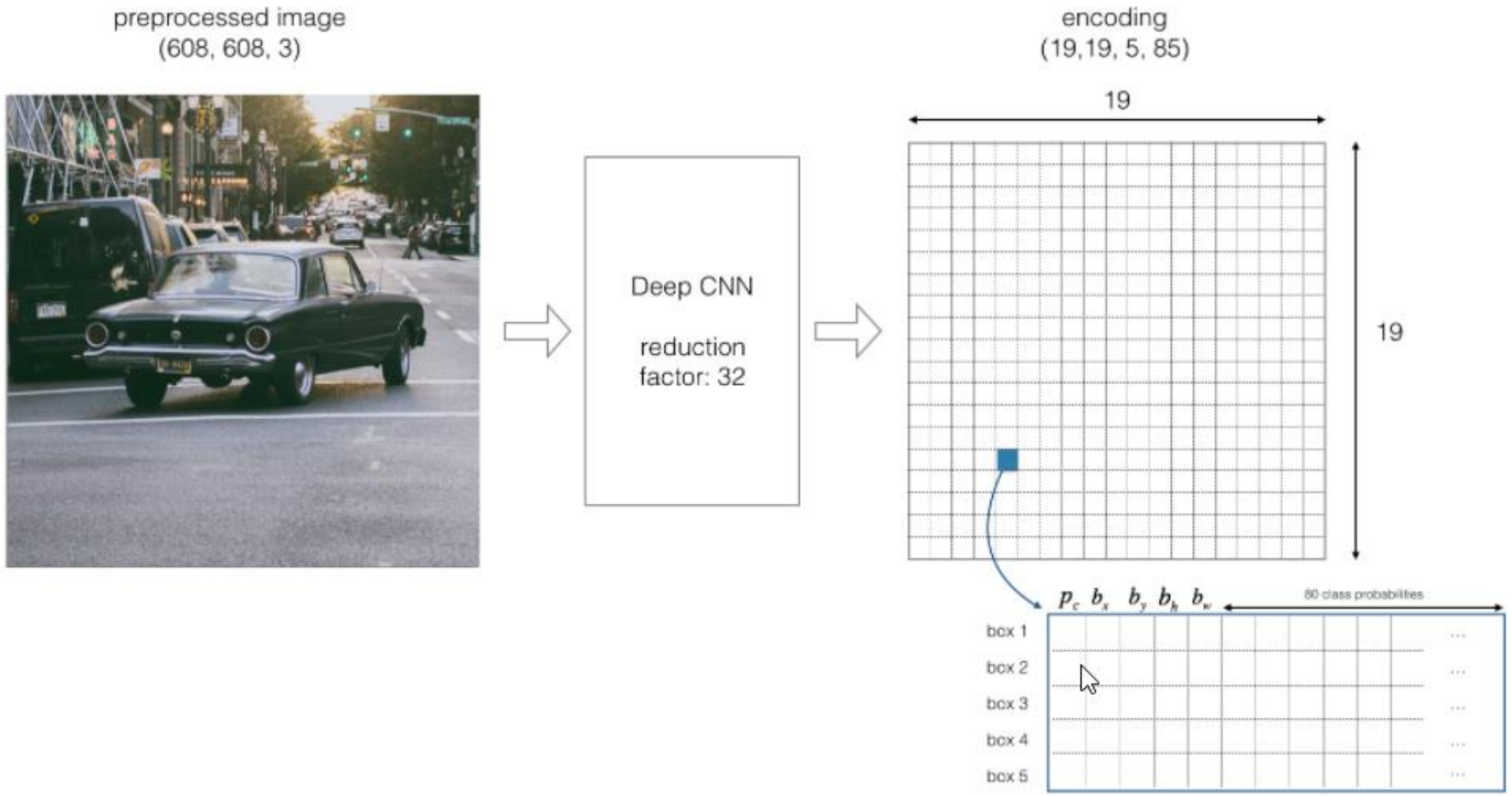
- Each cell predict if object is in the predetermined anchor box
- In this example, used 2 anchor box for each cell
- Predict if the anchor box has the object



Each cell predict a bounding box



Example here shows 5 anchor boxes for each cell

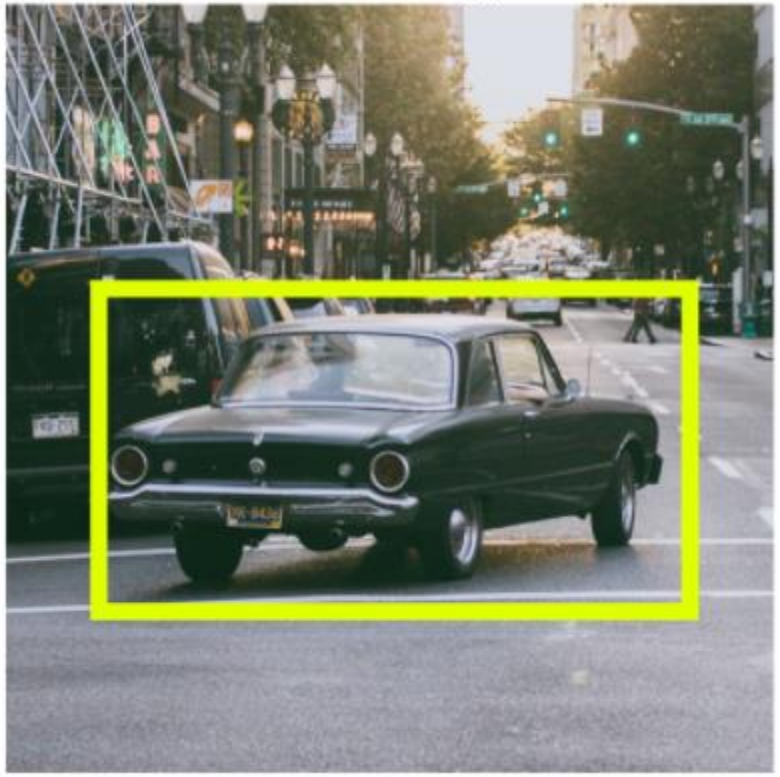


Multiple detections at same location

Before non-max suppression

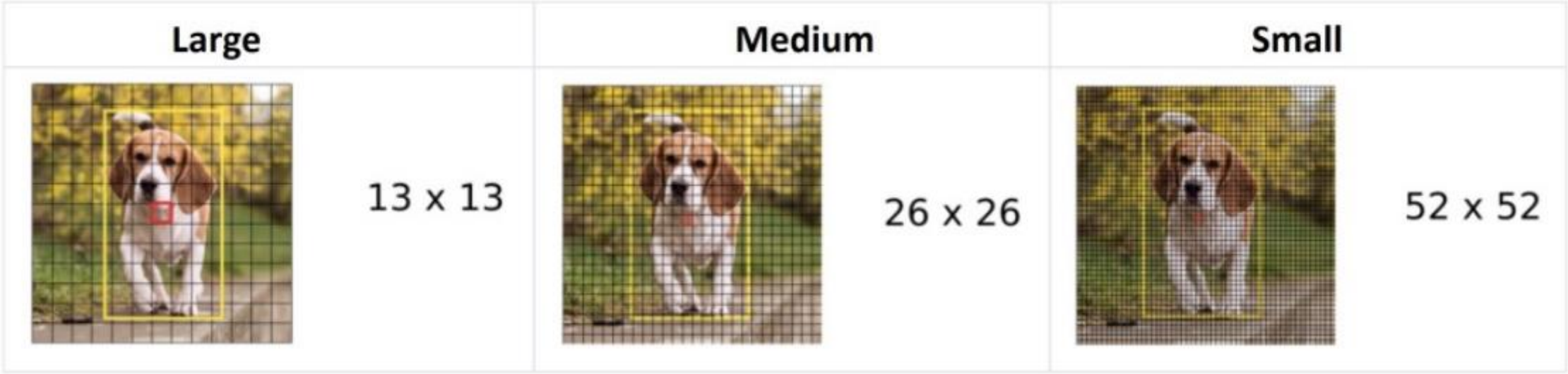
Non-Max Suppression

After non-max suppression



YoloV3 use detection at 3 different scales

- This means, if we feed an input image of size 416 x 416, YOLOv3 will make detection on the scale of 13 x 13, 26 x 26, and 52 x 52.

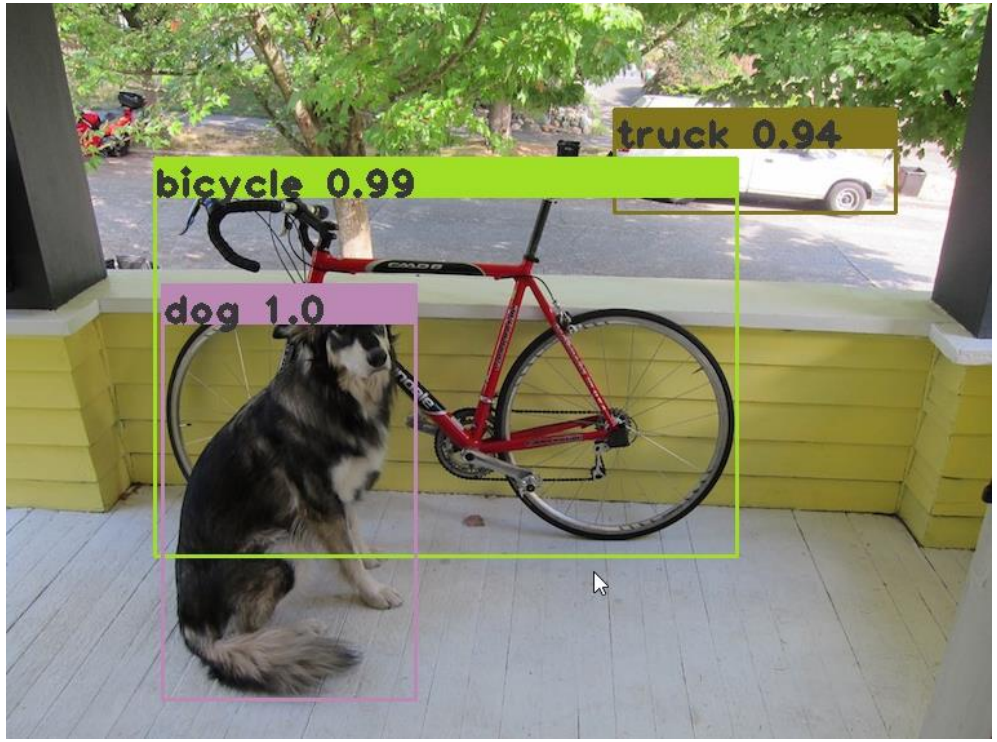


Experiment on YoloV3

- Use implementation from OpenCV
- Source code modified from https://pysource.com/2019/06/27/yolo-object-detection-using-opencv-with-python/

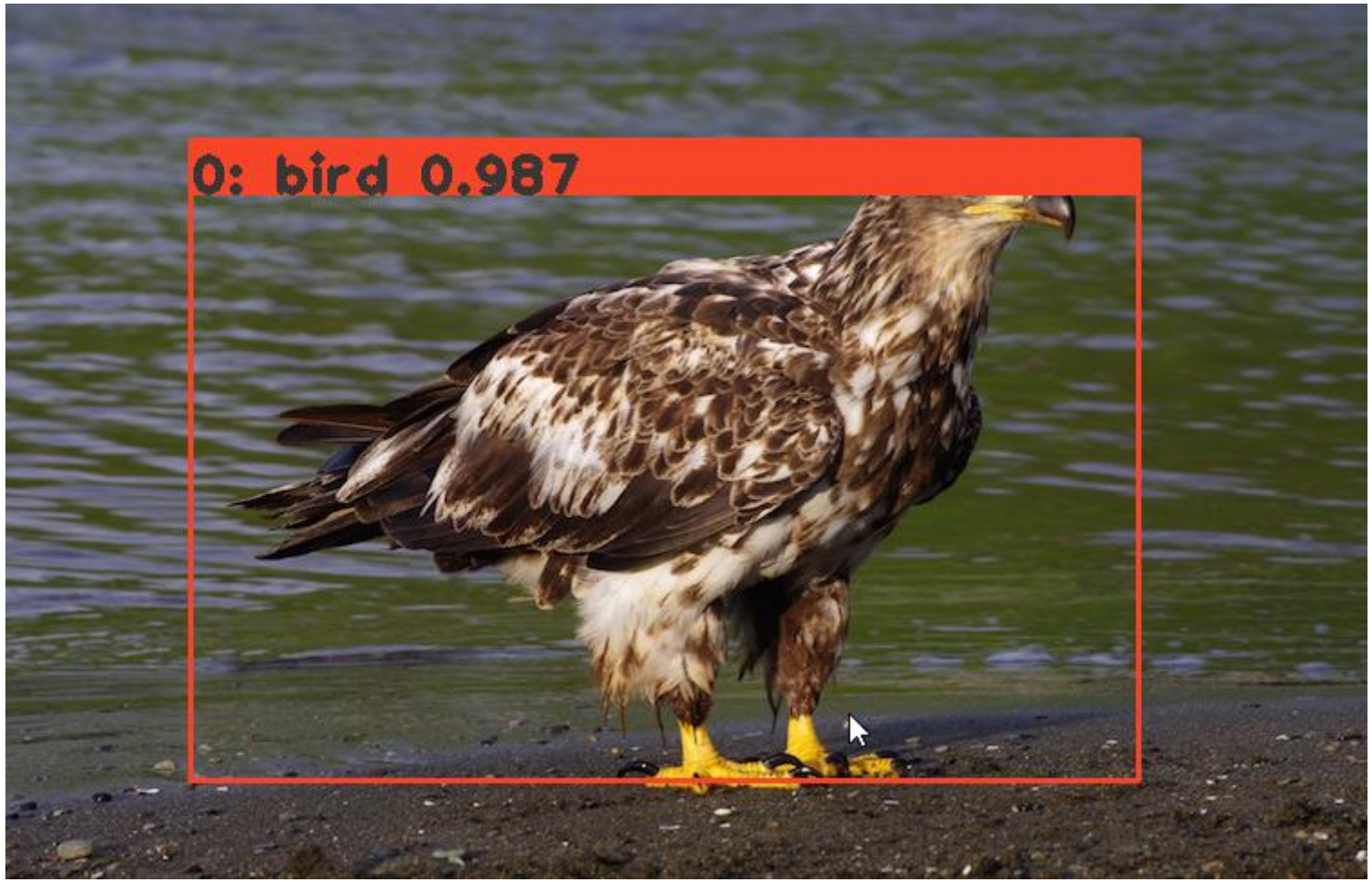
Practical

- **script_03_test_obj_det_on_image.py**
- Run on dog.jpg



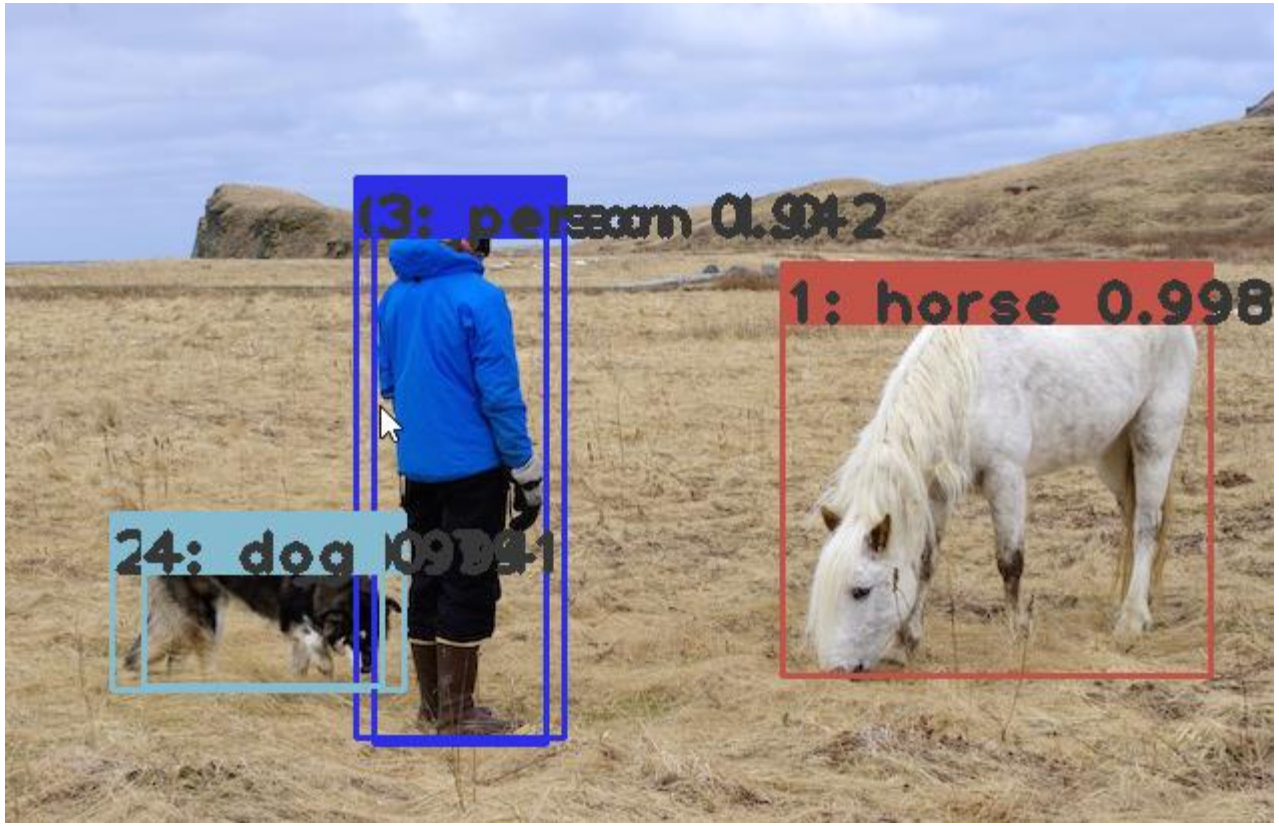
Practical

- Run on filename = 'eagle.jpg' -> Detect 1 object



Practical

- Run on filename = 'person.jpg'-> Detect 3 different objects, overlapping cases



Test on video frames from webcam

- **script_04_test_yolov3_opencv_webcam.py**

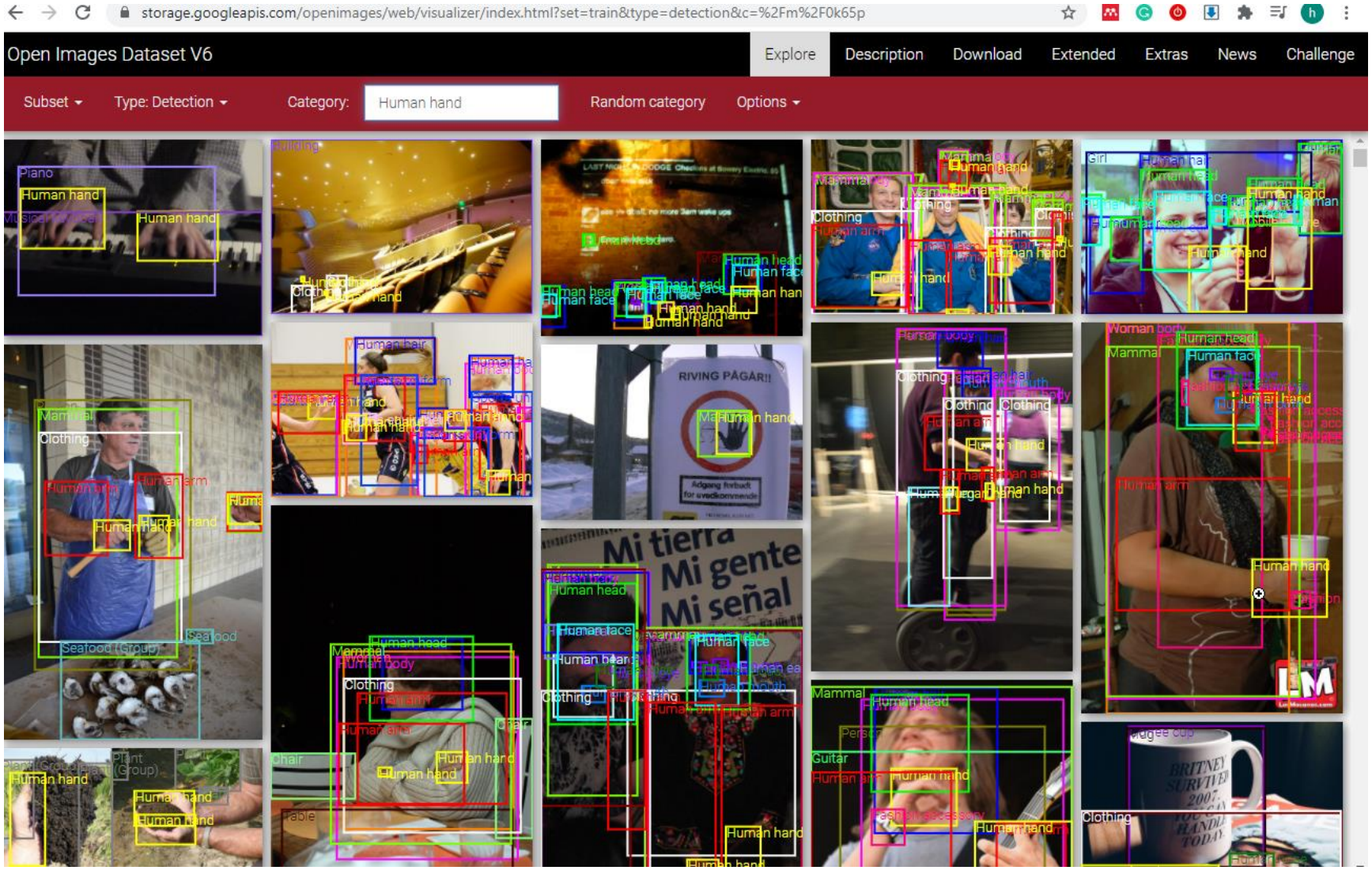
Using smartphone as webcam

- In online invigilation, two webcams provide sufficient coverage of the scene
  - Camera 1 monitor the face
  - Camera2 monitor the hand
- Refer to this article
  - https://www.howtogeek.com/671180/how-to-use-your-android-smartphone-as-a-webcam-on-windows-10/

What if you want to train your own object detector

- This process is more complicated
- Here is one reference
  - https://medium.com/analytics-vidhya/custom-object-detection-with-yolov3-8f72fe8ced79

Training require a lot of data



Annotation

- This is a manual process of assigning bounding box or segmentation mask
  - Some useful software
  - https://lionbridge.ai/articles/image-annotation-tools-for-computer-vision/

LabelImg is a graphical image annotation tool.

It is written in Python and uses Qt for its graphical interface.

Annotations are saved as XML files in PASCAL VOC format, the format used by ImageNet. Besides, it also supports YOLO format