Enrique Quintero

ITAI 1378

Prof. Patricia McManus

L09 Object Detection

Key Concepts, Methodologies and tools.

Object detection combines classification and localization of one or multiple objects in an image or video. In contrast, object classification assigns a label to the image with abstract data, i.e. it identifies a dominant object and classifies it. However, it does not have the ability to identify the position of the object or important data such as bounding boxes, annotations, scores, confidence or intersection union; these attributes are part of object detection.

Most used algorithms such as object detection R-CNN, Fast R-CNN, Faster R-CNN, SSD "Single Shot Multibox Detector or Yolo "You Only Look Once"

Table. Pros and Cons

Algoritm	Speed	Precision Small Objects	Complexity
R-CNN	SLOW	High	High
Fast R-CNN	Medium	High	Medium
Faster R-CNN	Medium	High	Medium
SSD	Fast	Low	Low
YOLO	Faster	Medium	Low

YOLO is ideal for real-time applications, while Faster R-CNN is suitable for tasks where accuracy is critical and SSD MobileNet v2 is a balance between both. In this hands-on test, we were able to see that balance with MobileNet v2.

Another important thing, I use a desktop computer with GPU and I learned that you must install a software called CUDA (https://developer.nvidia.com/cuda-toolkit) without this toolkit you cannot use the GPUs that the computer has. Additionally, as part of the code you must make a resource claim by calling CUDA and its position. You must

also install with PIP the version of the framework that supports GPUs, not all of them support it, for example Pytorch does not support CUDA 12.6.