Notes: Object Detection

One-stage vs. Two-stage Object Detectors

Source: <https://viso.ai/deep-learning/object-detection/>

* Deep learning-based object detectors require two separate tasks:
  + Find an arbitrary number of objects
  + Classify every single object and estimate its size
* Two-stage detectors approximate the regions for objects before using features for classification
  + First, the region containing an object is proposed
  + Then, features are extracted from the proposed region via bounding-box regression and used to classify an object
  + Two-stage detectors achieve higher accuracy but are typically slower
    - Many steps are required per image
  + Two-stage detectors may not be feasible for end-to-end usage as cropping images to determine object locations is often infeasible
  + Examples include:
    - Region convolutional neural network (RCNN)
    - And Faster R-CNN, Mask R-CNN, orGranulated RCNN
    - Feature Pyramid Network
* One-stage detectors predict bounding boxes without the region proposal step
  + These are well suited for real-time applications due to the time saved
  + Accuracy may be compromised to achieve the increased speed
  + Examples include:
    - You Only Look Once (YOLO)
      * Most recent is YOLOv4 or YOLO Representation (YOLOR)
    - Single-Shot Detector
    - RetinaNet