Paper title: You Only Look Once: Unified, Real-Time Object Detection

This paper introduced a new unified model for object detection, YOLO(You Only Look Once),. Comparison to other object detection system. YOLO possesses the fastest speed that is up to 45 frames per second even 150 fps after optimizing, which is the reason why it can be considered as real-time detection system. More to the point, user reasons globally about this image and only need to input once to get result like human visual system, which is why this model is named as YOLO. YOLO makes a different in the picture processing. Unlike the prior detection system repurposing classifiers to perform detection, YOLO reframe object detection as a single regression problem, straight from image pixels to bounding box coordinates and class probabilities. The whole algorithm can be divided into 3 parts. First, resize the input image to 448\*448, then run a single convolutional network on the image. Finally, threshold the resulting detection by the model’s confidence. Making practical comparison to other detection, authors conduct experiment on the PASCAL VOC 2007 and VOC 2012. The result shows that YOOLO perform more efficiently, and generalizes to new domains better. Certainly, there are still some drawbacks in YOLO. YOLO imposes strong spatial constraints on bounding box predictions and it struggles to localize object correctly.

To make a conclusion, YOLO is the fastest general-purposed object detector in the literature and YOLO pushes the state-of-the-art in real-time object detection.