

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

Joseph Redmo, Santosh Divvala, Ross Girshick, Ali Farhadi

May 9, 2016

01 . Introduction

Object Detection System 들은 대부분 classification 모델을 기반으로 했었다. 최근에는 R-CNN 모델도 나왔지만 이 역시 classification 모델에 기반한 것이며 실시간으로 이미디 탐지하는데에는 한계가 있었다.

You Only Look Once(YOLO) 은 classification 모델에서 벗어나 regression 문제를 정의하는 것을 통해 bounding box 을 설정해 클래스와 확률을 구합니다.

빠른 성능과 다른 실시간 시스템 대비 2 배 이상의 average precision 등을 보이는 것과 sliding window 가 아닌 CNN 을 이용해 class 에 대한 표현을 더 잘 학습한다는 것이 YOLO model 의 장점입니다.

02. Unified Detection

YOLO 의 핵심은 Unified Detection 이고 이를 가능케 한 것은 CNN 덕분입니다. 덕분에 bounding box 을 설정할 수 있으며 bounding box regression 과 multi-class classification 을 동시에 수행할 수 있습니다.

네트워크는 다음과 같이 설계되었습니다.

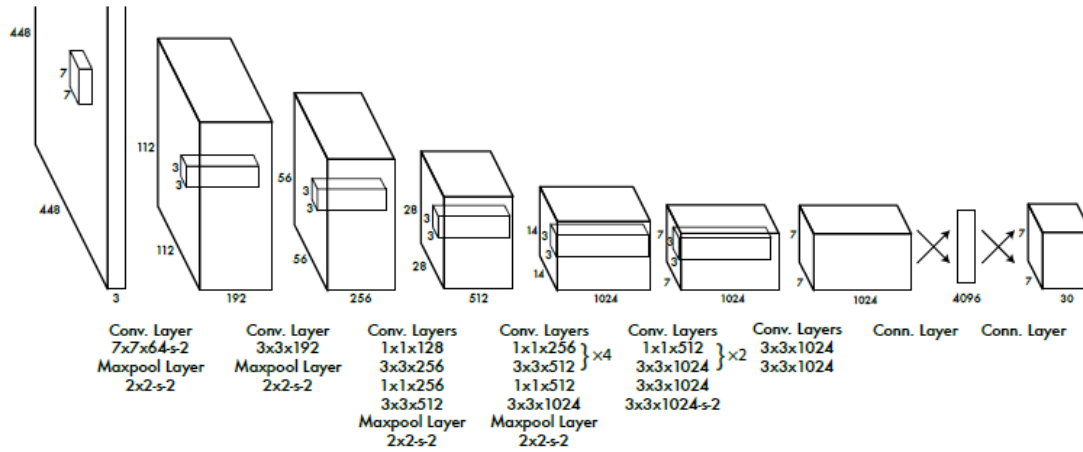


Figure 3: The Architecture. Our detection network has 24 convolutional layers followed by 2 fully connected layers. Alternating 1×1 convolutional layers reduce the features space from preceding layers. We pretrain the convolutional layers on the ImageNet classification task at half the resolution (224×224 input image) and then double the resolution for detection.

YOLO Model 은 CNN 기반으로 만들어졌습니다. 24 개의 convolutional layer 과 2 개의 fully connected layer 으로 구성됩니다. 또한 Fast YOLO 을 소개하며 단순히 24 개 layer 을 9 개로 줄인 것입니다.

04. Experiments

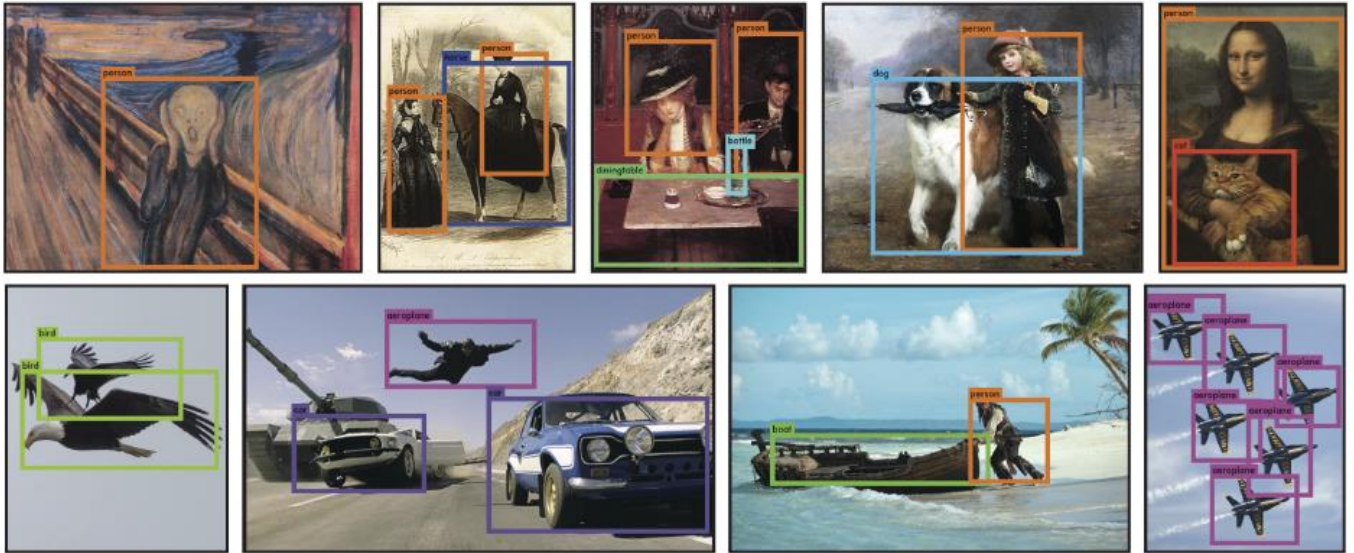
YOLO 은 실시간 성능을 보이면서 mAP(mean Average Precision) 은 63.4% 을 보입니다. 반면에 Fast YOLO 은 가장 빠른 object detection 알고리즘으로 나타났지만 52.7% 의 mAP 를 보였습니다.

Real-Time Detectors	Train	mAP	FPS
100Hz DPM [31]	2007	16.0	100
30Hz DPM [31]	2007	26.1	30
Fast YOLO	2007+2012	52.7	155
YOLO	2007+2012	63.4	45
Less Than Real-Time			
Fastest DPM [38]	2007	30.4	15
R-CNN Minus R [20]	2007	53.5	6
Fast R-CNN [14]	2007+2012	70.0	0.5
Faster R-CNN VGG-16[28]	2007+2012	73.2	7
Faster R-CNN ZF [28]	2007+2012	62.1	18
YOLO VGG-16	2007+2012	66.4	21

05. Qualitative Result

이러한 컴퓨터 비전 연구는 실험값 말고도 실제 실험 결과를 일부 보여주기도 합니다. 그 결과는 아래와 같습니다.

Figure 5: Generalization results on Picasso and People-Art datasets.



사람, 동물, 배, 비행기, 자동차 등을 잘 분류하고 있음을 볼 수 있습니다.

Reference

<https://arxiv.org/pdf/1506.02640.pdf>

https://deepbaksuvision.github.io/Modu_ObjectDetection/posts/04_01_Review_of_YOLO_Paper.html