

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

### 1 2-stage Detector

1.1 Localization과 Classification 를 순차적으로 수행

1.2 정확도는 높지만 시간 오래 걸림

### 2 1-stage Detector YOLO

2.1 Regional Proposal과 Classification를 동시에 해결

2.2 ROI를 먼저 추출하지 않고 전체 이미지에 대해 classification, box regression 수행 -> Fast R-CNN보다 background error가 두 배 이상 low

2.3 속도는 빠르지만 정확도 떨어짐 (trade off)

2.4 Single Convolutional Network로 이미지를 입력받아 여러 개의 bounding box와 각 box의 class를 예측 -> non-max suppression을 통해 최종 bounding box를 선정

2.4.1 Bounding Box prediction : 입력 이미지  $s \times s$  grid로 분할 -> 객체의 중심이 grid cell에 맞으면 grid는 객체를 탐지했다고 표기 -> 각 grid cell은 B개의 Bounding box와 각 bounding box에 대한 confidence score를 예측-> non-max suppression을 거쳐 최종 Bounding Box를 선정

### 3 YOLOv1 성능

3.1 Real-Time Detector에서 좋은 성능을 보이는 것을 확인할 수 있었고 Less Than Real-Time에서도 좋은 mAP를 가지고 있음

### 4 YOLOv1 한계

4.1 각 grid cell은 하나의 class만 예측

4.2 새로운 형태의 Bounding Box의 경우 정확히 예측할 수 없음

4.3 작은 Bounding Box의 Loss가 IoU에 더 영향을 줌