

Yolo 정리

○ 기존

경계박스를 직접적으로 예측하는 대신

○ Yolo 알고리즘

미리 크기가 정해진 앵커박스를 여러개 사용.

• 과정

1. Input image를 resize
2. img에서 작동하는 single convolution network 실행
3. 해당 모델을 통해서 나온 확률 값을 threshold로 잘라서 결과값을 보여줌.

(비 최대값 제거 과정)

Cf. 비-최대값 제거

Outputting the non-max suppressed outputs



- For each grid cell, get 2 predicted bounding boxes.
- Get rid of low probability predictions.
- For each class (pedestrian, car, motorcycle) use non-max suppression to generate final predictions.

**이것의 결과값은 아마도 모든
자동차와 보행자가 검출된 것입니다**

Andrew Ng

• 장점

1. 속도가 빠르다.

- 기존(r-cnn)

R-CNN같은 더 최신의 detection 시스템은 region proposal 방법을 사용합니다. region proposal은 image 안에서 가능성이 높은 bounding box를 생성합니다. 그리고 그 box속의 이미지를 classifier에 넣어 실행시킵니다. 분류 후에는 이전에 진행 했던 과정들을 중복된 detection을 제거하거나, 다른 장면에서의 object를 기반으로 re-score(재평가)를 하는 것 등을 위해서 성능 향상을 위해 재 사용합니다. 이러한 복잡한 pipeline은 각각의 component들을 따로 학습시켜야하기 때문에 느리고 최적화하기가 힘듭니다.

- yolo

그래서 우리는 object detection을 이미지 pixel로부터 bounding box좌표와 분류 과정을 single regression problem으로 재정의하였습니다. 우리 시스템을 사용하면 이미지를 한번만 보면 현재 image에 있는 object가 무엇이고 어디인는지 알 수 있습니다.

YOLO는 정말 쉽습니다. [Figure 1]을 보면 Single Convolution network는 물체의 bounding box와 bounding box안의 object가 무엇인지 동시에 예측합니다. YOLO는 전체 image에서 학습하며 직접적으로 detection performance를 최적화합니다. 이런 통합적인 모델은 전통적인 object detection방법에 비교해서 몇가지 이점을 갖습니다.

첫째로 YOLO는 매우 빠릅니다. 우리가 detection을 single regression problem으로 정의하는 것을 통해서, 우리는 복잡한 pipeline이 필요하지 않습니다. 우리는 테스트 시간에 새로운 이미지로 우리의 신경망을 실행시킬 수 있습니다. 우리의 기본적인 네트워크는 Titan X GPU상에서 batch과정없이 45FPS가 나옵니다. 그리고 빠른 버전은 150 FPS를 상회합니다. 또한 YOLO는 다른 real-time system보다 2배의 mean average precision(mAP)의 성능을 보여줍니다.

2. 이미지 예측 시, globally 하다.

두번째로 YOLO는 image를 예측시에 globally하다는 특징을 가집니다. sliding window나 region proposal 기반의 기술들과는 다르게 학습이나 테스트 중에 전체 이미지를 보기 때문에, 은연중에 각 class의 대표적인 표현들에 대해서 맥락적인 정보를 encoding합니다. 강력한 detection 방법인 Fast R-CNN 경우에는 배경정보에 따라서 object를 detect하는데 있어서 실수가 있는데, 이것은 Fast R-CNN이 큰 맥락을 보지 못하기 때문입니다. YOLO는 Fast R-CNN과 비교했을 때, background error가 절반정도 밖에 되지 않습니다.

3. Object의 일반적인 특징을 학습

세번째로 YOLO는 object의 일반적인 특징을 학습합니다. 즉, object의 일반화를 잘 합니다. 예술작품이나 자연의 이미지를 학습하거나 테스트할 때, DPM이나 R-CNN같은 강력한 detection방법들보다 학습을 더 잘합니다. YOLO는 일반화에 강력하기 때문에, 새로운 도메인이나 예상하지 못한 input이 들어와도 망가질 확률이 적습니다.

참고개념

비-최대값(NMS) 알고리즘

-> 국지적인 최대값만 찾아 그 값만 남기고 나머지 값은 모두 삭제하는 알고리즘

참고 사이트

- https://youtu.be/9s_FpMpdYW8

(올로 알고리즘 영상 - good)

- <http://research.sualab.com/practice/2018/05/14/image-detection-deep-learning.html>

(이미지 Detection 문제와 딥러닝: YOLOv2로 얼굴인식하기 - 파이썬 텐서로 구현하기 좋으나, 좀 어려움.)

- <https://dhhwang89.tistory.com/m/51>

([Deep learning] YOLO!, You Only Look Once : Unified, Real-Time object Detection - yolo 관련 논문 해설)

- <https://dhhwang89.tistory.com/m/138>

([Deep learning] Yolov3: An Incremental Improvement - 괜찮은 자료, 시간 있을 때 보기)