3 다음 사진은 근래에 개발된 object detector인데 2-stage detector, 1-stage detector. 2-stagedetector RCNN, 1-stage detector YOLO SSD FocalLoss-RetinaNet

Retina 망막

4 2stage detector, 1stage detector

Region Proposal, Classification   
클래스에 해당하는 부분을 찾은후 해당부분을 분류한다.

Selective Search  
각 부분을 후보 초기 영역으로 할당한후  
작은영역을 큰 영역으로 합쳐  
통합된 영역을 후보영역으로 만든다.

Bounding box regression  
bounding box regression(선형회귀 모델)이 존재한다

6  
결과적으로는 1stage detector가 더 빠르고 간단하지만 정확도가 떨어진다.

7 정확도 떨어지는 원인으로서 class imbalance를 요점으로 잡았는데  
해결을 위해 기존의 cross entropy loss reshape해서 onestage detector 비슷한 정확도를 성취하게 했다.

8 RCNN 2 stage detector proposal stage에서 candidate object를 줄여나가고   
classification stage 다음과 같은 알고리즘들을 이용해서 foreground 와 background 의 균형을 맞춰 class imbalance가 일어나지 않게합니다

 Fast R-CNN은 hard negative mining 대신에 heurisics method를 사용합니다.

 2장의 이미지에서 각각 64개의 region proposal을 추출하여 총 128 mini-batch를 구성합니다. 128개의 region proposal는 25%는 positive sample, 75%는 negative sample로 추출합니다. positive sample은 IOU가 0.5이상, negative sample은 IOU가 0.1~0.5 사이의 proposal을 지정했습니다. 이 방법을 휴리스틱(heuristocs) 방법

9 1 stage detector 2stage 에서 사용한 heuristics들을 사용하면 비효율적이기 때문에 사용 할 수 없습니다.  
이를 해결하기 위해 focal loss가 나오게 되었습니다.

11 기존은 다음과 같은 sliding window paradigm 으로 진행됬다고 합니다. 하지만, 딥러닝 기반의 방식이 나온뒤로는 바뀜

12 2stage detector, 1stage detector를 사용했을 때 coco set을 학습한 결과  
1stage detector는 빠른대신 정확도가 낮고 2stage는 정확도가 더 높음

13 class imbalance   
- 대부분의 locoation이 학습에 의미없는 negative가 되기 떄문에 비효율적이고  
- 모델의 학습을 잘못된 방향으로 이끌 수 있습니다.

15 robust loss function – focal loss는 class imbalance를 해결하기위해 디자인되서  
hard example detect가 어려운 example에 training을 더 잡중합니다.

16 focall loss는 cross entropy에서 시작했습니다. 변수 p를 모델이 추정한 클래스의 가능성

17 a 값은 가중치를 둠으로써 positive/ negative example을 balance 하는 값

18 focal loss 해당하는 부분인데 다음과 같이 reshape 함으로써 가중치를 둬서 train 하기 어려운 hard example들을 easy example 보다 더 집중해서 training 하는 방식 입니다.

-pt 작으면 modulating factor 1에 가까워지면서 unaffected  
-pt 1에 가까워지면 factor 0에 가까워져서 잘 분류된 example loss가 작아진다.

앞부분은 문맥상 loss가 커진다 이런말 이어댜 될것같다.

Focusing parameter 쉬운 example을 가중치 낫게 조정합니다.

19 기존 focal loss에서 balance 값인 a를 넣는게 정확도가 높았다.

22 retinanet은 backbone network와 2개 subnetwork로 구성되어있습니다.

Backbone은 전체 이미지를 convolutional feature map을 compute 하고   
first subnet은 backbone에서 나온 출력에서 object classification  
두번째 subnet은 bound box regression을 수행합니다

bound box regression .   
물체를 탐지하는 bounding box regression(선형회귀 모델)이 존재한다

23 FPN는 신경망을 통과하며 featuremap 만들어주고 위에서부터 다시 합쳐주면서 object detect

24 retina net 같은 경우는 ResNet에서 backbone으로 해서 모든피라미드 레벨을 채널 256개로 p3~P7까지 피라미드를 만듭니다.

26 p3~p7까지 피라미드 레벨에 다음과 같은크기 area로 만들고

각 피라미드에 다음과 같은 비율로 anchor를 만듭니다. 논문에서 보도되넋볻 ㅏ깊어졌을때는 anchor 사이즈를 다음과 같이 해서 적용

레벨마다 9개의 anchor **scale의 범위는 32~813 pixel**

27 IOU값이 0~0.4 background로 인식하고 0.4~0.5는 무시하였습니다.

28 classification subnet은 각 부분에 물체 존재 가능성을 예측합니다. 각 FPN에 FCN을 붙임

29 Box subnet 각 spatial location에 4개의 anchor가 regression을 진행합니다.  
rcnn 논문을 따라가 보면 중앙점 좌표와 width, height

34 resnet 50, 101 imagenet 사전 training 마지막을 제외하고 bias\*0 가우시안 weight 0.01  
마지막 레이어는 bias initialization   
최적화로는 SGD   
초기 학습률 0.01 이후로 감소시키며 90000번 학습시켰습니다.

36. retinanet cross entropy loss 사용했지만 이는 실패했고

Focal loss 써서 retinanet resnet 50으로 시작했더니 30.2가 되었다.

A balanced 추가해서 써서 더 정확도가 올라갔고

최종적으로 ~~

37

OHEM(Online Hard Example Mining)은 hard example과 easy example로 이루어진 dataset에서 hard example을 더 많이 샘플링하는 알고리즘입니다. OHEM이 고안된 이유는 detection dataset이 hard exaple보다 easy example을 훨씬 많이 포함하고 있기 때문에 학습에 악영향을 주게되고 이 문제를 개선하기 위해 고안되었습니다. 실제로 Fast R-CNN에 OHEM을 사용하여 모델의 학습 속도 개선과 성능 향상을 이뤄냈습니다.

OHEM Fst R-CNN 쓰지 않은이유

SGD(Stochastic gradient descent)를 사용하여 학습되는 ConvNet에서 기술적인 어려움때문에 hard negative mining을 사용하지 못합니다. 모델이 False Positive를 판단하고 이를 학습 데이터셋에 추가하고, mini-batch를 구성하는 과정이 끝날 때 까지 모델을 update할 수 없어, 학습이 느려지기 때문입니다.