

과제 #4 - 기말고사 문제풀이

신원인공지능학과

2023254023 김경호

1. Image Classification : 주어진 이미지가 어떤 카테고리에 속하는지 판단하는 과정.

태양 그림에서 이미지 전체를 하나의 레이블로 분류.

인러닝 모델이 이미지의 특징을 학습하고 각 카테고리에 대한 확률을 출력하도록 훈련

• Object Detection : 이미지 내의 여러 객체를 식별하고, 각 객체의 위치를 정확히 파악하는 과정, 이미지 내에서 객체의 종류뿐만 아니라 위치도 예측, 이를 통해 이미지에 여러 객체가 존재하는 경우 각각의 개체별로 식별

• Semantic Segmentation : 이미지를 구성하는 각 픽셀로 특정 클래스에 할당하는 과정.
모든 픽셀이 라벨링되며, 같은 종류의 객체들은 동일한 레이블로 표시됨

• Instance Segmentation : 같은 종류의 여러 객체를 개별적으로 구별

2. 2-stage Detector : 이미지에서 잠재적인 객체의 위치를 찾는 region proposal 단계가, 이렇게 제안된 영역을 분류하는 classification 단계로 구성됨.

정확도는 높지만, 계산 과정이 복잡하고 속도가 느릴 수 있음.

• R-CNN : 이미지에서 약 200 개의 후보 영역을 선택하고, 각 영역에 대해 CNN을 적용하여 특징을 추출한 뒤, SVM을 사용하여 객체를 분류. 각 후보 영역마다 CNN을 적용하여 계산 비용이 매우 높음.

• Fast R-CNN : R-CNN의 계산 비효율성을 개선하기 위해 전체 이미지에 대해 한 번만 CNN을 적용하고, 추출된 특징 맵을 기반으로 각 후보 영역의 특징을 재사용함. ROI Pooling 레이어를 통해 이미지 전체 분류와 비효율적인 ROI를 동시에 수행하여 성능을 향상.

• Faster R-CNN : Fast R-CNN을 더 발전시켜, 후보 영역을 제안하는 과정에서 RPN 사용.
RPN은 이미지 객체가 있는 부분의 위치를 비로 효율적으로 찾아내며, 이 결과를 Fast R-CNN과 통합하여 전체 프로세스의 속도와 정확도를 높임.

3. (1) IOU : 객체들 간에 두 바운딩 박스 사이의 겹치는 정도를 측정하는 지표.
 겹치는 영역과 두 박스의 합집합 영역의 비로 계산.
 이미지에서 빨간색 박스로 예측된 위치라고 하고, 초록색 박스로 실제 객체 위치라고 할 때,
 IOU는 이 두 박스가 겹치는 영역을 빨간색 박스의 초록색 박스의 합집합의 비율값이다.

(2) NMS : 여러 개의 겹치는 바운딩 박스 중에서 가장 확률이 높은 바운딩 박스를 선택하고
 나머지 겹치는 박스들은 제거하는 방법

(3) AP : 특정 클래스에 대한 객체의 평균값. 정답도-지정된 라인 아래 면적을 계산하여 얻어짐.
 코일이 레이블 클래스의 객체를 얼마나 잘 검출하는지 평가.

mAP : 다양한 클래스 각각에 대한 AP의 평균값.

객체 검출 성능의 다양한 종류의 객체들은 얼마나 잘 검출하는지를 종합적으로 평가하는 지표.

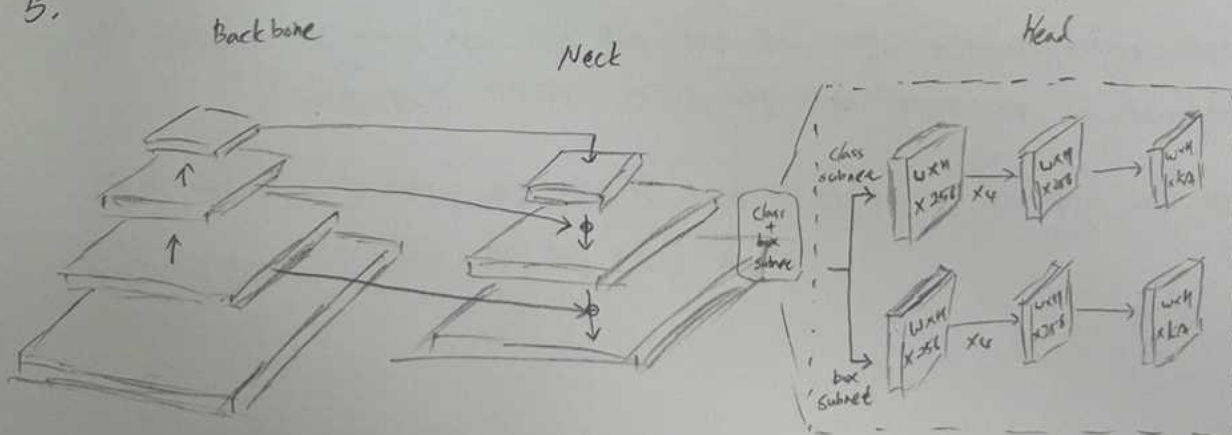
(4) Anchor Box : 객체 검출을 위한 일종의 참조 박스. 다양한 레이어에서의 객체 예측

4. 그리드 수평 크기 : 3 , 그리드 수직 크기 : 3

앵커 박스 수 : 9 , 앵커 박스 수 : 2

전체 층, 텐서는 (3, 3, 2x9) 텐서의 차원을 가진다. 즉, (3, 3, 18)의 차원을 가진다.

5.



- Backbone : 이미지로부터 기본적인 특징을 추출하는 역할. Darknet-53 사용.
53개의 합성곱 층을 포함하며, 이 층들은 각각의 스트라이드와 필터 크기는 가변다.
- Neck : 다양한 스케일에서 객체 특징을 추출하고 집계하는 역할.
여러 FPN 구조가 사용되어 서로 다른 레이어의 feature map을 상위 및 하위 레벨 간에 전달하여 더 풍부한 컨텍스트 정보를 제공한다.
- Head : 최종적인 객체 검출을 위해 각 그리드 셀에 대한 바운딩 박스, 객체 클래스, 객체 존재 확률 예측.

6. FPN : 각 레이어에서 추출된 피쳐 맵은 상향 샘플링 및 병합하는 과정을 통해 저층 정보를 통합함.
각각의 스케일에서 독립적으로 객체 클래스, 위치, 크기 및 클래스 예측.

PAU : 더 깊은 네트워크 레이어에서 피쳐 맵을 하위 레벨로 전달하고, 이는 다시 상위 레벨로 전달하는 양방향의 정보 흐름을 생성.
객체의 위치를 정확하게 파악하는 데 필수적인 저층의 특징과 객체를 식별하는 데 필요한 고층의 특징을 모두 활용.

7. PASCAL VOC : XML 형식으로 제공. 각 이미지에 대해 객체의 클래스, 바운딩 박스 좌표, 이름, 수를, 가려진 상태 등이 있음.

MS COCO : 각 객체에 대해 바운딩 박스의 좌표, 클래스, 그리고 인스턴스 세그멘테이션 마스크를 제공함.
JSON 형식으로 저장되며, 각 이미지는 객체의 카테고리, 위치, 크기 정보도 포함.

Google Open Image : 이미지에 대한 바운딩 박스, 객체 인스턴스 번호 및 객체 속성에 대한 정보는 CSU 파일 형식으로 제공.