[쿠글 week\_8 논문 요약]

10기 진태완

1. 전이학습 관련 (ppt 1번째 링크)

1.1 서론

이미지 세분화(Image Segmentation)는 컴퓨터 비전 분야에서 이미지 내의 각 픽셀을 의미 있는 영역으로 분류하는 기술이다. 최근 딥러닝의 발전으로 이 정확도와 효율성이 크게 향상되었으며, 해당 링크의 글에는 FCN, U-Net, SegNet, DeepLab, ResNet 등의 딥러닝 기반 모델을 중심으로 이미지 세분화 기술을 소개하고 있다.

1.2. FCN

FCN은 기존의 CNN을 변형하여 이미지 세분화에 적용한 모델로서, 주요 특징으로는 네트워크 뒷단에 전결합층 대신 컨볼추션 층을 사용하여 위치 정보를 유지하면서, 각 픽셀의 클래스를 예측할 수 있다는 점이 있다. 다양한 버전이 있으며, Skip Architecture를 사용하여 깊은 층의 의미적 정보와 얕은 층의 외관 정보를 결합해 성능을 향상시킨다.

1.3. U-Net

U-Net은 의학 이미지 세분화를 위해 개발된 모델로, 빠른 연산 속도와 높은 정확도를 가진다. U 형태의 아키텍처는 인코더와 디코더로 구성, 인코더는 이미지의 특성을 추출하고, 디코더는 이를 기반으로 원래 크기의 이미지를 복원한다. 또한, 다층의 출력 결과를 동시에 검증하여 더 정교한 세분화가 가능하다.

1.4. SegNet

SegNet은 FCN과 유사한 구조이며, 인코더-디코더 아키텍처를 사용한다. 주요 차이점은 Max Pooling Indices를 사용하여 디코딩 단계에서 업샘플링을 수행한다는 점이다. 이를 통해 FCN보다 적은 학습 파라미터로 높은 효율성을 달성할 수 있다. 인코더는 VGG16의 컨볼루션 층을 사용하며, 디코더는 이 인덱스를 이용해 피처 맵을 복원한다.

1.5. DeepLab

DeepLab 시리즈는 Atrous Convolution을 활용해 다양한 해상도의 특성 맵을 추출한다. 여러 버전으로 발전해왔으며, 각 버전마다 성능 향상을 위한 새로운 기법이 도입되었다. V3+ 버전은 Separable Convolution과 Atrous Convoluion을 결합하여 더욱 정밀한 세분화를 가능하게 한다.

2. YOLOV8

이 글은 객체 검출 기술 소개와 YOLO 알고리즘에 대해 다루고 있다. 기존의 검출 모델은 DPM과 R-CNN이 있는데, 이들은 느리고 복잡한 후처리 과정이 필요한 반면, YOLO는 객체 검출을 단일 회귀 문제로 간주하며, 이미지를 한 번만 보고 객체의 위취와 클래스를 예측하는 과정을 통해 높은 속도와 정확도를 제공한다.

YOLO의 특징을 세 가지로 소개하자면, 첫째, YOLO는 컨볼루션 네트워크를 사용해 여러 bounding box와 클래스 확률을 한 번에 계산해 매우 빠르다. 둘째, 이미지 전체를 고려해 예측하기 때문에 클래스의 형태 외의 주변 정보까지 학습하고, 이는 이전 모델들의 오류를 줄이는 효과를 가진다. 셋째, 물체의 일반적인 부분을 학습해 자연 이미지를 이용할 때 다른 모델에 비해 뛰어난 성능을 보인다.

YOLO는 객체 검출을 위한 단일 신경망 모델로, 이미지를 S x S 그리드로 분할하여 각 그리드 셀이 B개의 bounding box와 클래스 확률을 예측한다. 각 bounding box는 (x, y, w, h, confidence)로 표현되며, 테스트 단계에서는 클래스 확률과 confidence score를 곱하여 최종 객체 검출을 수행합니다. 실험 결과는 높은 정확성과 실시간 처리가 가능함을 보여준다.

YOLO는 하나의 CNN(Convolutional Neural Network) 구조로 디자인되었다. 파스칼 VOC 데이터셋을 활용하여 모델링되었다. 이 모델은 이미지를 S x S 그리드로 분할하고 각 그리드 셀이 B개의 bounding box와 클래스 확률을 예측한다. YOLO는 사전훈련된 ImageNet 데이터셋으로 컨볼루션 계층을 사전 훈련하고, 나머지 계층을 추가하여 객체 검출을 수행한다. 이 모델의 최종 출력은 클래스 확률과 bounding box의 위치 정보이다. 훈련 단계에서는 localization loss와 classification loss를 최소화하여 학습하며, 이를 위해 SSE(sum-squared error)를 사용한다. 객체의 존재 여부를 고려한 가중치 및 데이터 증강 기법을 사용하여 과적합을 방지하고, learning rate를 조정하여 최적의 학습을 진행한다.

추론 단계에서는 YOLO는 한 이미지 당 98개의 bounding box를 예측하고, 다중 검출 문제를 비 최대 억제 기법으로 해결한다. 그러나 YOLO는 공간적 제약으로 작은 객체를 명확하게 검출하지 못할 수 있으며, 크기에 관계없이 모든 bounding box에 동일한 가중치를 부여하여 정확성이 낮아질 수 있다.

DPM은 슬라이딩 윈도 방식을 사용하며 독립적인 파이프라인으로 구성되어 있다. 반면 YOLO는 단일 모델로 모든 절차를 수행한다. R-CNN은 region proposal 방식을 사용하며 YOLO와 비슷한 면이 있지만 복잡한 파이프라인으로 인해 속도가 느리다.

Fast YOLO는 빠르고 정확한 객체 검출 모델로, Fast R-CNN보다 2배 높은 정확도를 제공한다. YOLO는 다른 빠른 모델과 비교하여도 높은 성능을 유지한다. Faster R-CNN과 비교하면 YOLO는 속도와 정확도 모두에서 우세하다.

VOC 2007 에러 분석에서는 YOLO가 localization error가 높지만 background error가 낮은 것으로 나타났다. Fast R-CNN과 YOLO를 결합하면 성능이 향상되며, VOC 2012 결과에서도 YOLO가 뛰어난 성능을 보인다. 테스트에서는 YOLO가 예술 작품에서도 다른 모델보다 높은 정확도를 유지한다.

YOLO는 단순하면서 빠르고 정확하고, 또한 훈련 단계에서 보지 못한 새로운 이미지에 대해 객체를 잘 검출한다는 특징을 가진다. 따라서, 새로운 이미지에 강하며, 어플리케이션에서도 활용할만한 가치가 충분하다.