

Final report (Option 1)

Course name: Multimedia Processing Technique (DD026_1594)

Project number: Option 1 – 4

Project name: Survey the paper: Recent progress in semantic image segmentation

Paper link: <https://arxiv.org/ftp/arxiv/papers/1809/1809.10198>

Student name: Kim Kyungwon. 201810182

Github of project: <https://github.com/abc>

Report:

시맨틱 이미지 세분화의 최근 진행 상황

이미지 처리 및 컴퓨터 비전 영역의 핵심 애플리케이션 중 하나가 되는 시맨틱 이미지 분할은 의료 영역 및 지능형 운송과 같은 여러 영역에서 사용되어 왔습니다. 픽셀 수준 분류라고도 하는 시맨틱 이미지 분할은 동일한 객체 클래스에 속하는 이미지 부분을 함께 클러스터링하는 작업입니다. 시맨틱 이미지 분할 방법을 기존의 DNN 방법과 최근의 DNN 방법의 두 가지 범주로 나눕니다. 먼저, 분할을 위해 공개된 데이터 세트뿐만 아니라 전통적인 방법을 간략하게 요약한 다음, 8가지 측면에서 설명된 DNN을 기반으로 한 최근 방법을 종합적으로 조사합니다: 완전 합성 네트워크, 업샘플링 방법, CRF 방법과의 FCN 조인트, 확장된 합성 접근법, 백본 네트워크의 진전, 분할 피라미드 방법, 다단계 특징 및 다단계 방법, 감독, 약하게 감독되는 방법 및 비감독 방법. 마지막으로, 이 영역의 결론이 도출됩니다.

2. 데이터 세트 및 평가 지표

2.1 데이터세트

PASCAL VOC(Visual Object Classes) 챌린지(Everingham et al.2010년) 두 가지 구성 요소로 구성됩니다. (1) 공개적으로 사용 가능한 이미지 및 주석 데이터 세트 (2) 연례 워크숍 및 대회. Microsoft COCO 데이터 세트(Lin et al.2014년)에는 328k 이미지에서 총 250만 개의 레이블이 지정된 인스턴스가 있는 4세 어린이가 쉽게 인식할 수 있는 91가지 개체 유형의 사진이 포함되어 있습니다. ADE20K(Zhou et al.2017년)는 150개의 객체 및 물건 클래스가 있는 또 다른 장면 구문 분석 벤치마크입니다. 다른 데이터 세트와 달리 ADE20K에는 객체 분할 마스크와 부품 분할 마스크가 포함됩니다.

3. 전통적인 방법

픽셀 색상, HOG(Histogram of oriented gradients)(Dalal 및 Triggs)와 같은 의미론적 분할에 다양한 기능이 사용됩니다. 이미지 시맨틱 분할의 접근 방식에는 감독되지 않은 접근 방식과 감독되는 접근 방식이 있습니다. 구체적으로 말하면, 회색 영상에서 널리 사용되는 임계값(thresholding) 방법이 간단하다. 회색 영상은 수집 장비가 일반적으로 X선 CT 스캐너 또는 MRI(Magnetic Resonance Imaging) 장비인 의료 분야에서 매우 일반적입니다. K-평균 클러스터링은 클러스터링을 위한 감독되지 않은 방법을 나타냅니다. k-means 알고리즘은 사전에 클러스터 수를 제공해야 합니다. 분할 문제는 에너지 모델로 처리할 수 있습니다. Mobahi et al.에서 구현된 압축 기반 방법에서 파생됩니다.

4. 최근 세분화된 DNN

CNN(Convolutional Neural Network)(LeCun 및 Bengio1995년) 생물학적 프로세스에서 영감을 얻은 공유 가중치 아키텍처를 사용합니다. 뉴런 간의 연결 패턴은 동물의 시각 피질의 조직을 모방합니다. 시야의 제한된 영역. 또한 공유 가중치 아키텍처 및 번역 불변 특성에 따라 시프트 불변 또는 공간 불변의 속성을 갖습니다. CNN은 뛰어난 구조로 인해 이미지 분류, 분할 및 감지에서 놀라운 결과를 얻었습니다.

4.1 완전 컨볼루션 네트워크(FCN)

이 논문(Long et al.2014년)은 이미지 분할 영역에 ANNFN을 도입한 첫 번째 작품입니다. 주요 통찰력은 완전 연결 계층을 완전 컨볼루션 계층으로 대체하는 것입니다. FCN은 VGG-Net에 구현되어 당시의 PASCAL VOC(2012년 평균 IU 62.2%에서 20% 상대적 개선)의 분할에 대한 최첨단을 달성하는 반면, 일반적인 이미지의 경우 추론에 1/5초 미만이 소요됩니다.

4.2 업샘플 방법: 보간법 대 디콘볼루션

FCN 아키텍처 외에도 의미론적 분할에는 디콘볼루션 계층도 채택됩니다. Noh et al.에서 사용된 디콘볼루션 네트워크. (2015년) 픽셀 단위 클래스 레이블을 식별하고 분할 마스크를 예측하는 디콘볼루션 및 풀링 해제 레이어로 구성됩니다. 논문의 FCN과 달리(Noh et al.2015년), 네트워크는 개별 객체 제안에 적용되어 최종 의미론적 분할을 위해 결합된 인스턴스 단위 분할을 얻습니다. 계산 효율성과 원본 이미지의 우수한 복구로 인해 업샘플 단계는 이중 선형 보간을 광범위하게 채택합니다. 디콘볼루션은 컨볼루션의 역 계산입니다. 입력 크기를 복구할 수도 있습니다. 따라서 기능 맵 크기를 원래 입력 크기로 복구하기 위해 분할에 적용할 수 있습니다

4.3 CRF 및 기타 전통적인 방법과 FCN 조인트

Deeplab의 연구에 따르면 DCNN(Deep Convolutional Neural Networks)의 최종 계층에서의 응답은 정확한 객체 분할을 위해 충분히 지역화되지 않았습니다(Chen et al.2016b). 그들은 최종 DCNN 계층에서 완전히 연결된 CRF(Conditional Random Field)를 결합하여 이러한 열악한 현지화 속성을 극복합니다. 이는 레이블 컨텍스트와 고차 관계의 혼합을 포함하여 풍부한 정보를 MarkovRandomField(MRF)에 결합하여 이미지 의미론적 분할을 해결합니다.

4.4 확장 컨볼루션

대부분의 의미론적 분할은 원래 이미지 분류 작업을 위해 고안된 CNN(Convolutional Neural Networks)의 적응을 기반으로 합니다. 그러나 이미지 시맨틱 분할 작업과 같은 조밀한 예측은 분류와 구조적으로 다릅니다. 논문(Chen et al.2016b)은 확장된 컨볼루션을 사용하여 다중 스케일 컨텍스트 정보를 체계적으로 집계하는 모듈을 제시했습니다.

4.5 백본 네트워크의 진행 상황

백본 네트워크는 네트워크의 주요 구조를 나타냅니다. 모두에게 알려진 바와 같이 시맨틱 분할에 사용되는 백본은 이미지 분류 작업에서 파생됩니다. 결과적으로 그들은 ImageNet 분류 데이터 세트에서 훨씬 더 깊은 모델보다 훨씬 우수한 성능을 보이는 잔차 네트워크의 얇은 아키텍처를 도출합니다.

4.6 분할의 피라미드 방법

연구원들은 더 강력한 백본 네트워크를 채택하는 것 외에도 피라미드 전략을 CNN에 결합하려고 시도합니다. 대표적인 것이 피라미드 방식입니다. 4가지 피라미드 방식이 있습니다.

1. 이미지 피라미드(Adelson et al.1984년)는 원하는 중지 기준에 도달할 때까지 연속적으로 다운샘플링되는 이미지 모음입니다.
2. 객체를 여러 규모로 강력하게 분할하기 위해 ASPP(Atrous Spatial Pyramid Pooling)를 제안하는데, ASPP는 여러 샘플링 속도로 필터를 사용하여 유효 시야(FOV) 및 컨볼루션 피쳐 레이어를 조사한 다음 여러 스케일에서 객체 이미지 컨텍스트를 캡처합니다
3. 피라미드 풀링은 풀링 크기의 다양한 척도를 채택한 다음 출력에 대해 원래 크기로 업샘플링 프로세스를 수행하고 마지막으로 결과를 연결하여 혼합 기능 표현을 형성합니다.
4. 기능 피라미드는 다양한 축척에서 물체를 감지하기 위한 이미지 작업의 기본 구성요소입니다.

4.7 다단계 기능 및 다단계 방법

다중 모델은 이미지 작업을 처리하는 앙상블 방식입니다(Li et al.2015년; 비올라와 존스2001년). 다단계 전략과는 별도로 의미론적 분할에는 다단계 방법이 사용됩니다. 의미론적 분할의 정확도와 속도를 향상시키기 위해 LC(deep layer cascade) 방법을 제안합니다. LC는 단일 심층 모델을 여러 하위 모델의 계단식으로 취급하고 대부분의 쉬운 영역을 분류합니다.

4.8 감독, 약한 감독 및 비 감독 방법

시맨틱 이미지 분할의 대부분의 진행은 감독 방식에서 수행됩니다

5. 결론

의미론적 이미지 분할은 이미지 처리 및 컴퓨터 비전 영역의 핵심 응용 프로그램입니다. 전통적인 의미론적 이미지 분할에 대한 간략한 검토 외에도, 이 문서는 다음과 같은 측면에서 특히 DCNN을 기반으로 한 의미론적 이미지 분할의 최근 진행 상황을 종합적으로 나열합니다. 지금까지 시맨틱 이미지 분할을 보다 정확하거나 빠르게 또는 정확도와 속도 모두에서 수행하는 방법이 점점 더 많이 등장하고 있습니다.