

리뷰 - Polygon-RNN

Share



Unknown User (8a7f808563eba3750163edfa6cb90018)

최종 수정: Jun 24, 2018

개요

원제 : Annotating Object Instances with a Polygon-RNN

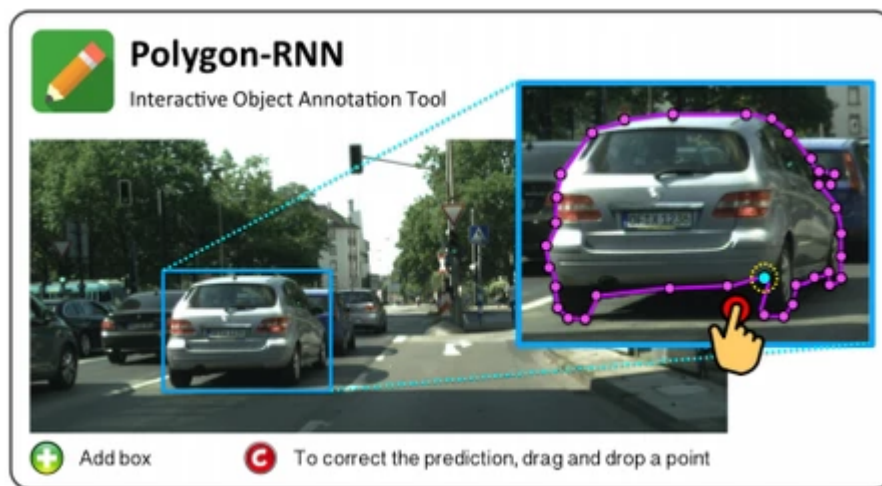
CVPR 2017

이미지 내의 물체 segmentation 데이터를 생성하기 위해,

시스템이 polygon을 제안하여, 수작업으로 찍어야 할 꼭지점의 수를 현저하게 줄일 수 있다.

논문의 목적

segmentation annotating. 이미지 세그멘테이션의 DL을 위한 데이터 만드는 것을 쉽게 하는 것.



이미지 내에서 box를 주고,

그 안에서 polygon 꼭지점(vertex)를 자동으로 제안하고

이를 수정한다.

클릭해야 하는 꼭지점 수를 줄이는 것이 목적

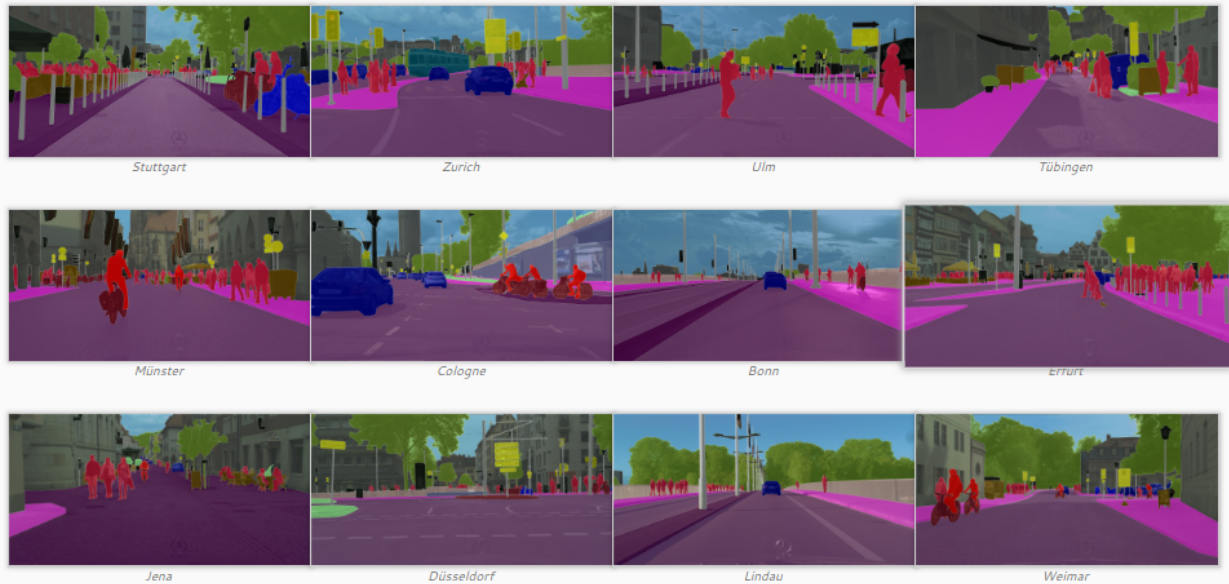
기존 방법과 차이

기존 방법은 pixel 단위의 segmentation. 요건 polygon으로 표현한 영역으로.

Cityscapes 데이터베이스

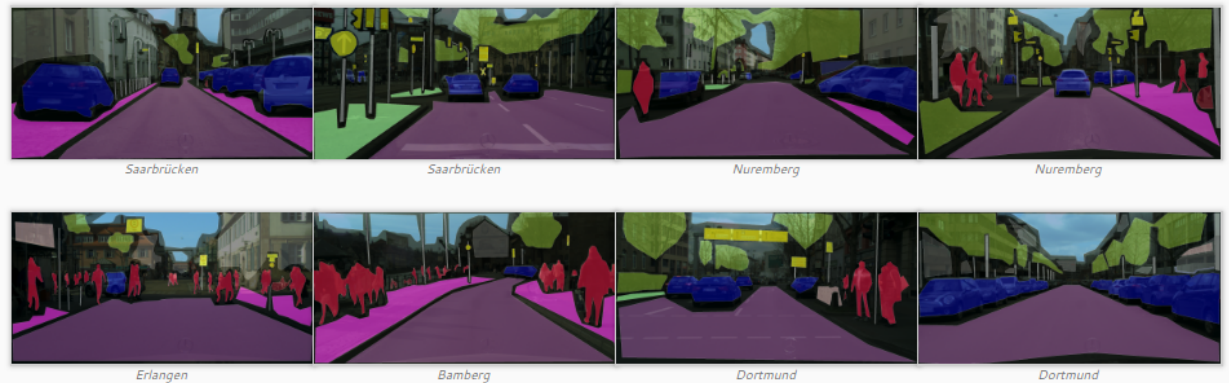
Fine annotations

Below are examples of our high quality dense pixel annotations that we provide for a volume of 5000 images. Overlaid colors encode semantic classes (see [class definitions](#)). Note that single instances of traffic participants are annotated individually.



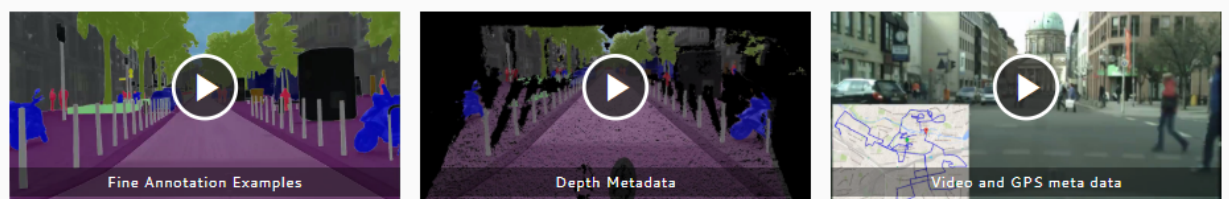
Coarse annotations

In addition to the fine annotations, we provide coarser polygonal annotations for a set of 20 000 images in collaboration with [Pallas Ludens](#). Again, overlaid colors encode the semantic classes (see [class definitions](#)). Note that we do not aim to annotated single instances, however, we marked polygons covering individual objects as such.



Videos

The videos below provide further examples of the Cityscapes Dataset. The first video contains roughly 1000 images with high quality annotations overlaid. The second video visualizes the precomputed depth maps using the corresponding right stereo views. The last video is extracted from a long video recording and visualizes the GPS positions as part of the dataset's metadata. Note that the images are blurred for privacy reasons.



Polygon

이미지에서 특정 물체의 윤곽선을 나타내기 위한 다면체.

연속된 꼭지점으로 표현한다.



모델 구조

본 논문의 구조

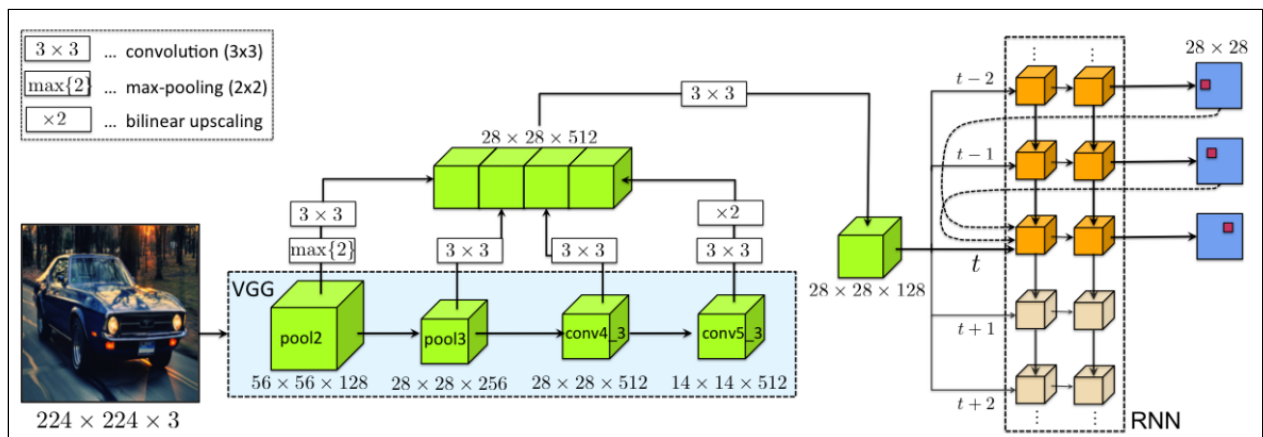


image representaion을 위한 CNN

VGG-16 기반.

skip connection 사용

RNN

ConvLSTM 기반.

t-2, t-1의 2개의 이전 출력과 CNN으로 구한 image representation을 입력으로 함.

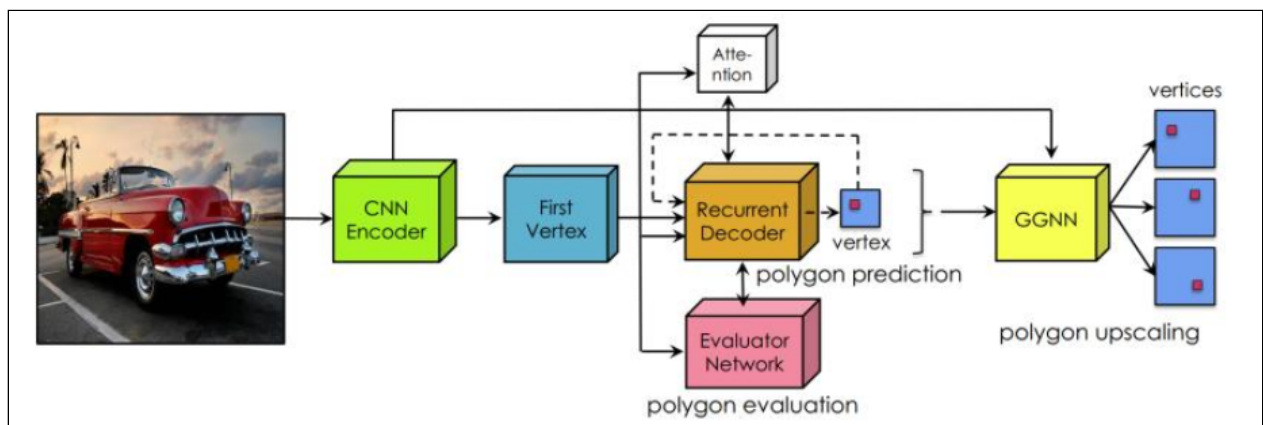
RNN의 출력은 1개의 꼭지점을 의미. 따라서 RNN입력은 이전 2개의 꼭지점과 image 특징.

궁금함

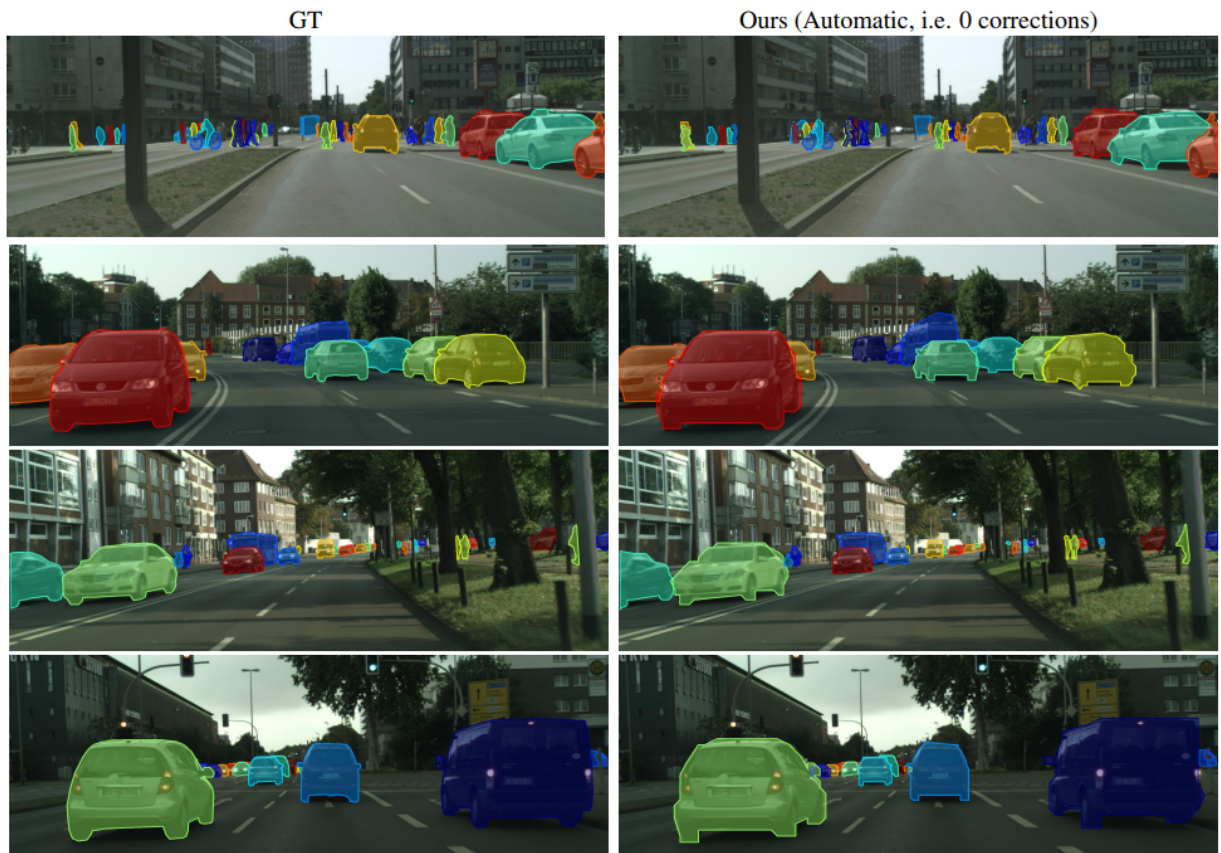
- 입력 box는 임의의 사이즈인데 어떻게 고정된 입력의 CNN으로 처리할 까.
- RNN 출력은 28x28인데 어떻게 박스 사이즈를 처리?

요 궁금점 외에 학습에 대한 자세한 사항은 설명하고 있지 않다. 코드도 오픈하고 있지 않다.

이후 논문의 구조



결과

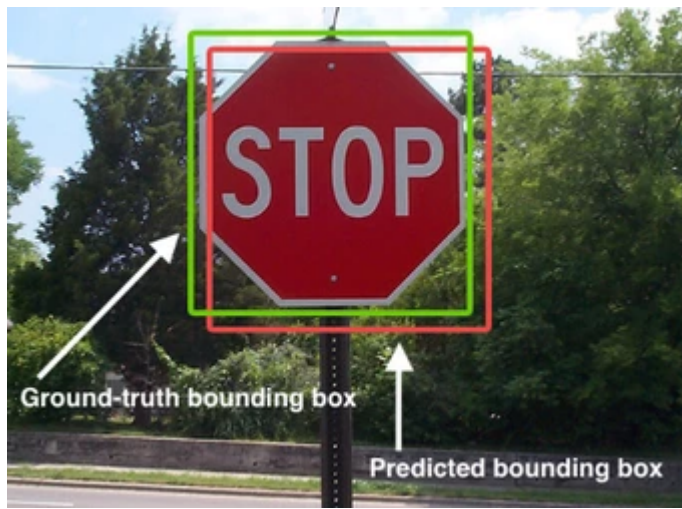


CityScapes 전체에 대해 사람이 직접 꼭지점을 찍는 것 보다 4.7배 적게 한다. IOU는 78.4%

자동차의 경우 7.3배, IOU는 82.2%

IOU

Intersection over Unison



$$\text{IoU} = \frac{\text{Area of Overlap}}{\text{Area of Union}}$$

이후 논문

CVPR 2018.

제목 : Efficient Interactive Annotation of Segmentation Datasets with Polygon-RNN++

활용 관련

github의 코드를 보면 python으로 되어 있고, matplotlib GUI 프로그램이다.

동영상을 보면 GUI로 세그멘테이션 polygon을 생성한다.

각종 물체 인식을 위한 데이터 생성에 사용할 수 있겠다.

고찰

특이한 성능 지표

segmentation task의 결과로 논문의 성능을 보이지 않고, annotation task라고 어플리케이션 관점으로 성능을 보이고 있다.

이건 segmentation 자체의 성능을 지표로 보이기 힘들어 그런거 아닌가 싶다.

실제 활용 가능한 어플리케이션

이후 논문의 결과를 보면 오직 CityScapes만으로 학습시키고, 다른 도메인에 적용을 했다. 의료 영상, 실내 물체 인식. 꽤 강력한 것 같다.

실제 사용할 정도의 완성도가 있는 것 같다.

단일 물체 segmentation

단일 물체를 대상으로 하는 박스(Bol, Box of Interest)는 사람이 정해 주어야 한다.

Faster R-CNN, YOLO와 같이 다중 물체를 대상으로 하지 않는다.

Reference

- paper : <https://arxiv.org/abs/1704.05548>
- 프로젝트 홈 : <http://www.cs.toronto.edu/polyrnn/>
- 이후 논문 Polygon-RNN++
 - paper : Efficient Interactive Annotation of Segmentation Datasets with Polygon-RNN++, CVPR2018 : <https://arxiv.org/pdf/1803.09693.pdf>
 - 학습된 모델 : <https://github.com/davidjesusacu/polyrnn-pp>
 - 설명 동영상 : <https://www.youtube.com/watch?v=evGqMnL4P3E>
- CityScapes 데이터베이스 홈 : <https://www.cityscapes-dataset.com/>

👍 좋아요 처음으로 좋아하는 사람이 되볼까요?

레이블 없음 ✎