Instance Level Object Segmentation in Videos

Course Project
Introduction to Visual Computing
Ecole CentraleSupelec, Spring 2019

Kai-Wei TSOU, Ayush K. Rai

Motivation

- Decrease Accident Rates involving Self Driving Cars
- Multiple Applications in Other Fields
  - Robotics
  - Biomedical Imaging
  - Remote Sensing

- Inspired from CVPR 2018 Workshop Autonomous Driving Video Segmentation Challenge on Kaggle

AUTOS

US safety agencies to investigate fatal Tesla crash in Florida

PUBLISHED SAT, MAR 2 2019 • 12:14 AM EST

REUTERS

KEY POINTS

- A report on Friday's crash released by the Palm Beach County Sheriff's Department did not indicate if Autopilot was engaged at the time of the crash that killed the 50-year-old Tesla Model 3 owner.
- The National Highway Traffic Safety Administration (NHTSA) and the National Transportation Safety Board (NTSB) said they are sending teams to investigate a fatal crash in Florida on Friday involving a Tesla car and a semi-trailer.
- The two agencies are investigating several crashes involving the use of Tesla's driver assistance system Autopilot including another fatal crash in California in March 2018.

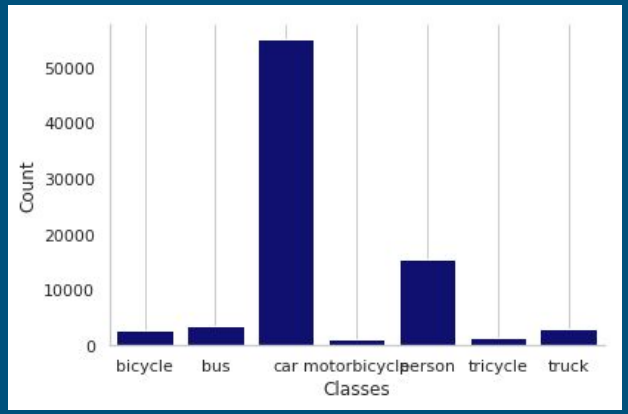
Source : https://www.cnbc.com/2019/03/02/us-safety-agencies-to-investigate-fatal-tesla-crash-in-florida.html

Task Description

DataSet Description

- Kaggle Competition Dataset : Too Large!!
- Subsample of the Competition Dataset of random 10 Videos
  - Train Set (~6k)
  - Val (~2.5k)
  - 7 Classes (car, motorcycle, bicycle, pedestrian, truck, bus and tricycle)
- Dataset : Highly Biased on Cars

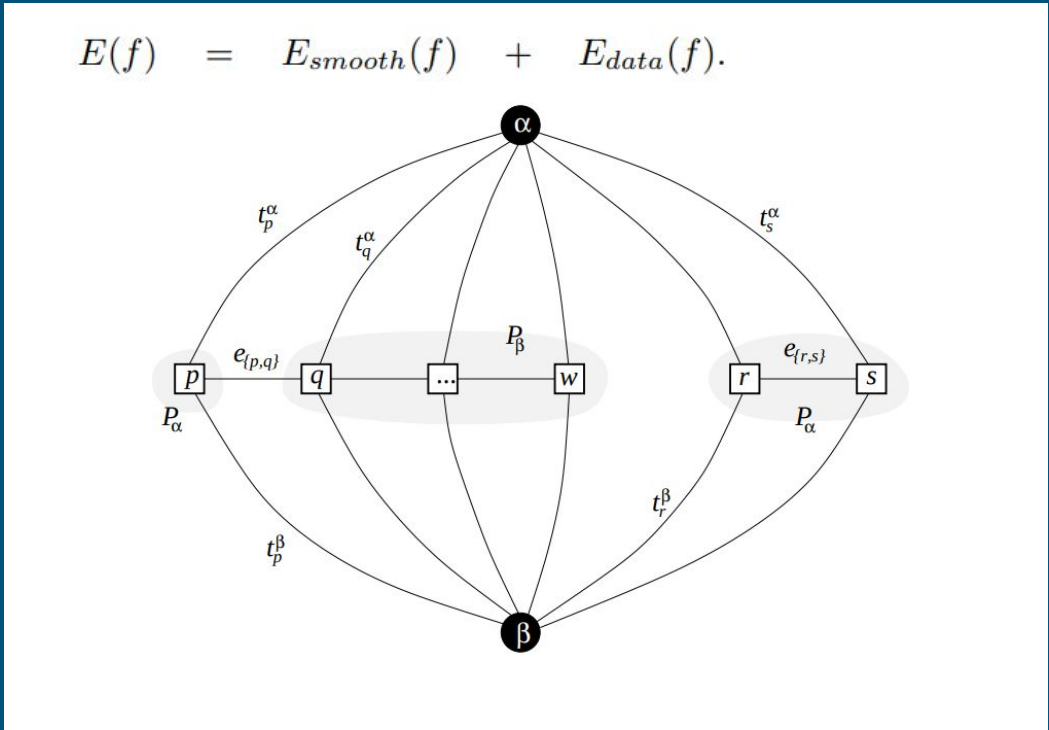
  (Rectified by adding more data from other datasets Like Berkeley Deep Drive, CityScapes or Carla Driving Simulator)



Related Work I : Energy Min via Graph Cuts

Two Variants

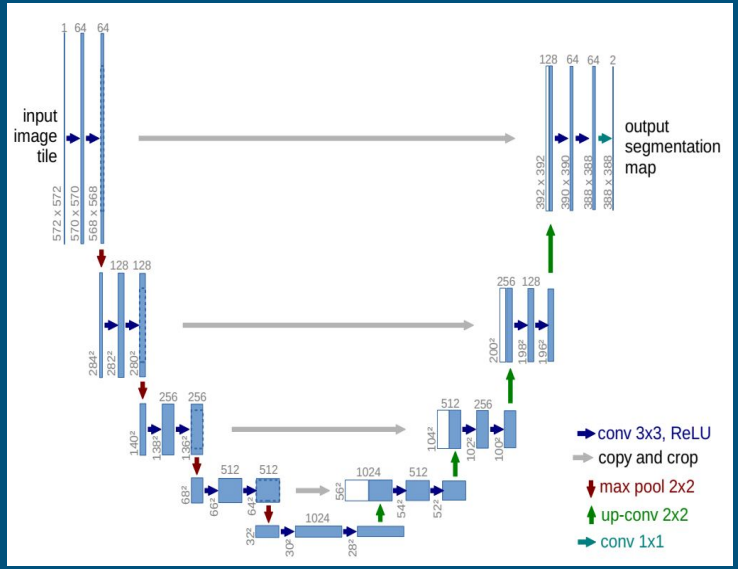
- Alpha Expansion
- Alpha Beta Swap



Fast Approximate Energy Minimization via Graph Cuts : Yuri Boykov, Olga Veksler and Ramin Zabih ICCV 1999

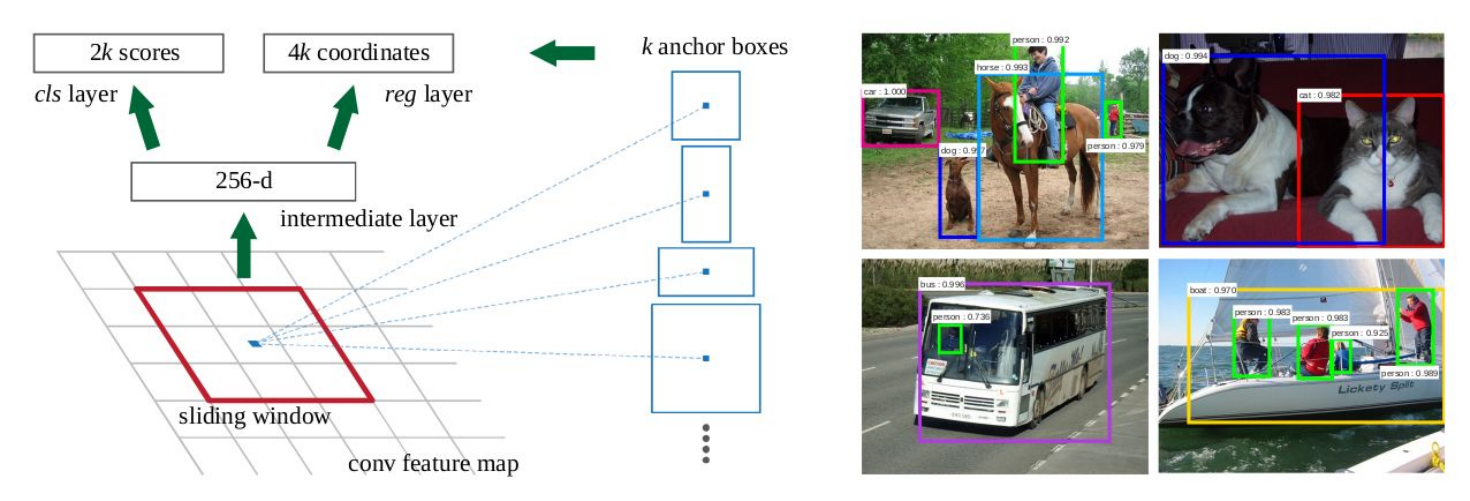
Related Work II : U-Net

- Popular Architecture for Semantic for Semantic Segmentation.
- Categorical Cross Entropy Loss
- Dice Loss
- Losses did not achieve convergence!!



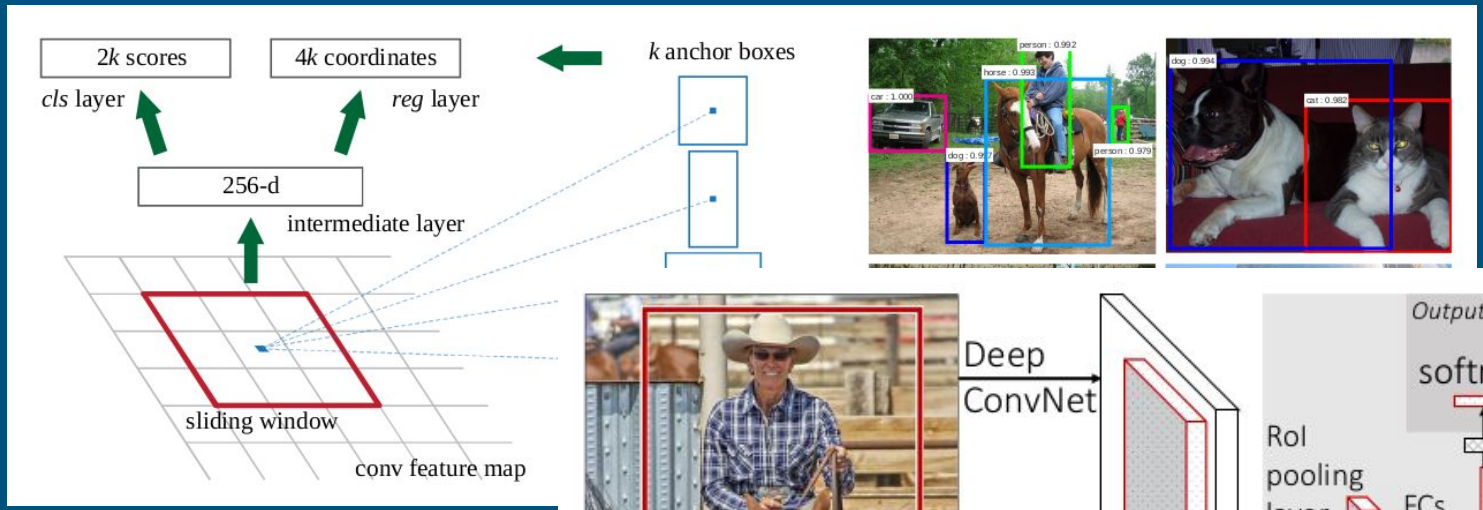
Olaf Ronneberger, Philipp Fischer, Thomas Brox UNet : Convolutional Neural Networks for Biological Image Segmentation MICCAI 2015

Adopted Approach : Mask RCNN I



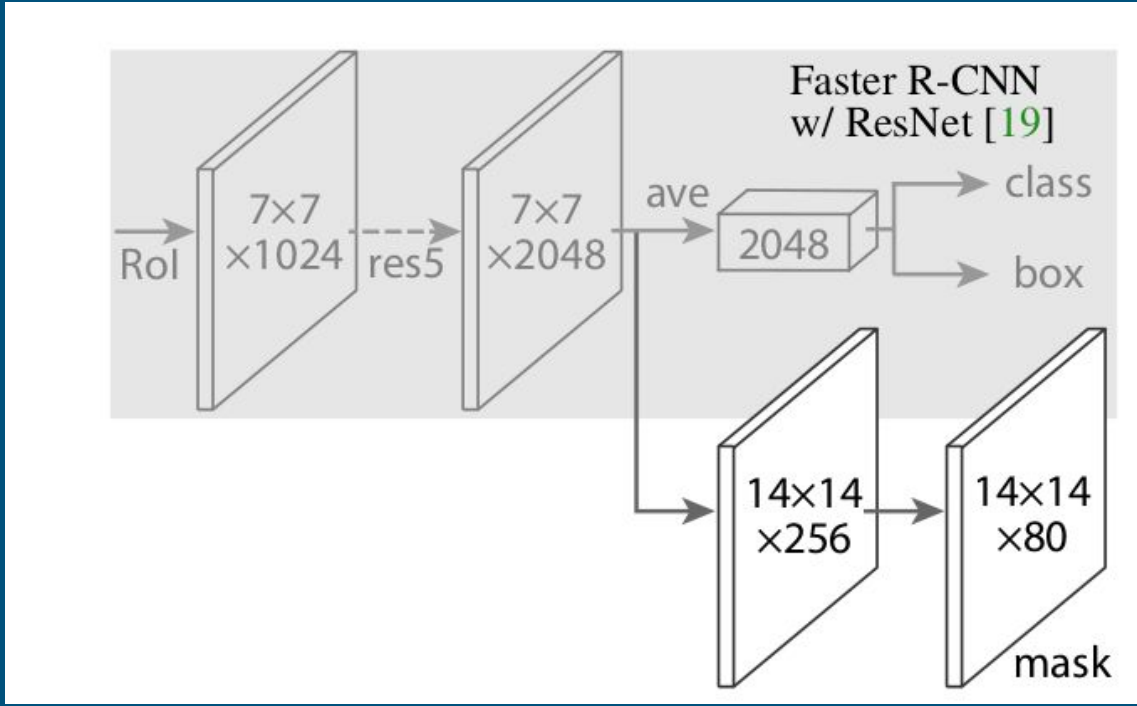
Ren, Shaoqing, et al. "Faster r-cnn: Towards real-time object detection with region proposal networks." Advances in neural information processing systems. 2015.

Adopted Approach : Mask RCNN II



Girshick, Ross. "Fast r-cnn." Proceedings of the IEEE international conference on computer vision. 2015.

Mask RCNN = Faster R-CNN + ResNet

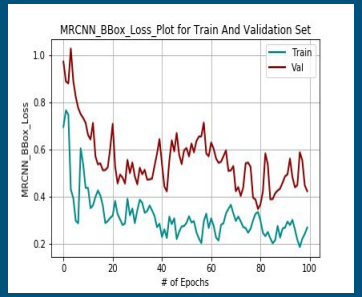


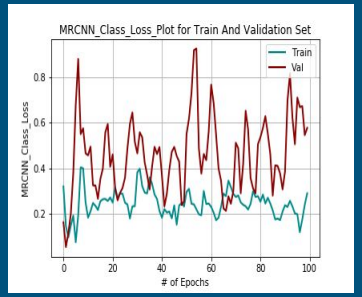
He, Kaiming, et al. "Mask r-cnn." Proceedings of the IEEE international conference on computer vision. 2017.

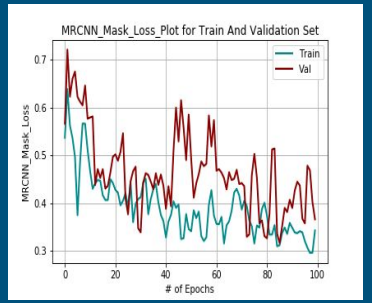
Result - Demo Video

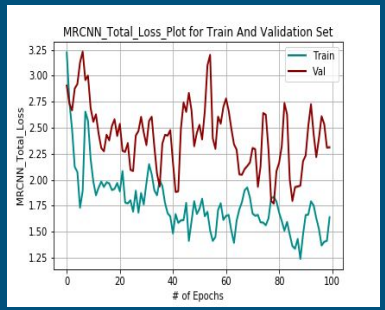
https://www.youtube.com/watch?v=nRgVNnI4-AM

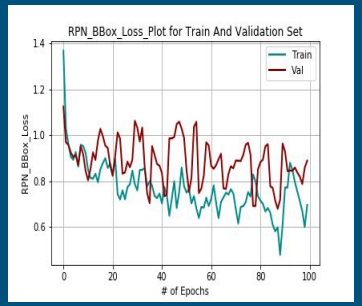
Evaluation I : Loss Plots

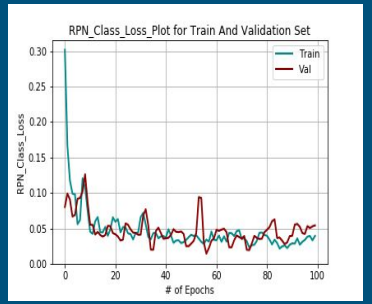












Evaluation II

Mean Intersection Over Union Score (IOU) per class (Validation Set)

Note : We only detected masks if its confidence score was >=0.9

Car	0.2899
Motorbicycle	0.0
Bicycle	0.0
Person	0.0096
Truck	0.0035
Bus	0.0132
Tricycle	0.0

Challenge : Combine Temporal Information with Mask-RCNN

- Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting
- Optimizing Video Object Detection via a Scale-Time Lattice by Chen et al. CVPR 2018
- Deep Spatio-Temporal Random Fields for Efficient Video Segmentation Chandra et al CVPR 2018

Conclusion and Take Home Message

- Instance level understanding of Images/Videos is a highly active Area of Research.
- Mask R-CNN : Meta Algorithm for Instance Segmentation
- Winner of many Kaggle Competitions on Instance Segmentation like 2018 Data Science Bowl.
- Also supports Multi-Task Learning Algorithm : Human Key Point Detection,
- Multiple feature backbone architectures can be used like ResNet/ResNext/FPNs (Feature Pyramid Networks)

Thank you