AUTONOMOUS DRIVING

GROUP MEMBERS: CHAOFU ZHANG, QIANHUI CEN, YUAN LE

Content Table

BACKGROUND

DATASET

EDA

EXPERIMENTAL SETUP

DEEP LEARNING NETWORK

RESULTS

SUMMARY AND CONCLUSION

Background

Problem: how to predict autonomous driving

To do:

build an algorithm to estimate the 6 poses of vehicles from one picture in real-world traffic environment and anticipate how a vehicle would do autonomous driving based on the analysis of the surroundings(positions of nearby automobiles)

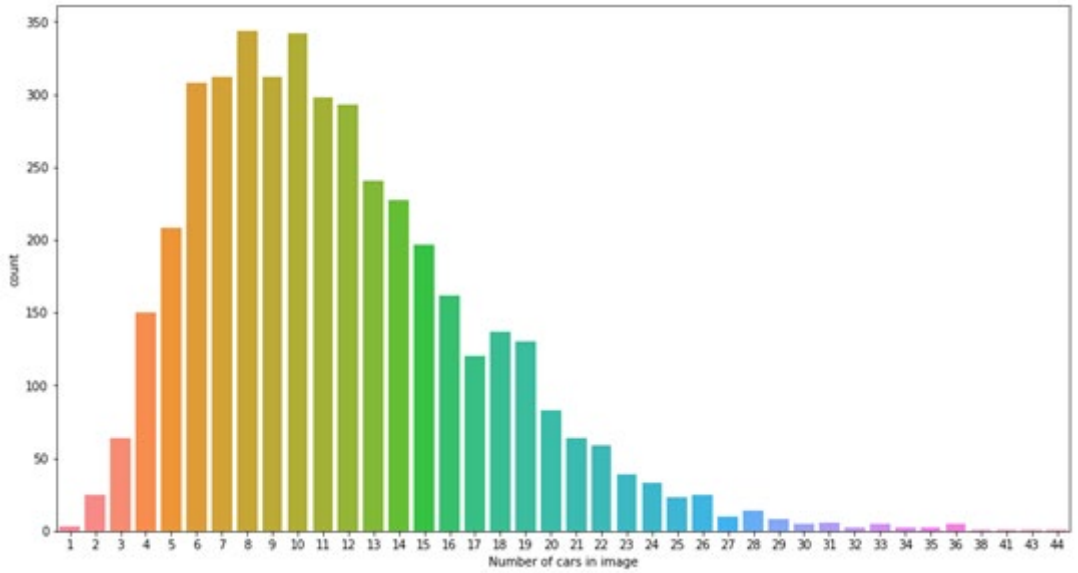
Motivation: explore the truth of autonomous driving and try to find out how to improve self-driving to make it wide-spread.

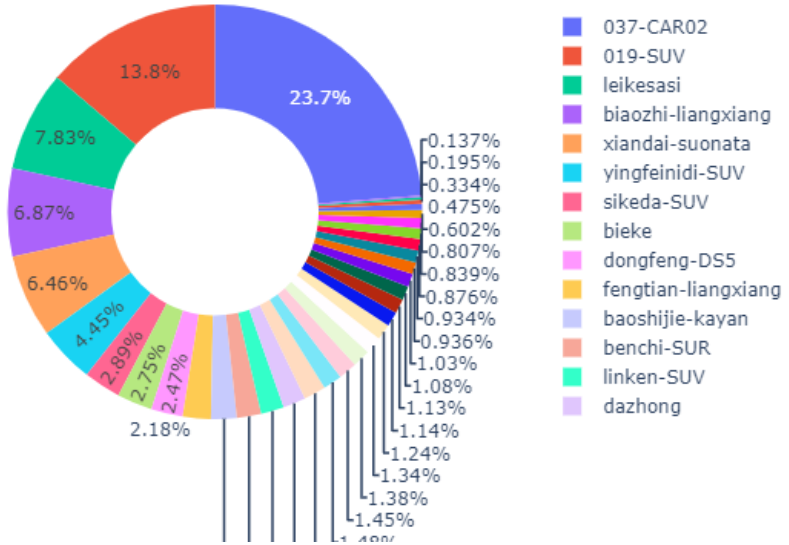
Dataset

Dataset: autonomous driving from Peking University

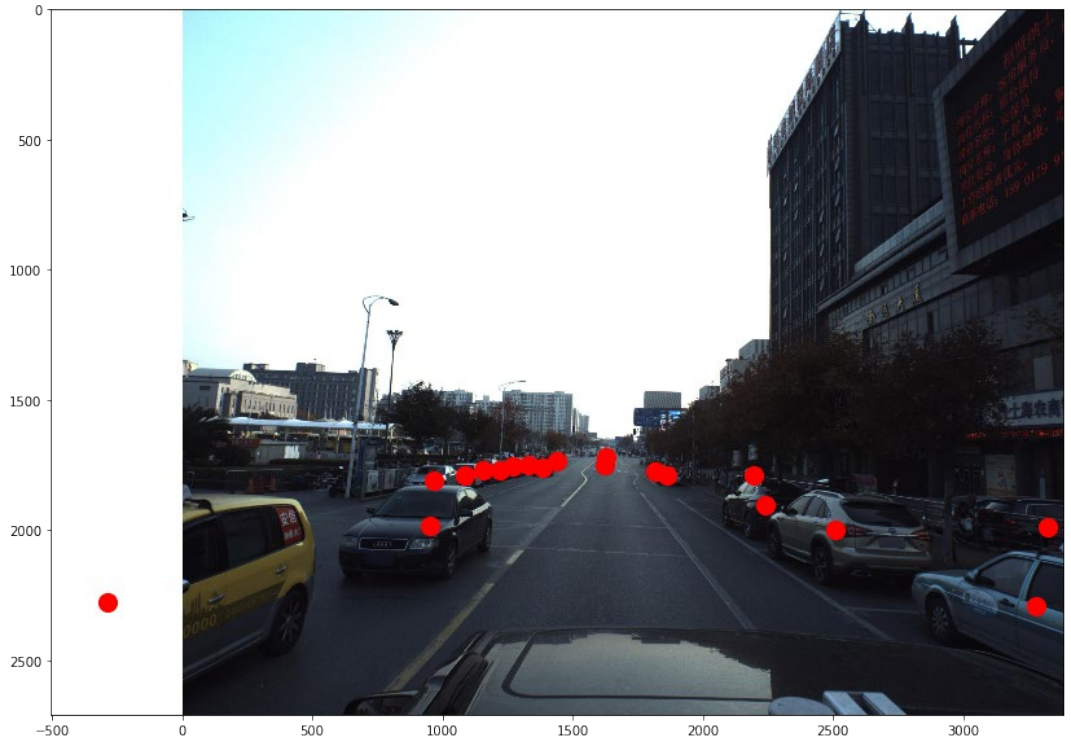
https://www.kaggle.com/c/pku-autonomous-driving

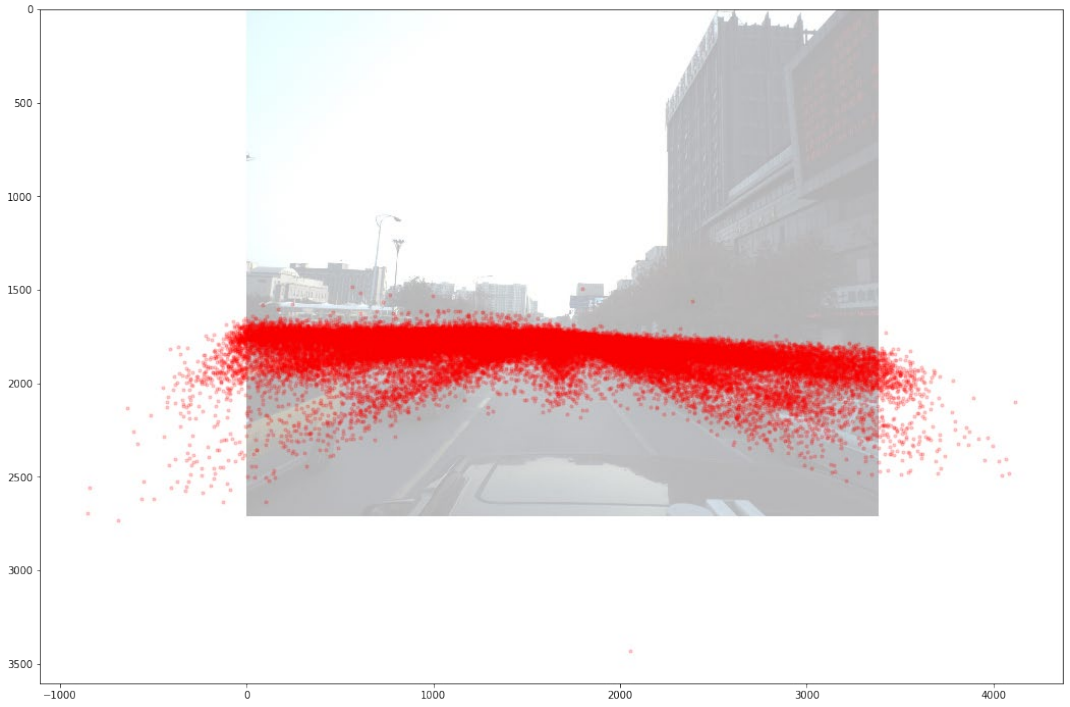
- Images:6283 real-world images (train: 4220/ test: 2063)
- Pose Information: model type, yaw, pitch, roll, x, y, z (string type)
- Image Masks: binary
- Car Models: over 60000 labeled 3D cars
- Camera Information: camera intrinsic parameters


count
Number of cars in image


23.7%
13.8%
7.83%
6.87%
6.46%
4.45%
2.89%
2.75%
2.47%
2.18%
0.137%
0.195%
0.334%
0.475%
0.602%
0.807%
0.839%
0.876%
0.934%
0.936%
1.03%
1.08%
1.13%
1.14%
1.24%
1.34%
1.38%
1.45%
037-CAR02
019-SUV
leikesasi
biaozhi-liangxiang
xiandai-suonata
yingfeinidi-SUV
sikeda-SUV
bieke
dongfeng-DS5
fengtian-liangxiang
baoshijie-kayan
benchi-SUR
linken-SUV
dazhong

EDA


0
500
1000
1500
2000
2500
−500
0
500
1000
1500
2000
2500
3000

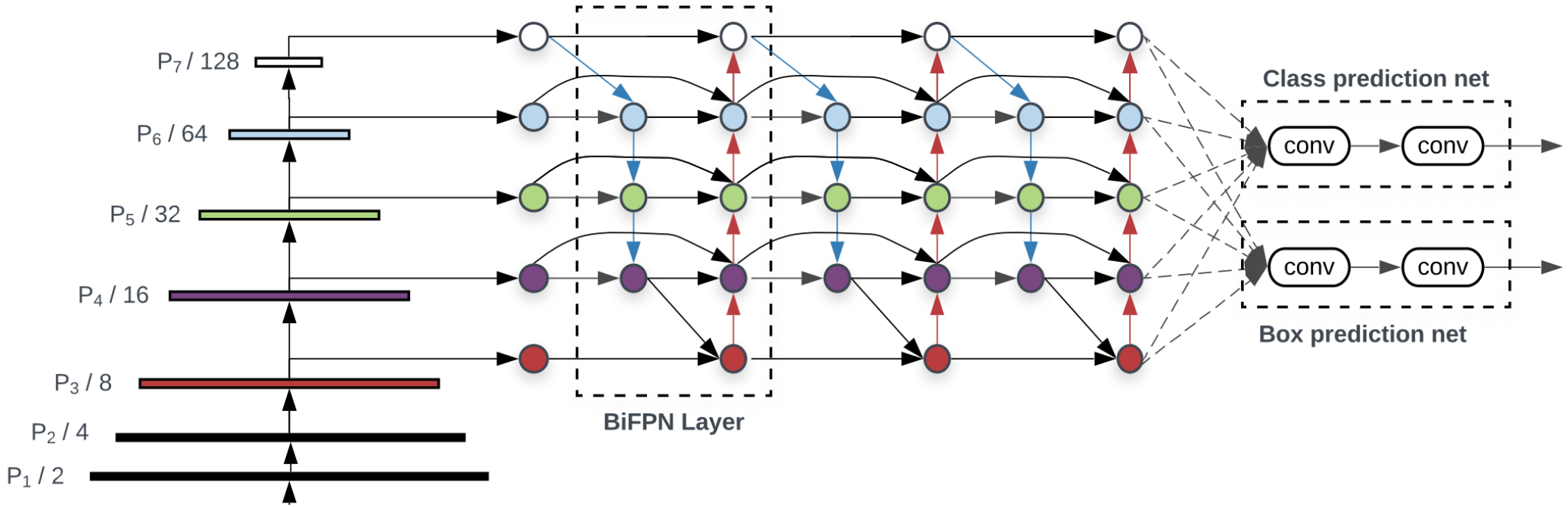

0
500
1000
1500
2000
2500
3000
3500
−1000
0
1000
2000
3000
4000

Experimental Setup

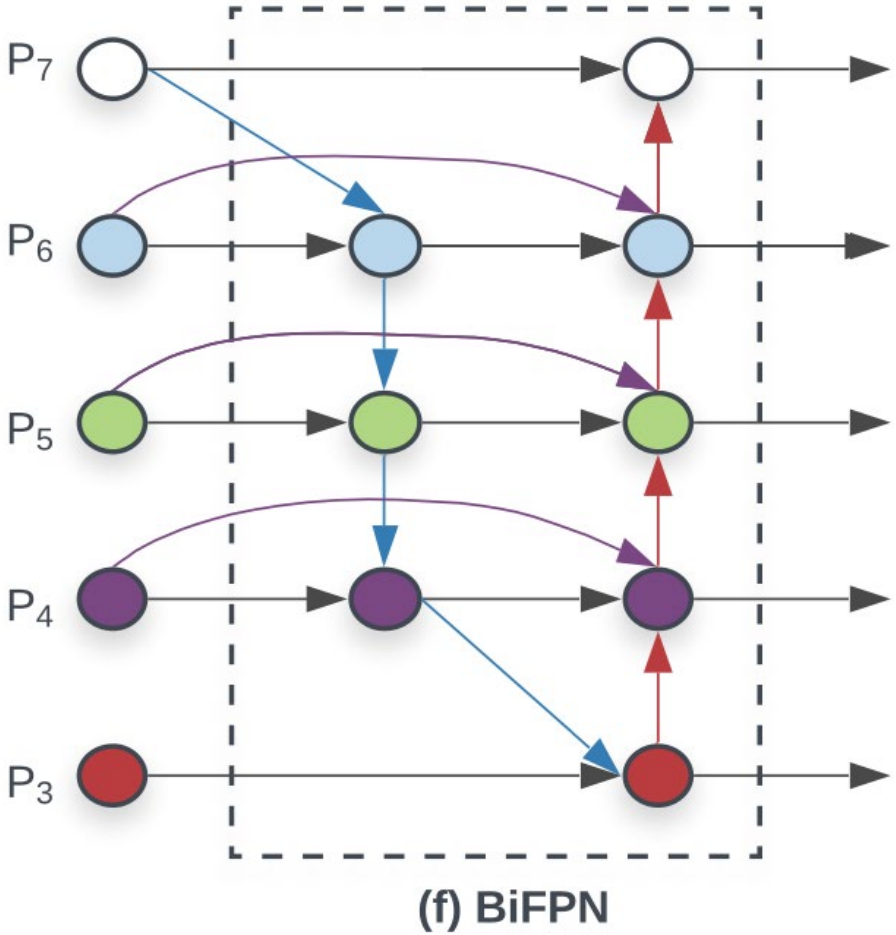
- Resize image (keep same ratio):
  - Original Size: (2710, 3384, 3) > Cut the top Half (Since in all images, the top half part is sky) then, in order to keep the same ratio, keep the height then padding with mean of the image to generate background. Left padding with of ⅛ width, right padding fill in rest width to 2048 pixels.
  - After Resize: (640, 2048, 3)
- Get Mask and regression:
  - based on the resized image, find the each car on the new resized image and its 6D regression pose.
- Hyperparameter tuning:
  - Batch Size = 4, Epoch = 10

Related Works

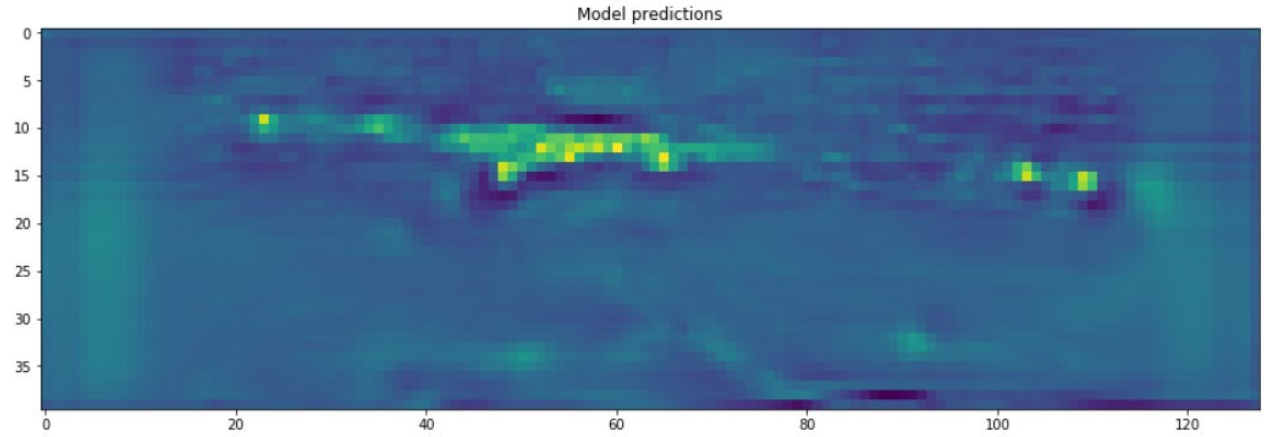
- You Only Look Once: Unified, Real-time Object Detection
  - https://arxiv.org/abs/1506.02640
- EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks
  - https://arxiv.org/abs/1905.11946
- EfficientDet: Scalable and Efficient Object Detection
  - https://arxiv.org/abs/1911.09070
- Feature Pyramid Networks for Object Detection
  - https://arxiv.org/abs/1612.03144

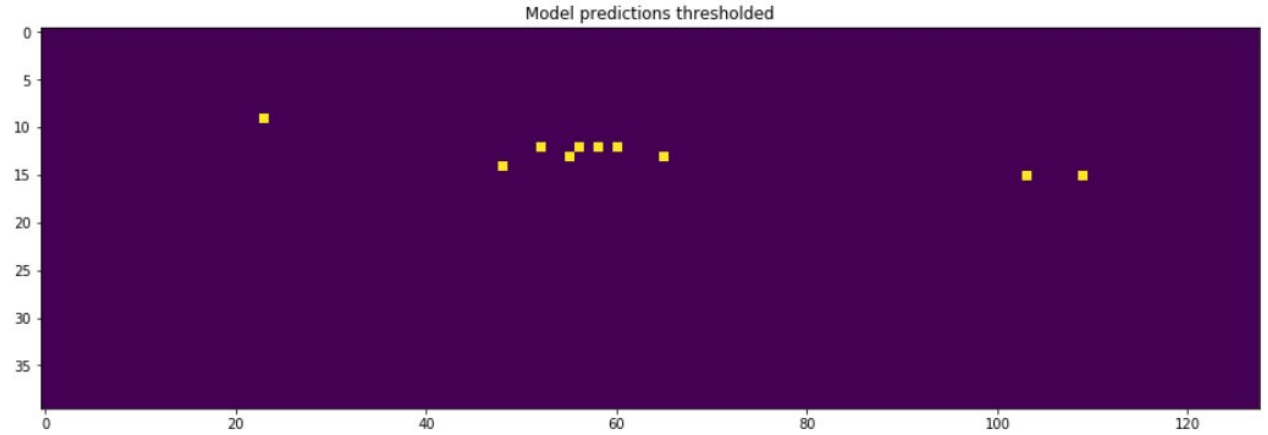


Deep Learning Network



Deep Learning Network


Model predictions
0
5
10
15
20
25
30
35
0
20
40
60
80
100
120


Model predictions thresholded
0
5
10
15
20
25
30
35
0
20
40
60
80
100
120

Results

Summary and Conclusion