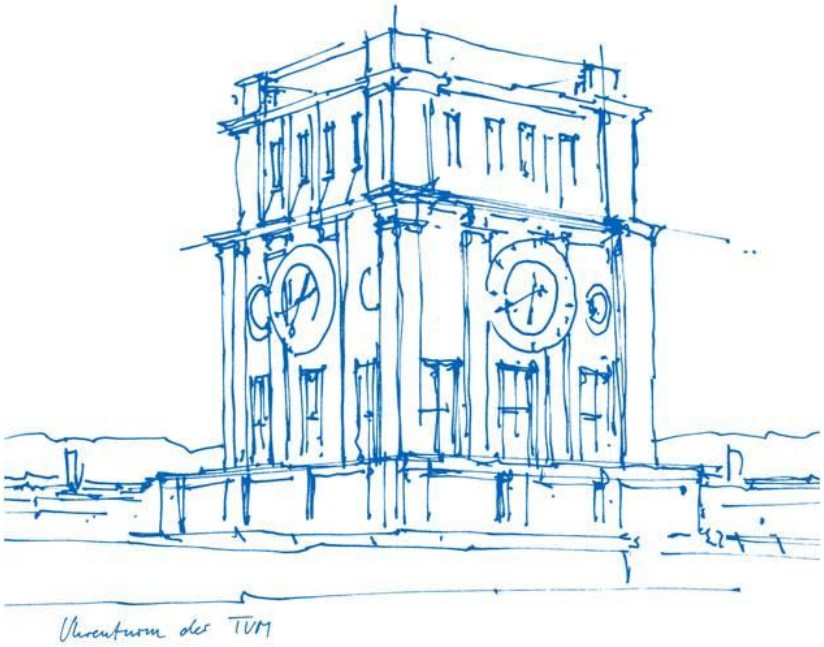
Sensor Modality Fusion

Group members: Nikhitha Radhakrishna Naik and Chenguang Huang

Technische Universität München

Faculty of Informatik

03 December 2019



Contents

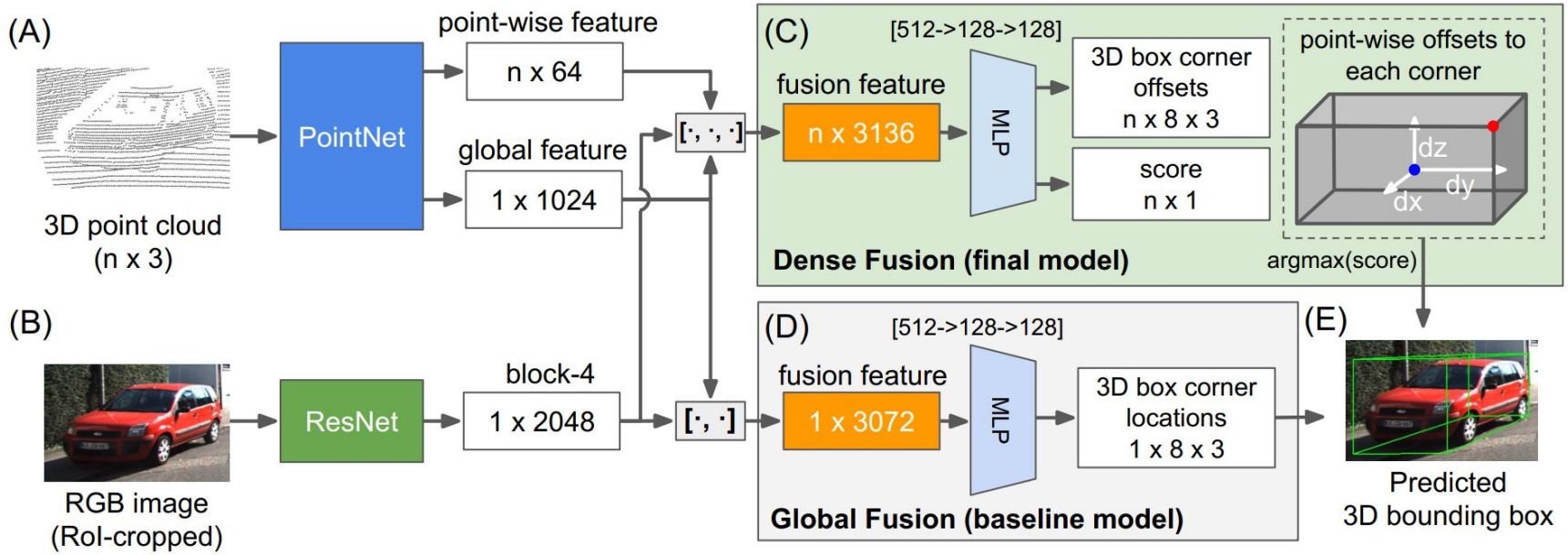
1. Recap
2. New approach
3. Trained networks and analysis
4. Plan for next week

Recap

3D Object Detection with low level sensor fusion:

PointFusion(CVPR, 2018)

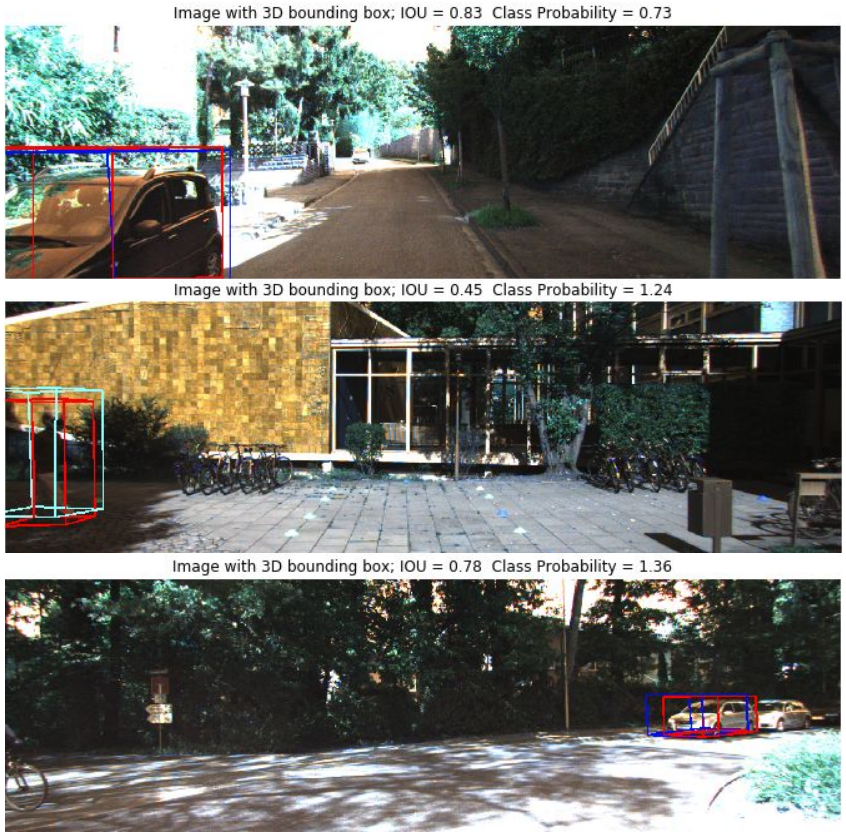
extract features of point cloud and image respectively and concatenate them.



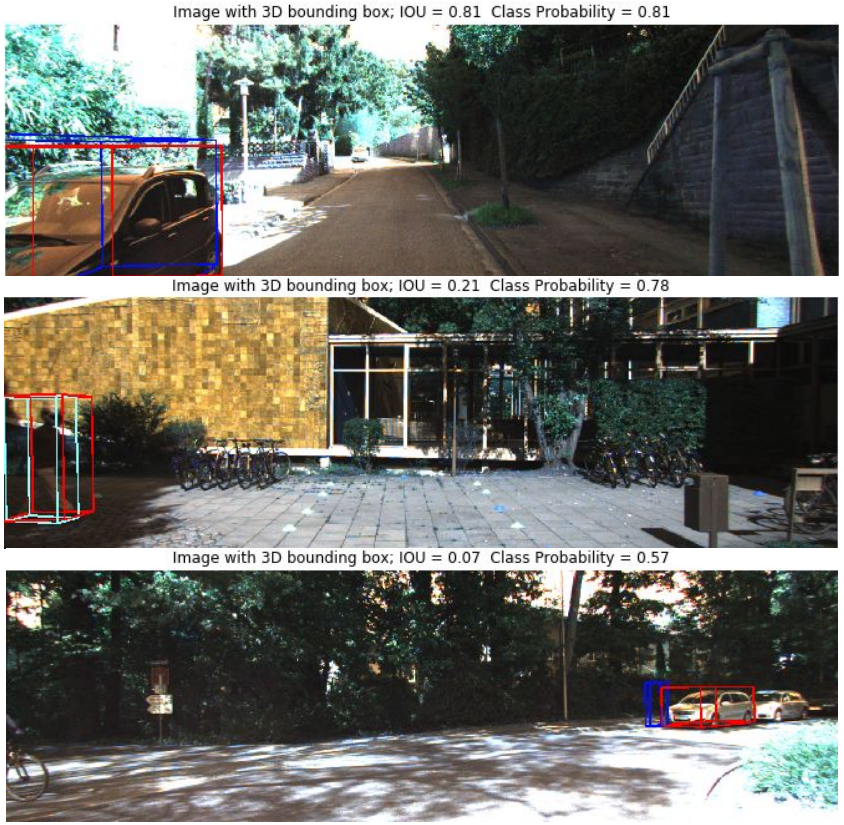
- Problem
  - Depend solely on image input and 2D detector for region proposal

Recap

Result: prediction visualization



reproduced result



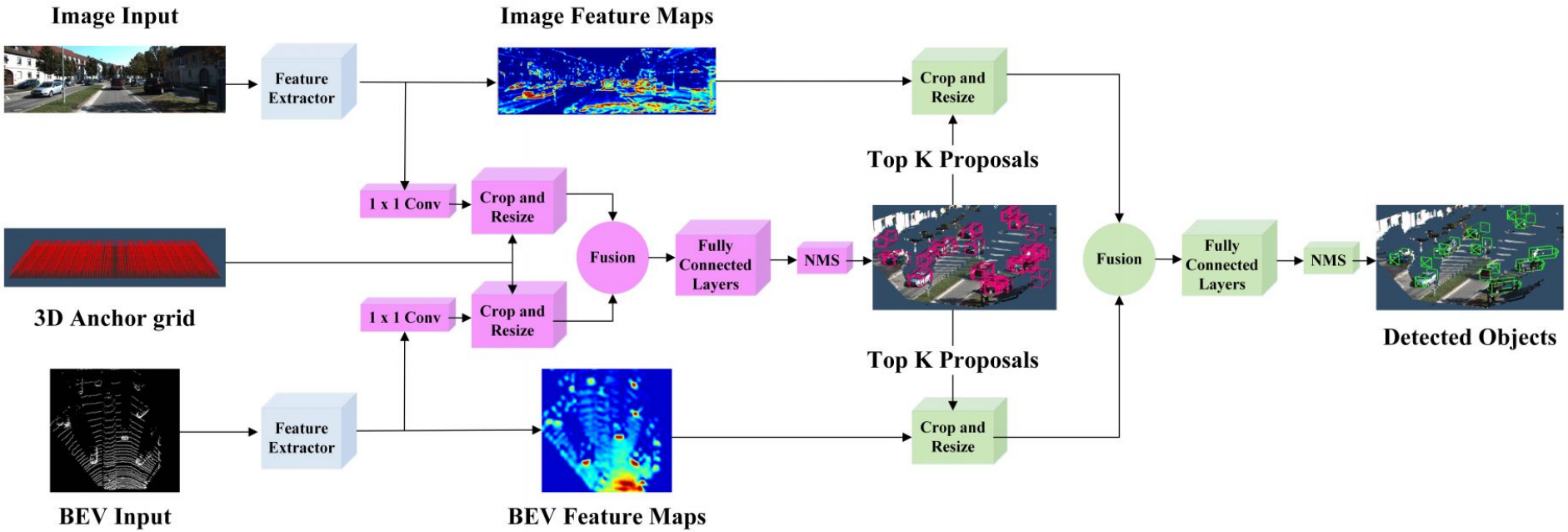
result from unofficial implementation

New approach

3D Object Detection with low level sensor fusion:

Aggregate View Object Detection

Separate region proposal and detection step, and both steps use fused data

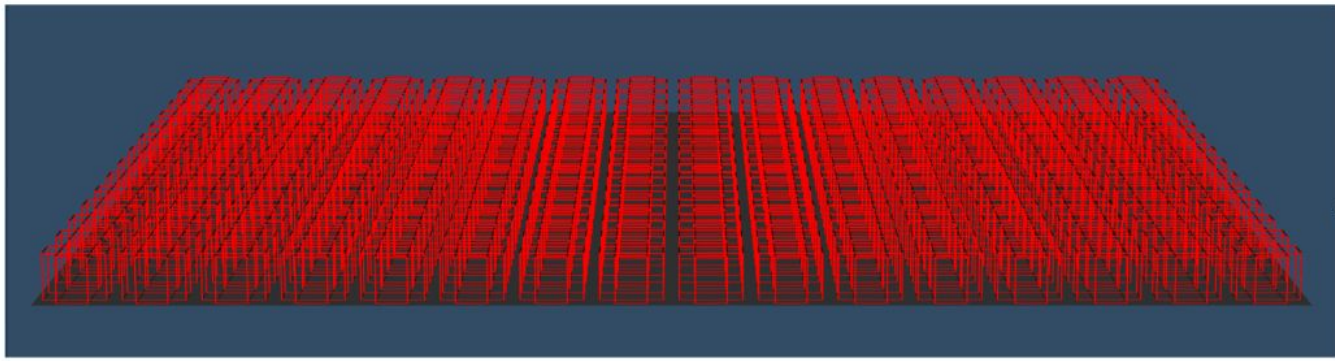


New approach

3D Object Detection with low level sensor fusion:

Aggregate View Object Detection

Anchor grid



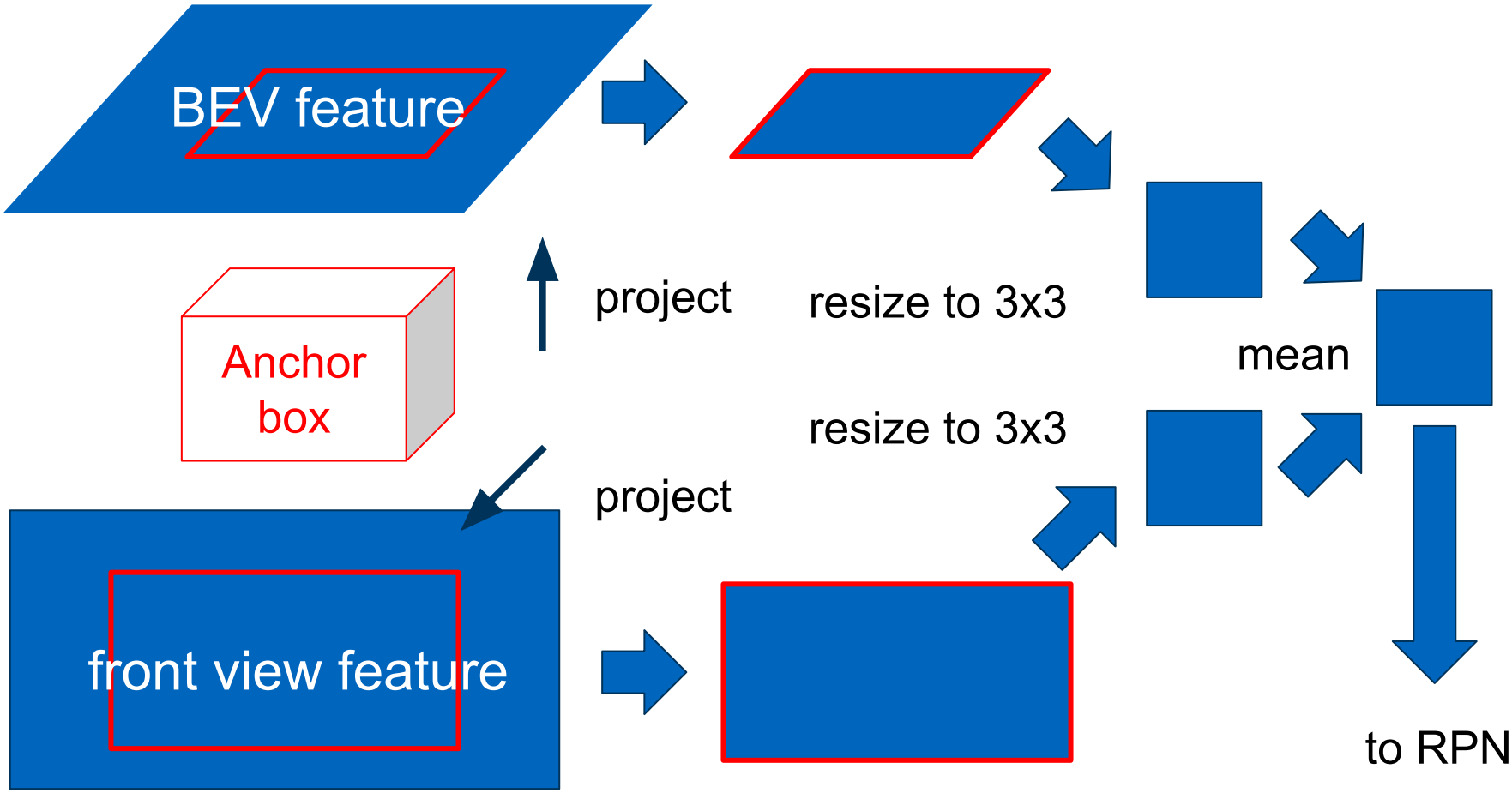
1.Create grid points in BEV
2.Create 3d boxes centering at these points and determine the sizes of the boxes by clustering

New approach

3D Object Detection with low level sensor fusion:

Aggregate View Object Detection

Fusion for region proposal



New approach

3D Object Detection with low level sensor fusion:

Aggregate View Object Detection

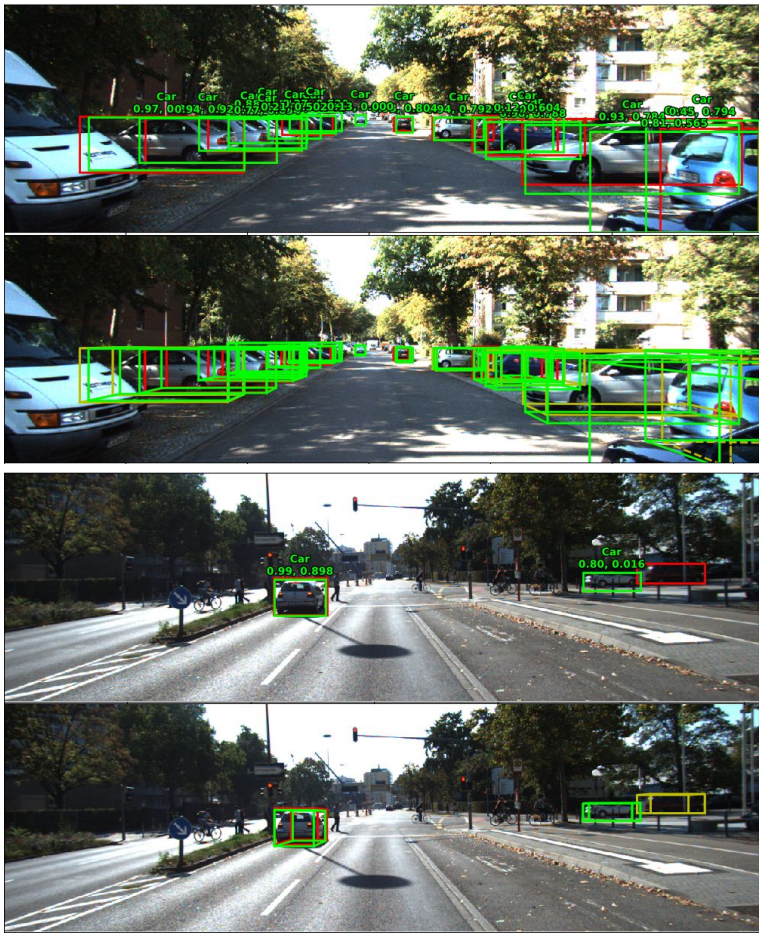
Some reproduced results

New approach

3D Object Detection with low level sensor fusion:

Aggregate View Object Detection

Some reproduced results



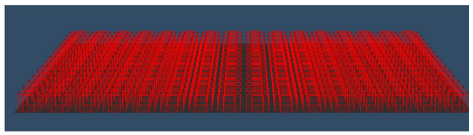
New approach

3D Object Detection with low level sensor fusion:

Aggregate View Object Detection

Analysis

- Advantage of AVOD
  - Propose regions **with fused data** rather than only using 2D detector
- Drawback
  - Only perform region proposal in selected regions(in anchor grid)
- Motivation
  - Propose regions in dense feature map



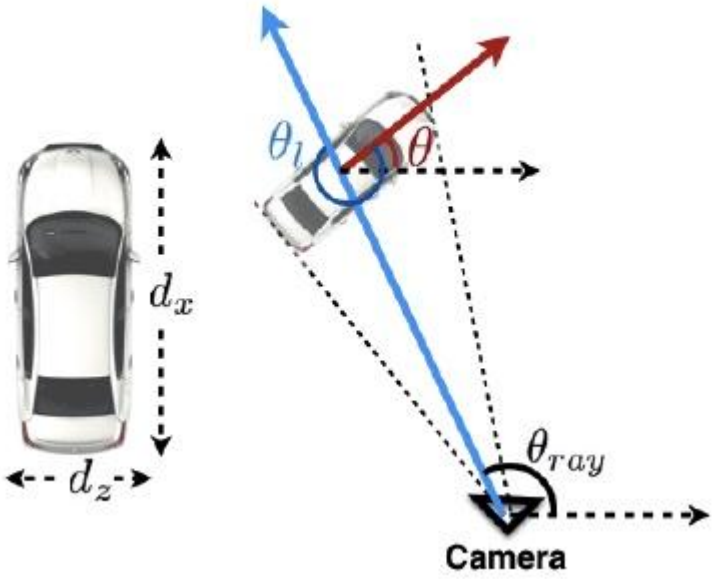
3D Anchor grid

Deep3DBox

- Built on 2D Object detection
- Performs 3D Object detection based on the 2D detections

Deep3DBox

Parameters: Dimensions, local orientation



Results from the paper

Training and Results

- Image size: 224x224
- Epochs : 10
- Used 2D detections
- Pre-trained VGG network

Plan for next few weeks

- Train image-based network longer and evaluate their results
- Implement the metrics for evaluating performance
- Improve the AVOD network
- Compare the performances from all the different networks

Thank you!