



# Dataset Bias Analysis on Autonomous Driving

Zheng Lyu

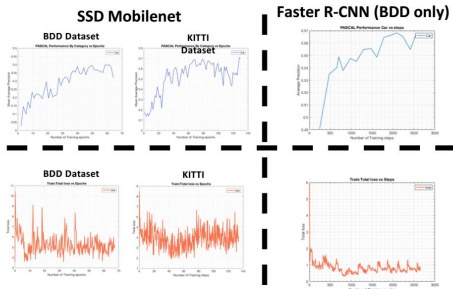
Department of Electrical Engineering, Stanford University

## Background & Motivation

- In deep learning, the bias of dataset can have critical effect on the final performance of the model.
- KITTI has been considered as a typical standard trustable dataset and is widely used.
- In this work, the robustness of the dataset is evaluated with both SSD mobilenet v1 coco (SSD) and Faster R-CNN ResNet101 coco (ResNet-101). KITTI, Apolo and new Berkeley DeepDrive (BDD) dataset s are used for analysis.



## Training Process



## Test Results

**SSD Mobilenet**

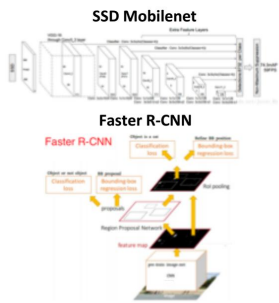
Eval	Train	KITTI 24,000 training steps			BDD 8500 images(12,000 steps)		
		mAP	AP@Car	AP@Person	mAP	AP@Car	AP@Person
	KITTI(1500)	0.517	0.697	0.337	0.312	0.505	0.1189
	BDD(1500)	0.06	0.108	0.013	0.278	0.399	0.171
	Apollo(100)	0.06	0.094	0.00001	0.213	0.417	0.01

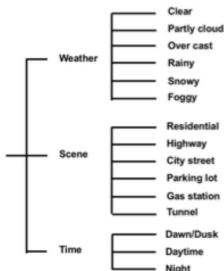
**Faster R-CNN**

Eval	Train	KITTI From google model zoo with 800,000 training steps			BDD 8000 images(8,000 steps)		
		mAP	AP@Car	AP@Person	mAP	AP@Car	AP@Person
	KITTI(300)	0.829	0.906	0.751	0.535	0.621	0.448
	BDD(300)	0.252	0.273	0.231	0.508	0.573	0.443
	Apollo(100)	0.213	0.47	0.02	0.320	0.583	0.056

## Schematic of the Models



## Bias Categories



## Examples

