

Autonomoose: Toward All-Weather Autonomous Driving

Prof. Steven Waslander

University of Toronto
University of Waterloo



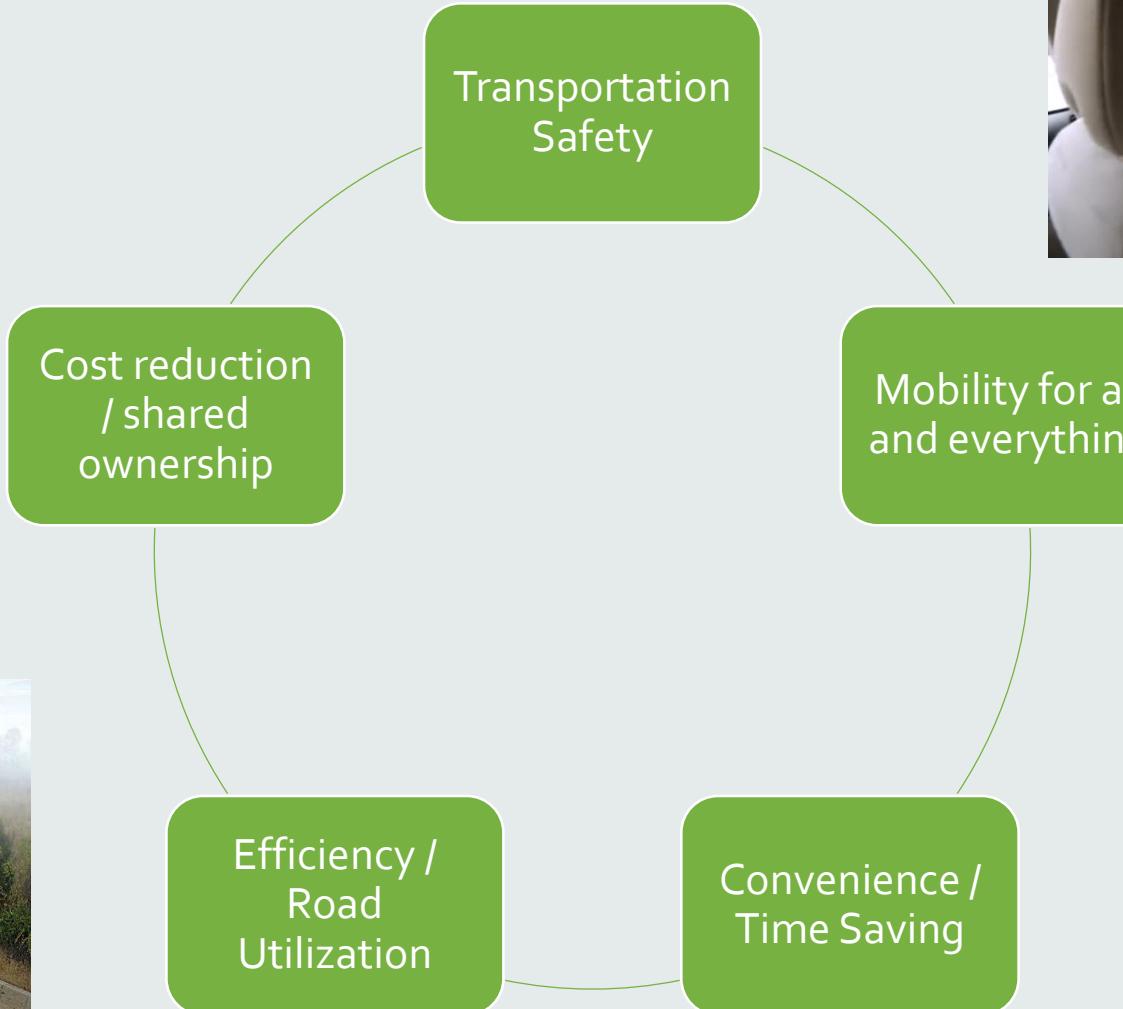
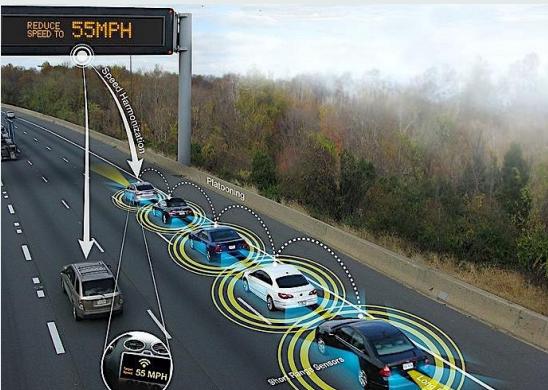
June 15th, 2018

Outline

- The State of Autonomous Driving in Canada
- Waterloo's Autonomoose Program
- 3D Object Detection with Deep Learning
- Dataset Creation for Deep Learning
- Mitigating the Effects of Canadian Weather

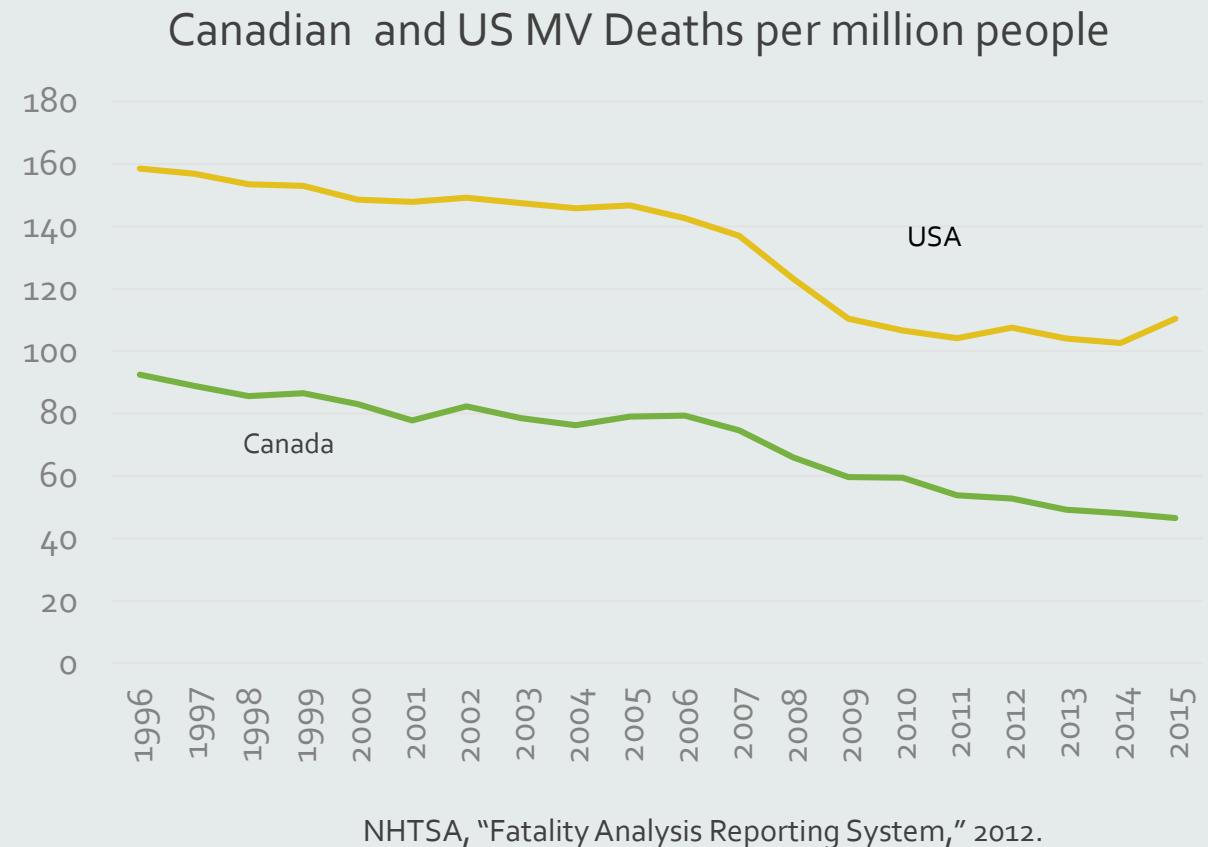


The case for autonomous driving



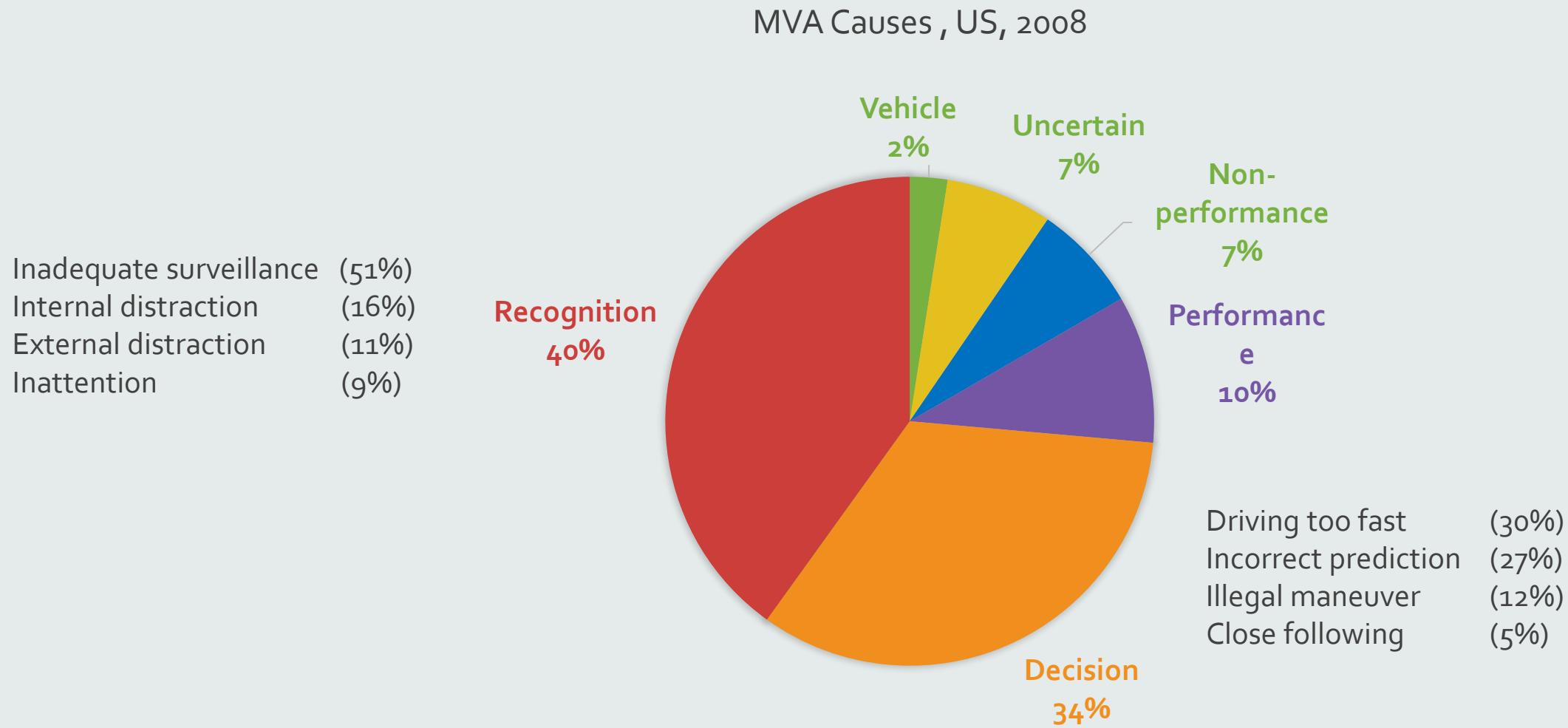
Driving is Still Dangerous

- 1,858 MV deaths in Canada in 2015
 - 46.5 deaths / million people
 - Drunk, drowsy and distracted
- 5.1 deaths / billion km in Canada 2015
 - 1 death in 198 million km
- 320 collisions / billion km in Canada 2015
 - 1 collision in 3.1 million km (reported)
 - 1 in 450,000 km (estimated, US)



Human error dominates MVA causes

- 2008 NHTSA study showed 94% of 5226 accidents were attributable to human error



Major AV R&D Investments

- Intel / BMW - \$15 billion
- Waymo - \$1.1 billion
- Ford - \$1 billion
- Toyota - \$1 billion
- Uber / Volvo - \$680 million
- GM - \$580 million
- Many more from Baidu, Aptiv, Nissan, Nvidia and others

Total Public AV investments (2014-2017)



Includes automaker and tech company investments, acquisitions, gifts, and partnerships

Source: *The Brookings Institution*

BROOKINGS

Human vs. Machine – Driving Strengths

Aspect	Human	Machine
Sensing	Excellent vision Strong inertial (inner ear)	Strong Multi-sensor Direct measurement
Perception	Excellent detection, tracking and prediction	Poor detection, tracking and prediction

Human vs. Machine – Driving Strengths

Aspect	Human	Machine
Sensing	Excellent vision Strong inertial (inner ear)	Strong Multi-sensor Direct measurement
Perception	Excellent detection, tracking and prediction	Poor detection, tracking and prediction
Information Processing	Single focus Limited attention	Excellent multiple source processing Perfect attention
Memory	Strong recall of strategies and data Flexible	Precise recall of detail Rigid
Reasoning	Intuitive, approximate Excellent at handling ambiguity Excellent recall and learning	Deductive, precise Poor at handling ambiguity Excellent fleet wide learning and recall

Brandon Schoettle, U Mich.

Human vs. Machine – Driving Strengths

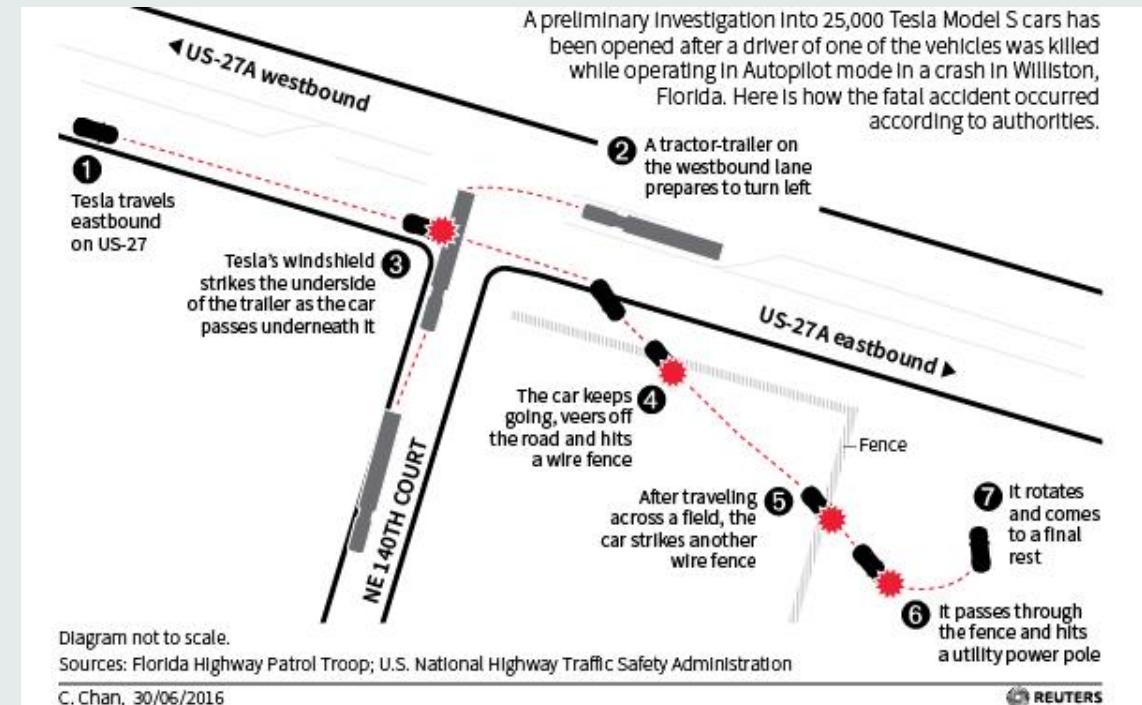
Aspect	Human	Machine
Sensing	Excellent vision Strong inertial (inner ear)	Strong Multi-sensor Direct measurement
Perception	Excellent detection, tracking and prediction	Poor detection, tracking and prediction
Information Processing	Single focus Limited attention	Excellent multiple source processing Perfect attention
Memory	Strong recall of strategies and data Flexible	Precise recall of detail Rigid
Reasoning	Intuitive, approximate Excellent at handling ambiguity Excellent recall and learning	Deductive, precise Poor at handling ambiguity Excellent fleet wide learning and recall
Reaction Speed	Slow, inconsistent	Fast, limited by computation
Command Output	Sufficient, inconsistent	Strong, Precise

Brandon Schoettle, U Mich.

Tesla Accident – Limited capability

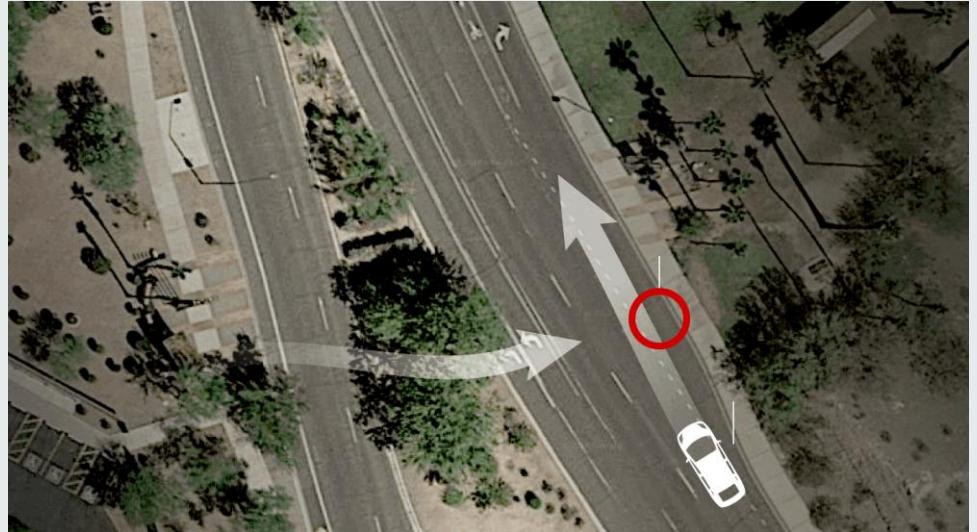
- Tesla Autopilot fatality – May 7th, 2016

- Transport truck left turn across four lane road
- Sun behind white truck not visible to camera
- Radar detected truck, but discarded as overhead sign
- Driver completely inattentive



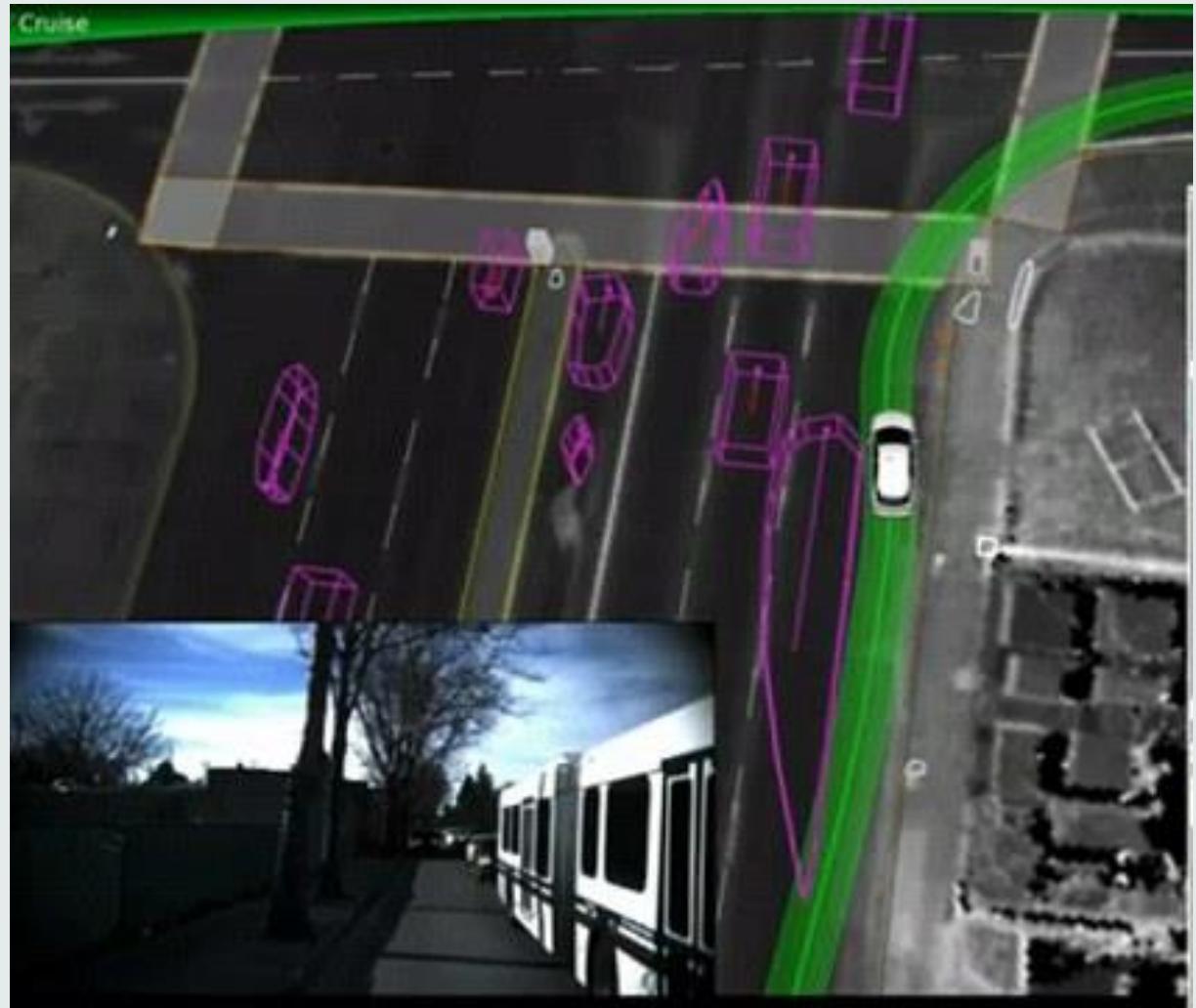
Uber Fatality

- Elaine Herzberg, 49, walking her bike across 4 lanes at night, completely crossed path of car before being hit
- Detected in lidar data at 150 m, or 6 seconds before the crash
- Visible in vision at 30-40 m or 2 seconds
- Volvo emergency braking system alert 1.3 seconds out (20 m)
 - To avoid erratic driving on false positives, Uber disabled Volvo's emergency brake
 - Warning message posted to driver on console in dash
 - Driver was supposed to monitor messages, and brake if needed.



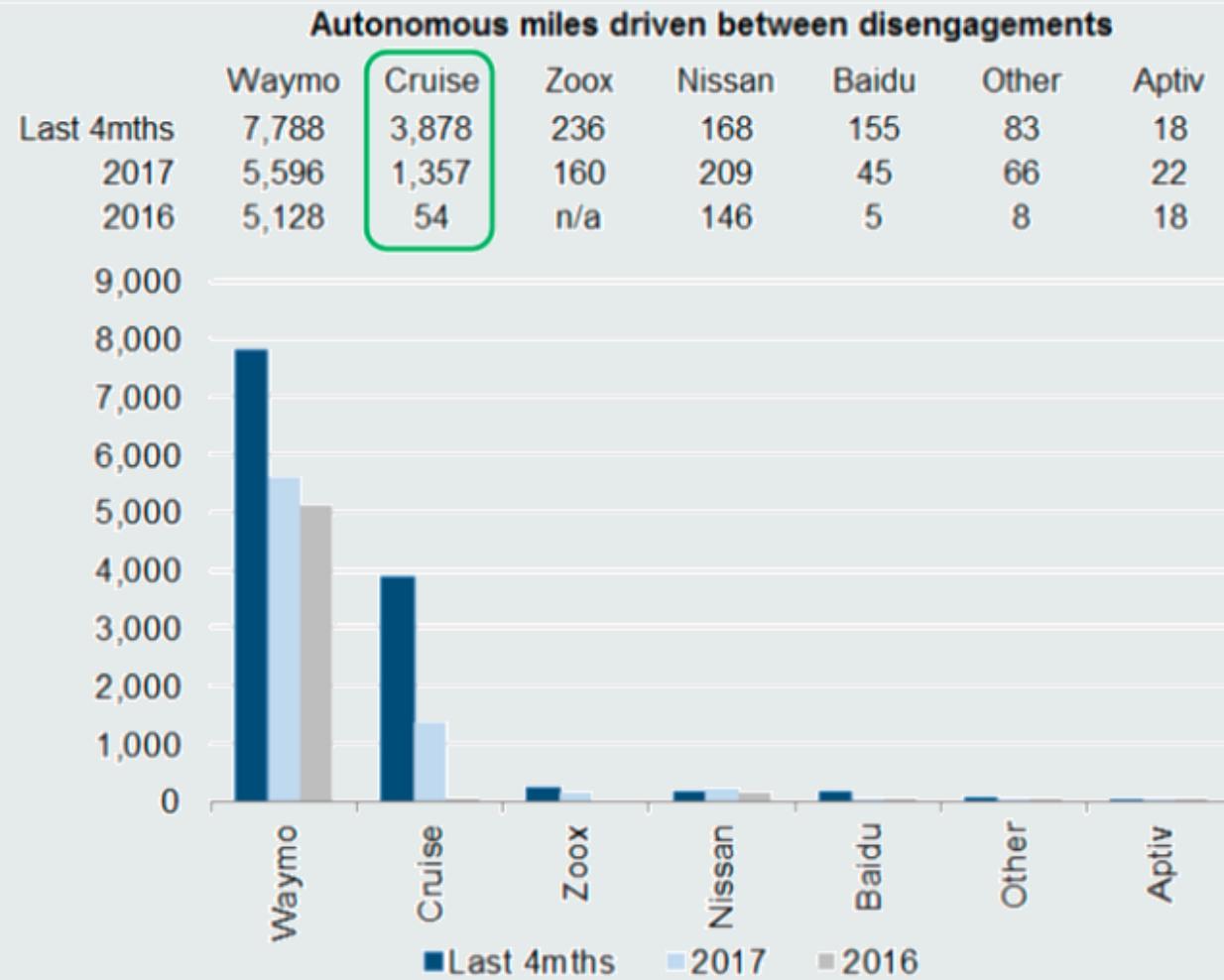
Google car (now Waymo) accident - driver behavior prediction

- Google car hits a bus – Feb 14, 2016
 - Google car had to maneuver around temporary sign
 - Google car predicted the bus would stop
 - Bus squeezed through a small gap
 - Driver unable to react in time



Autonomous Driving Disengagement Rates

- California requires annual public reporting on autonomy testing (Dec 1, 2016-Nov 30, 2017)
 - Miles driven
 - Disengagements and causes
- State-of-the-art: Waymo
 - 1 disengagement every 12,500 km
 - 1 collision per 157,000 km
 - All collisions reported
 - No at fault collisions



Autonomous Driving in Ontario



MINISTRY OF TRANSPORTATION

HOME | ABOUT THE MINISTRY | NEWS | F

Vehicles

- ▶ Vehicles Home
- ▶ Vehicle Registration
- ▶ Cycling in Ontario
- ▶ Insure a Vehicle
- ▶ Accessible Parking Permit
- ▶ Road-Building Machines
- ▶ Low Speed Vehicles
- ▶ Electric Vehicles
- ▶ Automated Vehicles
- ▶ Three-Wheeled Vehicles
- ▶ Buying or Selling a Used Vehicle
- ▶ Personalized Plates
- ▶ Replacing Plates and Stickers

Automated Vehicles - Driving

Beginning January 1, 2016, Ontario is launching a pilot project to test automated vehicles on Ontario's roads.

What is an automated vehicle?

AVs are driverless or self-driving vehicles that use sensors and artificial intelligence to navigate roads.

Why is Ontario pilot testing automated vehicles?

AVs have the potential to deliver environmental benefits such as decreased GHG emissions tied to enhanced fuel efficiency; reduced traffic congestion; increased mobility for goods and services; minimized driver employment through a made-in-Ontario sector.

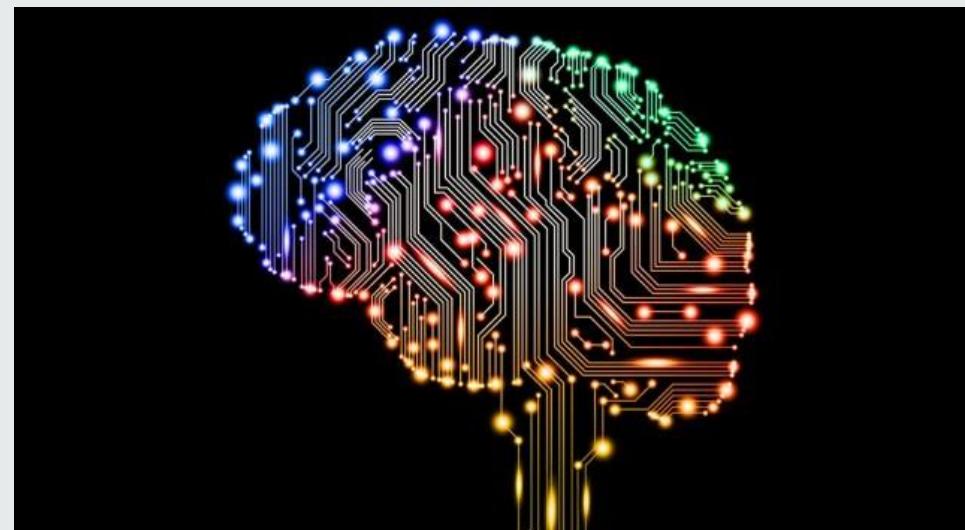
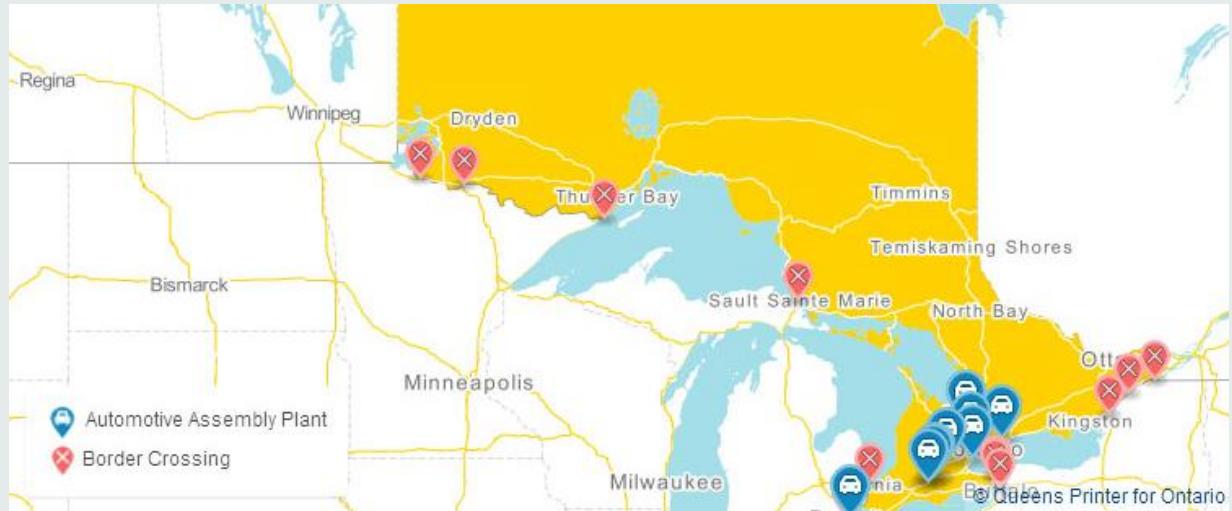
About the pilot:

- This pilot is restricted to testing on designated roads.
- The pilot will run for 10 years and will be overseen by the Ministry of Transportation.
- Only vehicles manufactured and sold in Ontario will be eligible.
- The driver must remain in the vehicle during the test period.
- The driver must hold a full class 5 driver's license.
- Eligible participants must have a valid driver's license and insurance coverage.
- All current Highway Traffic Act regulations apply.



Canada's Role in AVs Going Forward

- AVs are leading robotic AI charge
 - Essential feature, if cost barriers are resolved
 - Penetration will be rapid and ubiquitous
 - Fleets, rideshare first
 - 15-30 year transition for bulk of the fleet
 - Transformational impact on society
- Ontario Automotive Corridor
 - Long history, large investment
 - Manufacturing and now more R&D
- Birthplace of Modern AI
 - Major well-funded initiative a bold strategic move
 - Critical to AV success in perception, prediction
- Huge demand for talent, innovation
 - Strong public education system
 - Should aim to be the smartest country on earth



Biggest Requirements for Canadian AV Industry

- Legislative clarity for deployment
 - National standards, regulations must be set now
 - Transport Canada & Innovation, Science and ED
 - Legality and liability must be defined
 - Audi leaving features off in Canada for now
- Safety assessment and test program
 - Publicly verifiable performance
 - Expanded concept of safety rating
 - Operational domain specification
- Research into Canadian Challenges
 - Adverse weather
 - Remote operation
 - Driving culture
 - Wildlife

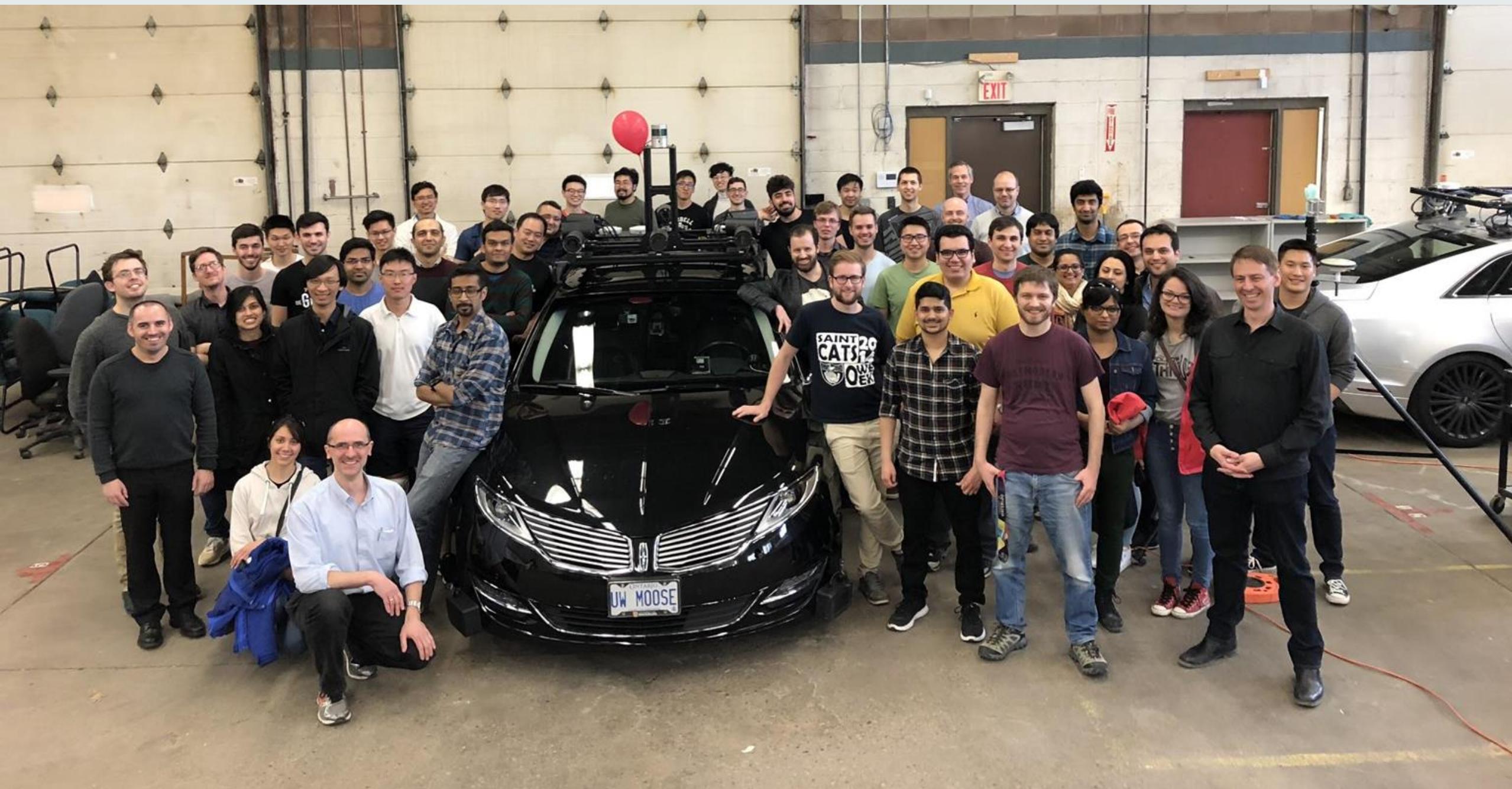


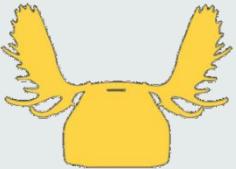
Outline

- The State of Autonomous Driving in Canada
- Waterloo's Autonomoose Program
- 3D Object Detection with Deep Learning
- Dataset Creation for Deep Learning
- Mitigating the Effects of Canadian Weather



The Autonomoose Team



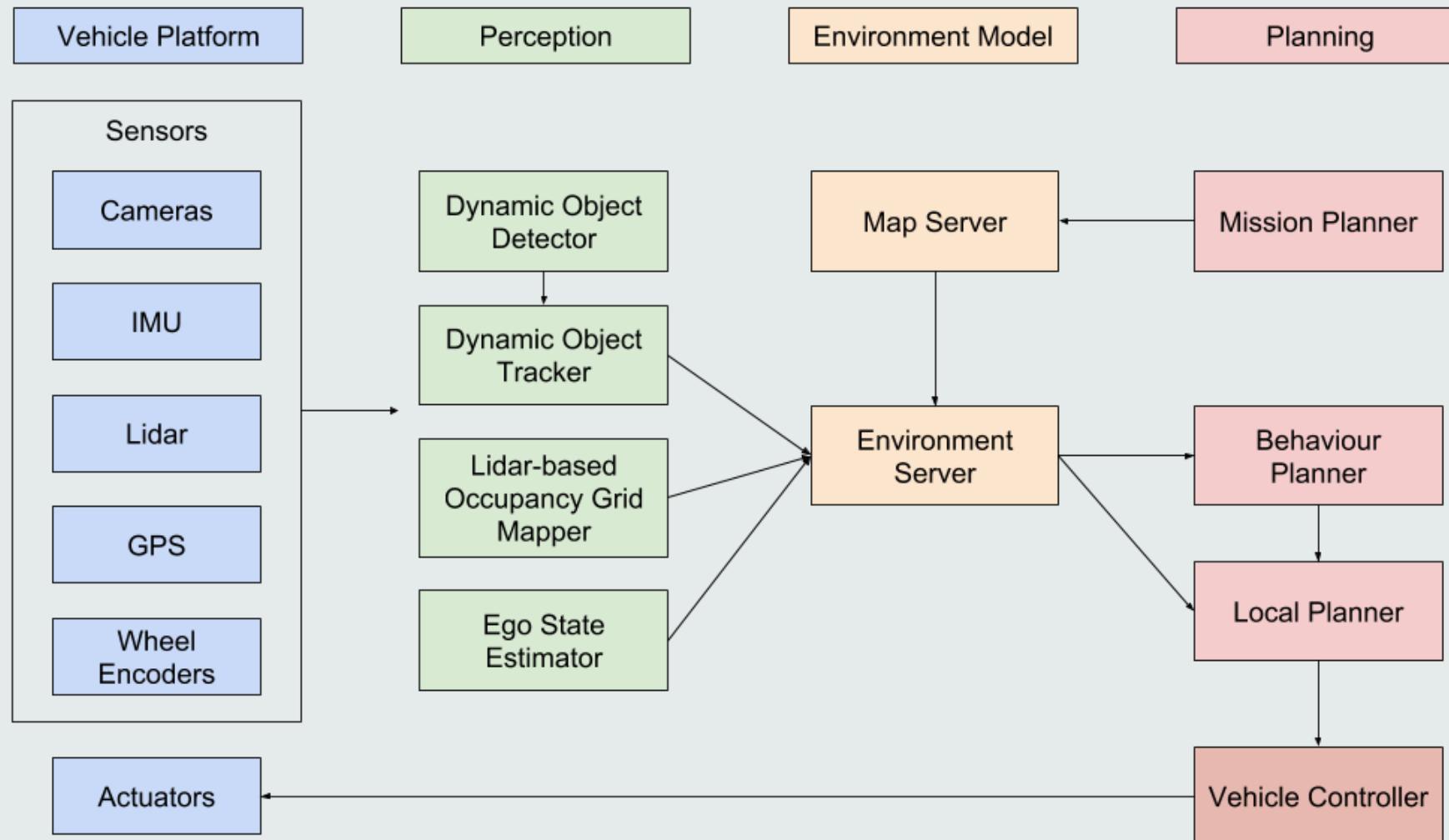


The Autonomoose: Waterloo's AV Program

- Three core faculty
 - Myself, MME
 - Krzysztof Czarnecki , ECE
 - Sebastian Fischmeister, ECE
- Over 80 students and engineers
- Under Waterloo Centre for Automotive Research
 - Ross McKensie, Managing Director
 - 15 faculty in AV research
- Vehicle arrived June 2016



Baseline Autonomous Driving Architecture

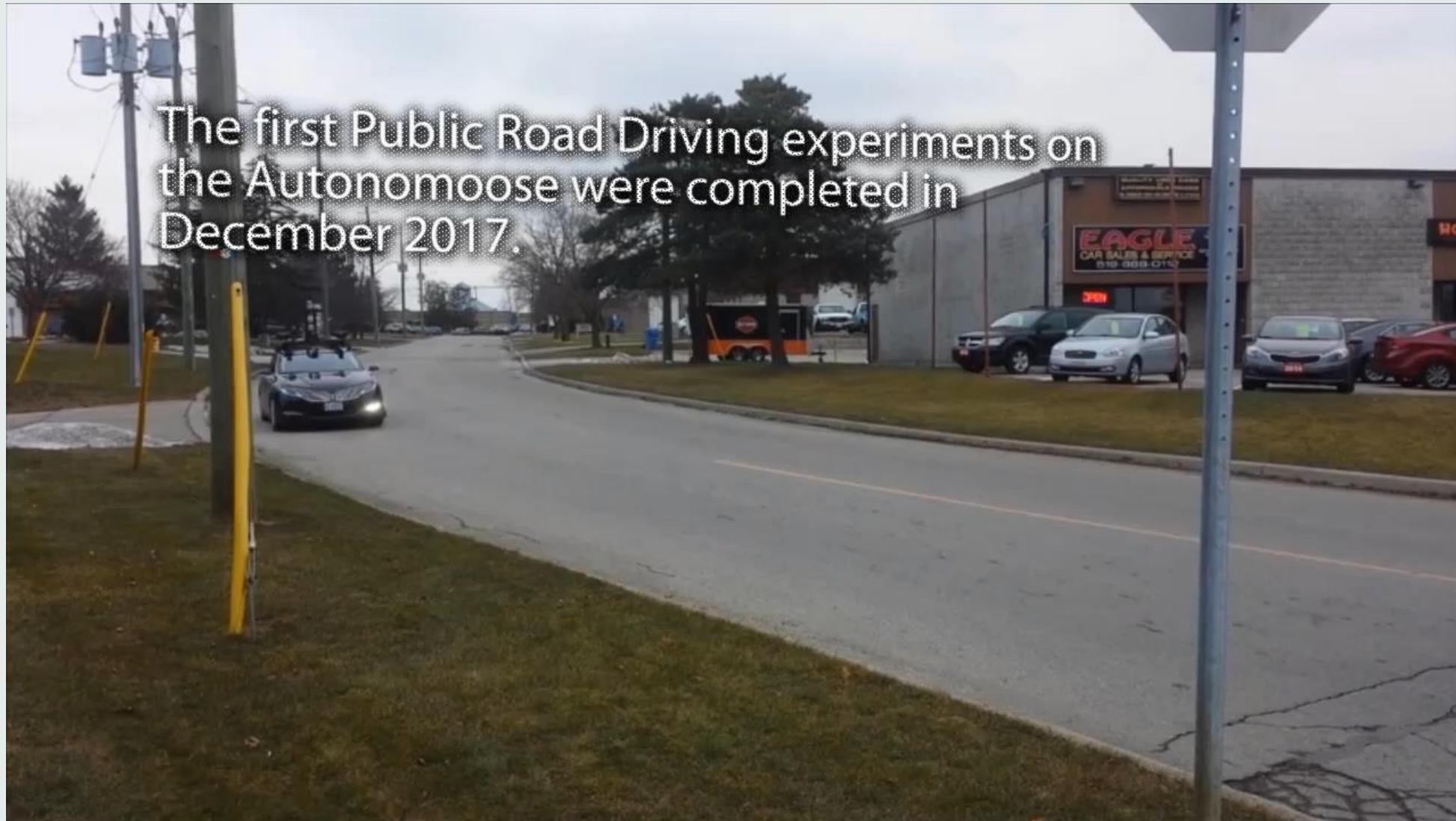


Autonomous in Six Months - CES 2017 Demonstration



<https://www.youtube.com/watch?v=KjQosq-DIMg>

First Public Road Drive – December 2017



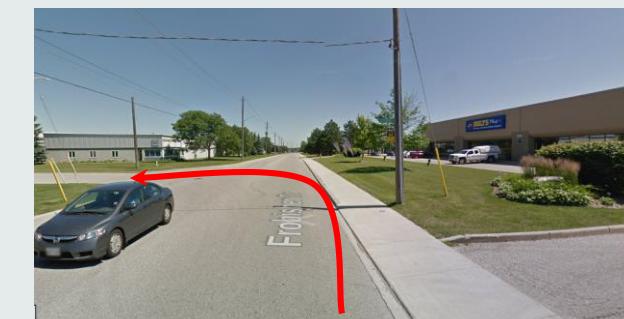
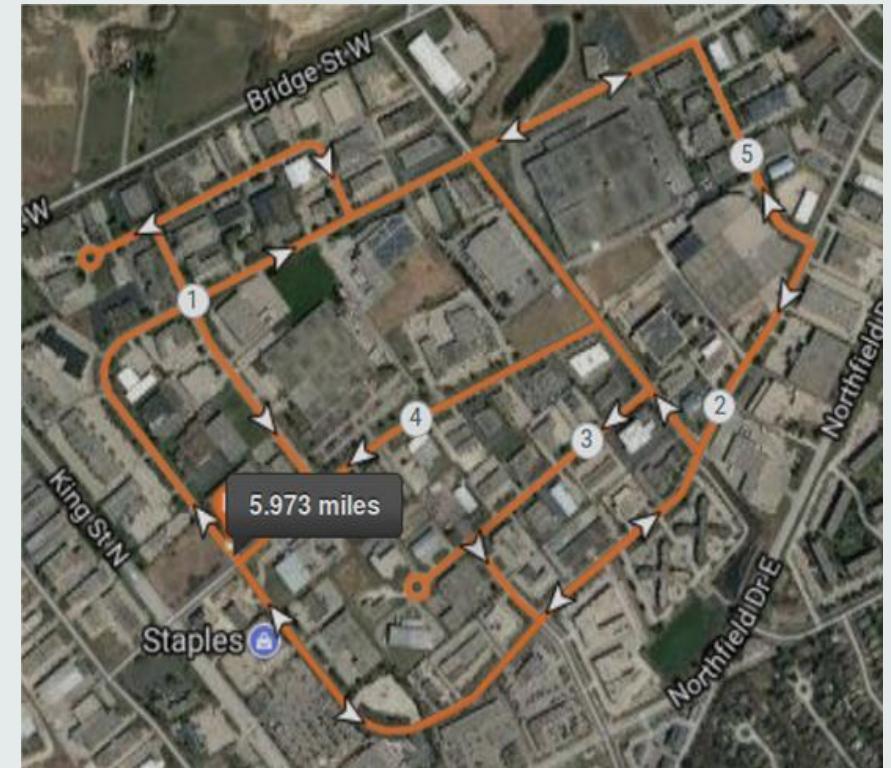
The first Public Road Driving experiments on the Autonomoose were completed in December 2017.

<https://www.youtube.com/watch?v=i7S-JZYdb74>

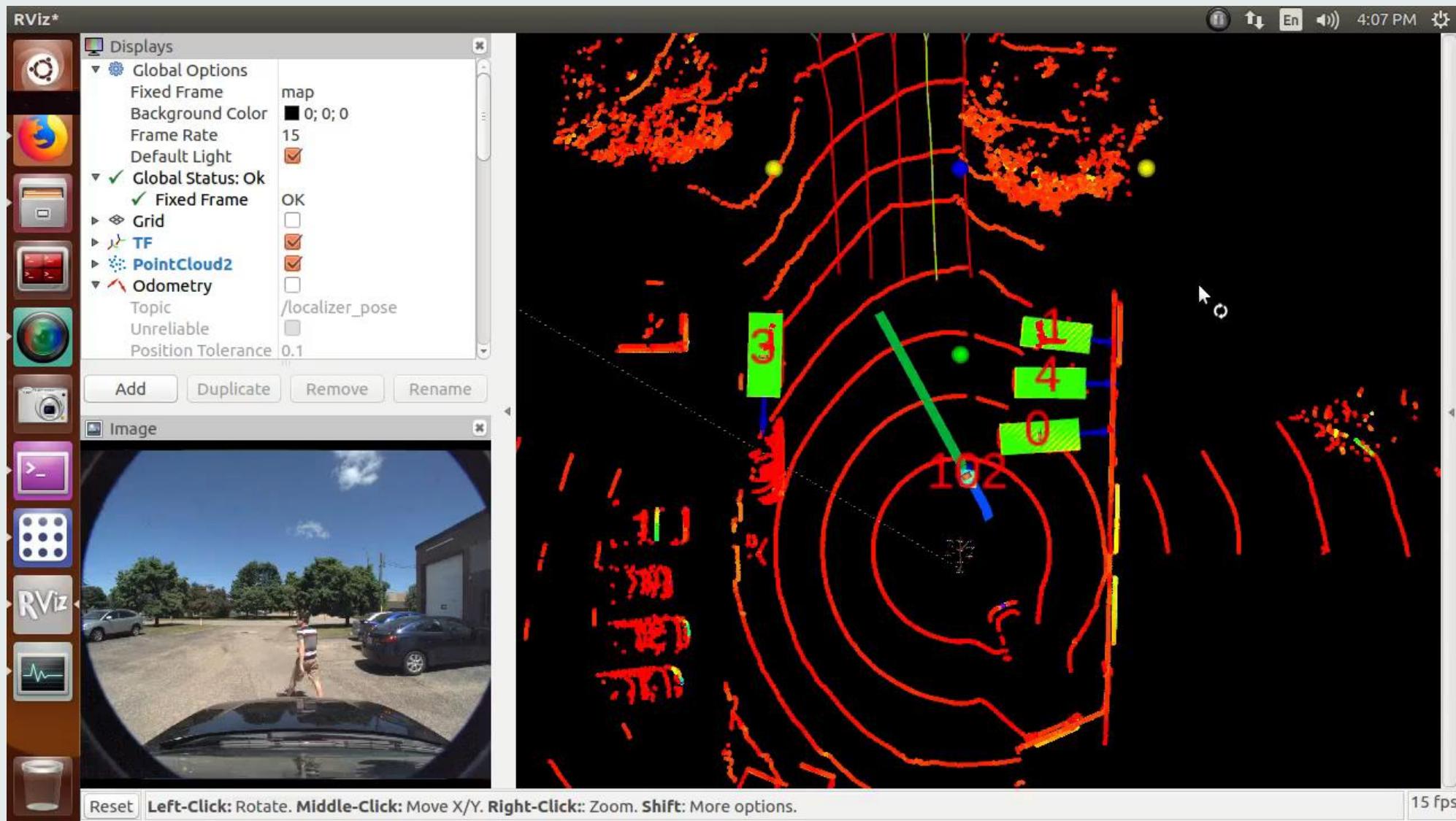
The public drive challenge

- Goal: 100 km on public roads with MTBI of 5 minutes
- Colby test lap
 - Industrial area with limited traffic
 - 2 km loop, one stop sign
- Bathurst test circuit
 - All residential, two lane roads
 - 9.6 km sequence

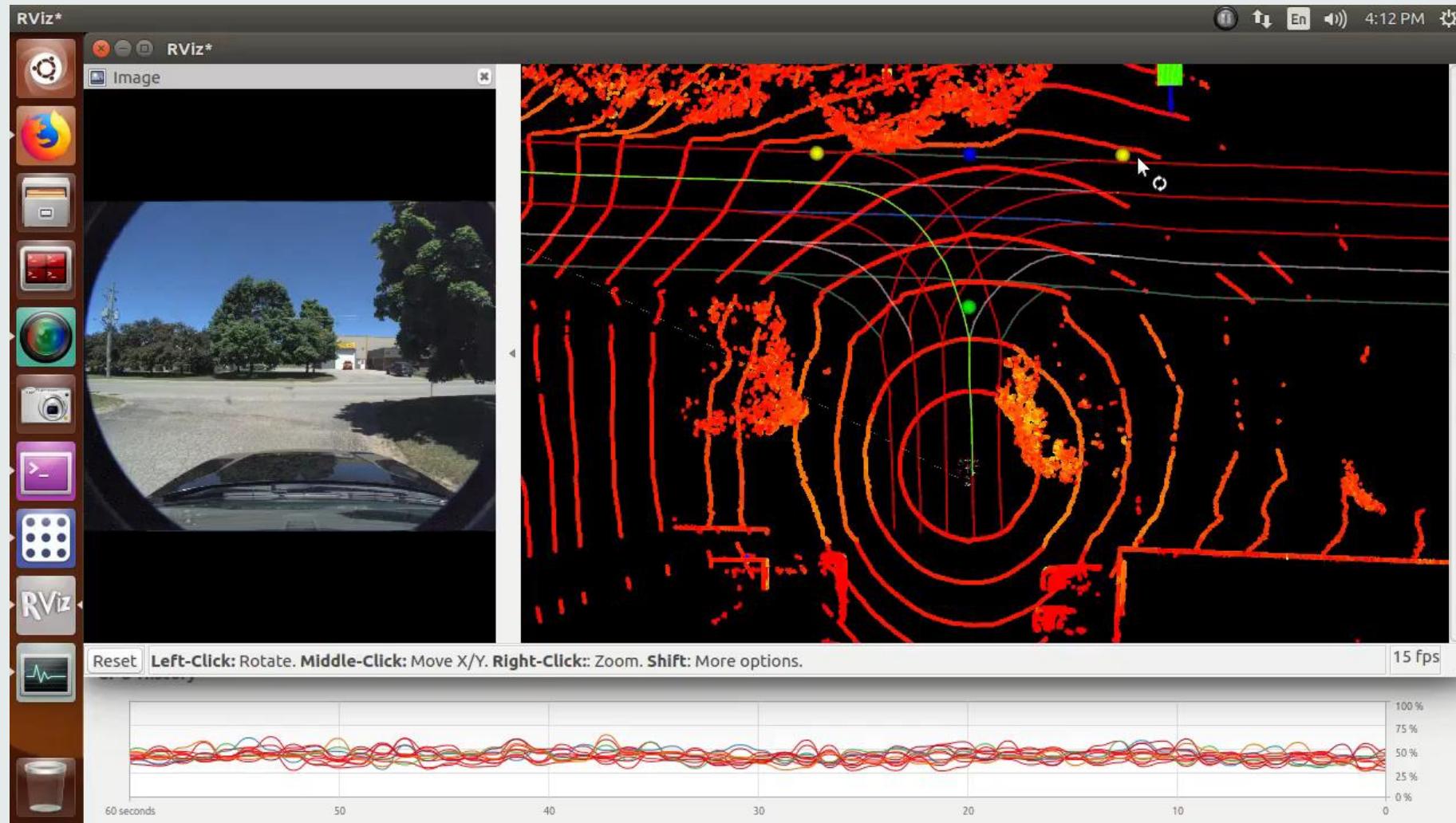
Bathurst



Detection, Tracking and Prediction - Pedestrians



Detection, Tracking and Prediction - Driving



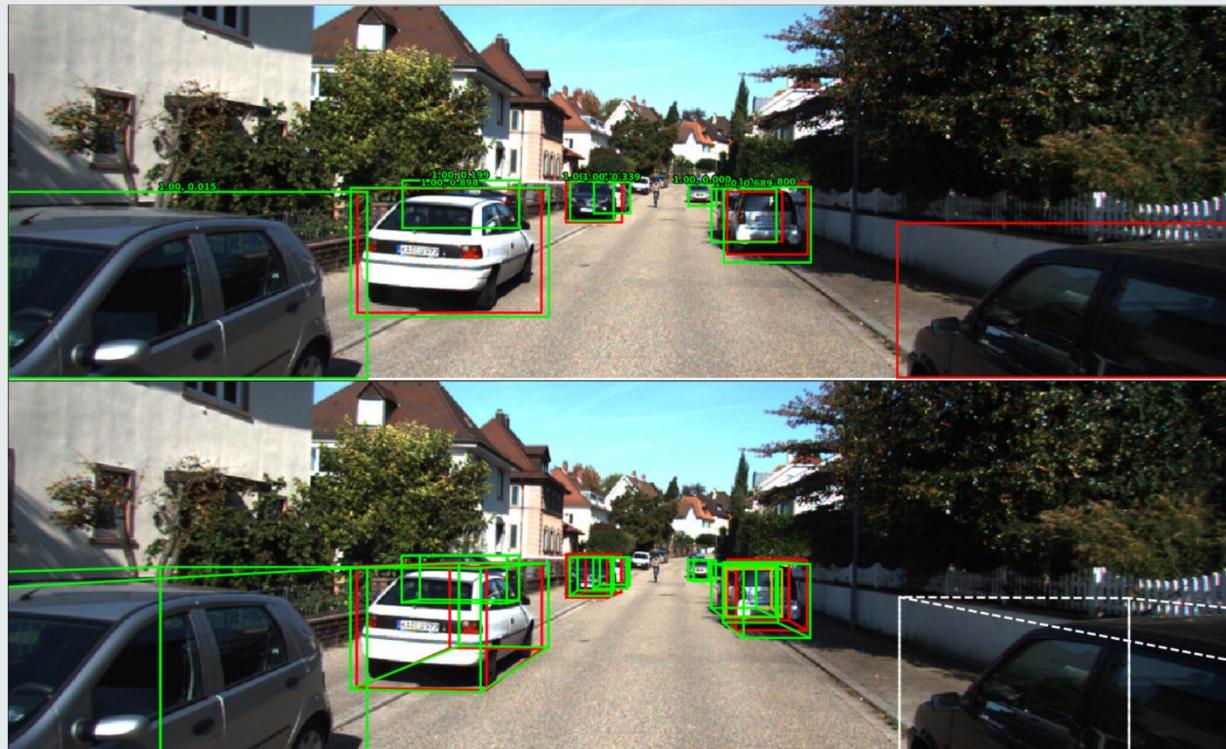
Outline

- The State of Autonomous Driving in Canada
- Waterloo's Autonomoose Program
- 3D Object Detection with Deep Learning
- Dataset Creation for Deep Learning
- Mitigating the Effects of Canadian Weather



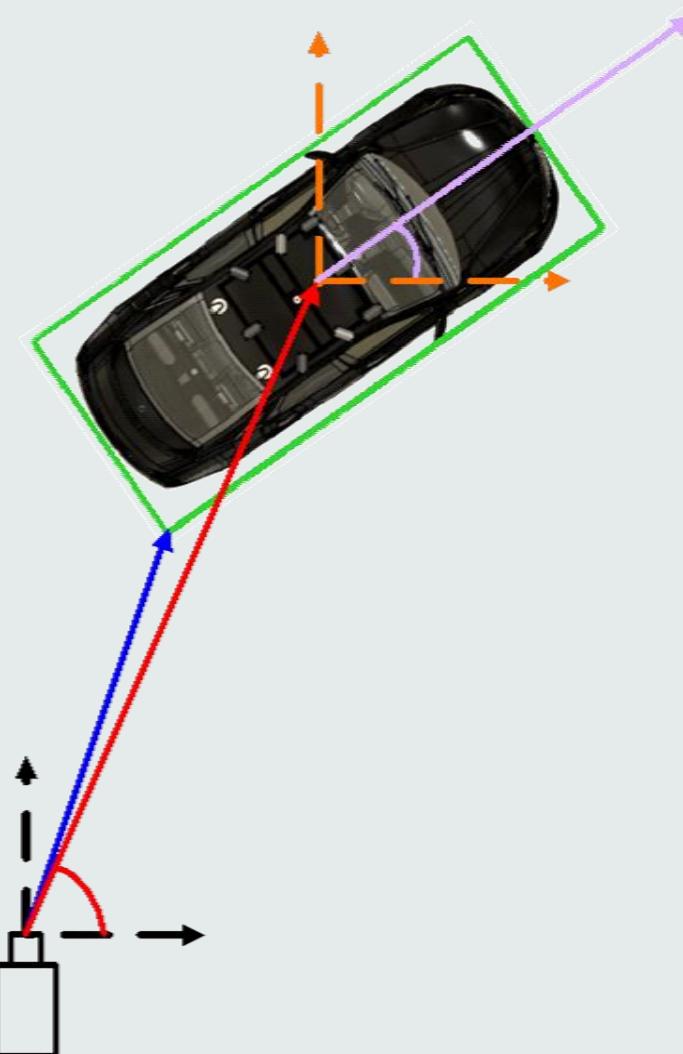
3D object detection and datasets

- Multi-channel deep learning approach that combines depth and RGB information
- 3D position, orientation and extent far more challenging than 2D
 - Best algorithms at 70% mean AP (detection box accuracy)

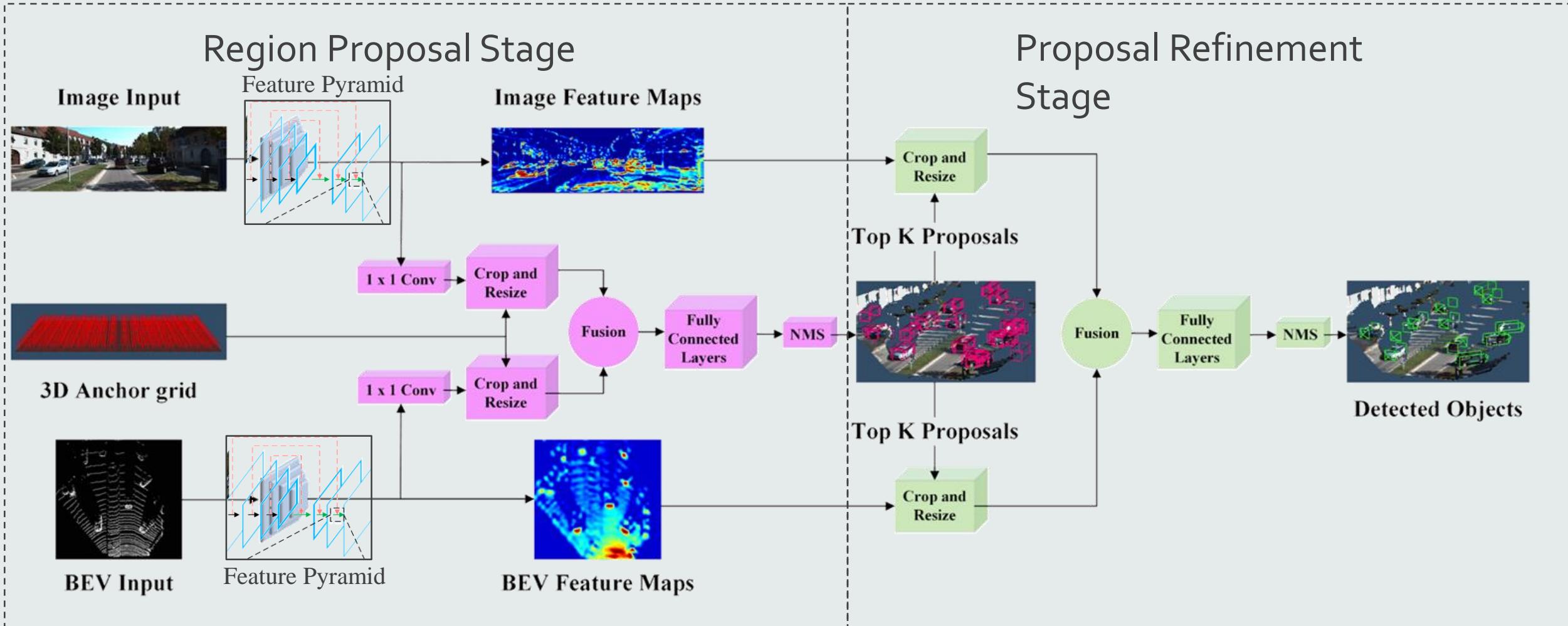


3D Object Detection Objectives

- Class identification
 - Car, pedestrian, bicycle
- Centroid Estimation
 - X,Y,Z position in camera frame
- Bounding box extent
 - Intersection over Union with ground truth
- Vehicle Orientation
 - Forward direction



AVOD-FPN Meta-Architecture



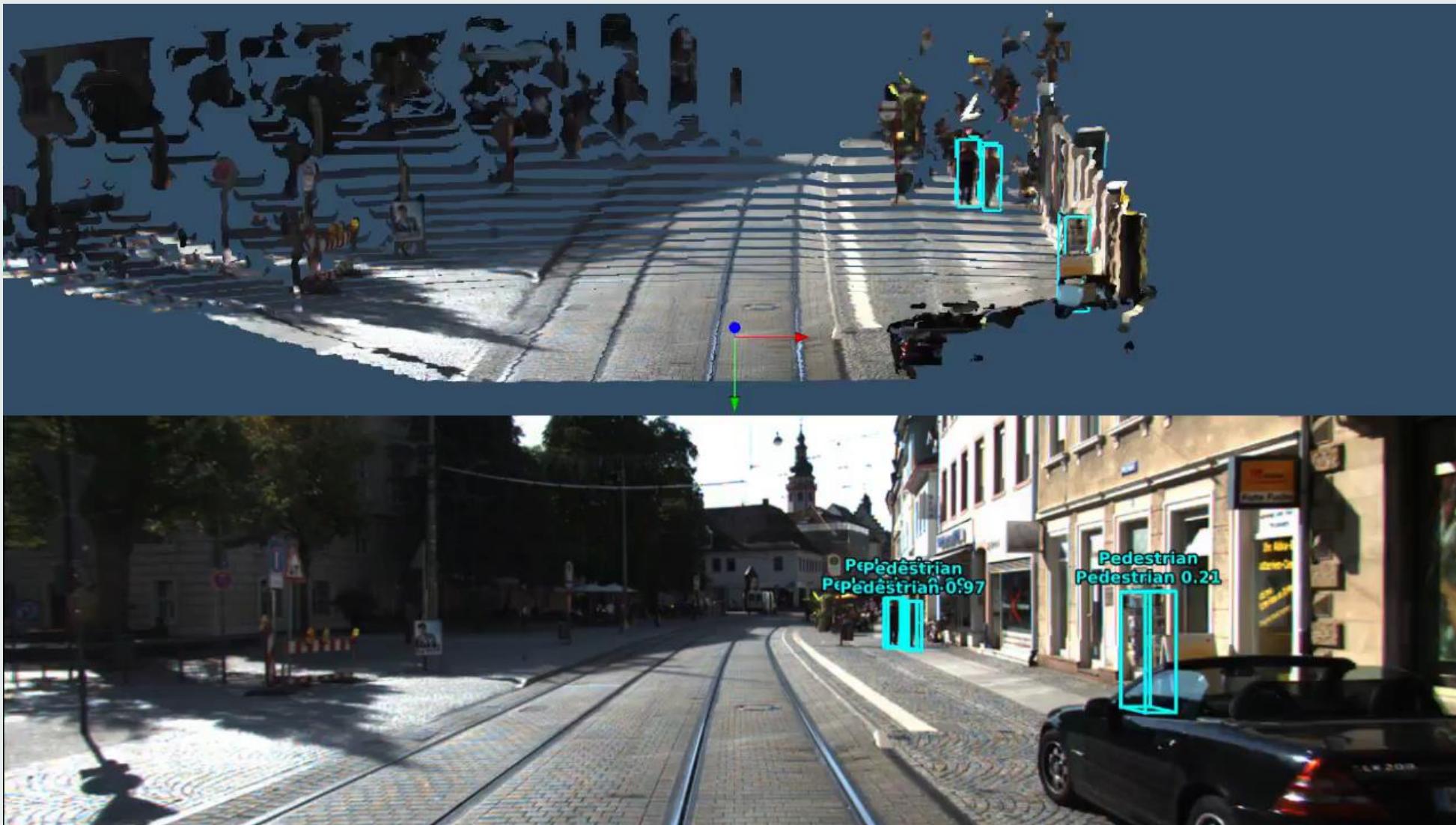
3D Object Detection: Results



3D Object Detection: Results



3D Object Detection Results



AVOD-FPN, AVOD Results

Car

	Method	Setting	Code	Moderate	Easy	Hard	Runtime	Environment
1	AVOD-FPN		code	71.88 %	81.94 %	66.38 %	0.1 s	Titan X
J. Ku, M. Mozifian, J. Lee, A. Harakeh and S. Waslander: Joint 3D Proposal Generation and Object Detection from View Aggregation . arXiv preprint arXiv:1803.08168 (2018)								
2	F-PointNet			70.39 %	81.20 %	62.19 %	0.17 s	GPU @ 3.0 Ghz
J. Ku, M. Mozifian, J. Lee, A. Harakeh and S. Waslander: Joint 3D Proposal Generation and Object Detection from View Aggregation . arXiv preprint arXiv:1803.08168 (2018)								
3	DF-PC_CNN			66.22 %	80.28 %	58.94 %	0.5 s	GPU @ 3.0 Ghz
J. Ku, M. Mozifian, J. Lee, A. Harakeh and S. Waslander: Joint 3D Proposal Generation and Object Detection from View Aggregation . arXiv preprint arXiv:1803.08168 (2018)								
4	AVOD		code	65.78 %	73.59 %	58.38 %	0.08 s	Titan X
J. Ku, M. Mozifian, J. Lee, A. Harakeh and S. Waslander: Joint 3D Proposal Generation and Object Detection from View Aggregation . arXiv preprint arXiv:1803.08168 (2018)								
5	VxNet(LiDAR)			65.11 %	77.47 %	57.73 %	0.03 s	GPU @ 2.5 Ghz

Pedestrian

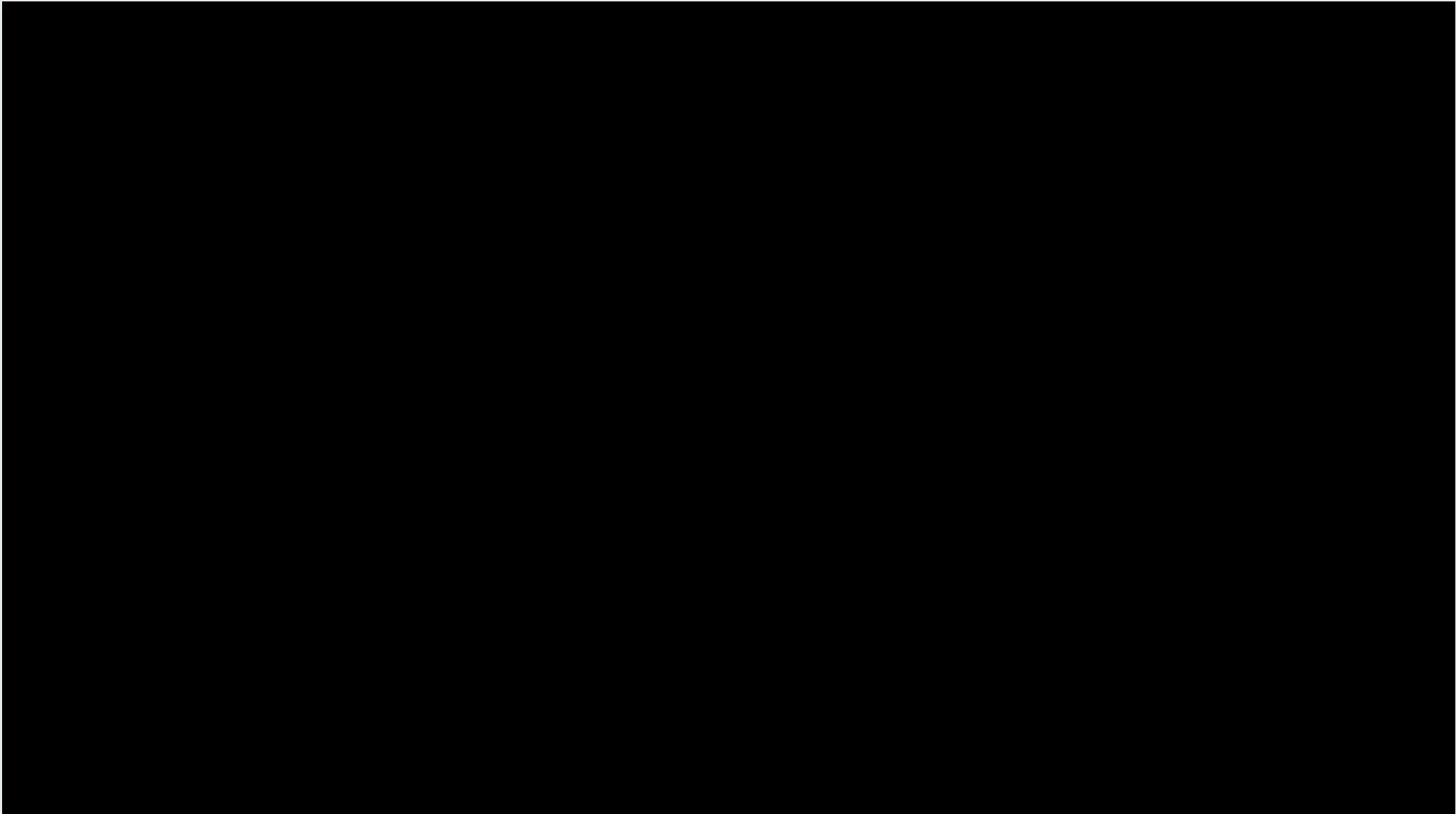
	Method	Setting	Code	Moderate	Easy	Hard	Runtime	Environment
1	F-PointNet			44.89 %	51.21 %	40.23 %	0.17 s	GPU
2	AVOD-FPN		code	39.00 %	46.35 %	36.58 %	0.1 s	Titan X
J. Ku, M. Mozifian, J. Lee, A. Harakeh and S. Waslander: Joint 3D Proposal Generation and Object Detection from View Aggregation . arXiv preprint arXiv:1803.08168 (2018)								
3	VxNet(LiDAR)			33.69 %				
4	AVOD		code	31.51 %				
J. Ku, M. Mozifian, J. Lee, A. Harakeh and S. Waslander: Joint 3D Proposal Generation and Object Detection from View Aggregation . arXiv preprint arXiv:1803.08168 (2018)								

Cyclist

	Method	Setting	Code	Moderate	Easy	Hard	Runtime	Environment
1	F-PointNet			56.77 %	71.96 %	50.39 %	0.17 s	GPU @ 3.0 Ghz (Pyth)
2	VxNet(LiDAR)			48.36 %	61.22 %	44.37 %	0.03 s	GPU @ 2.5 Ghz (Pyth)
J. Ku, M. Mozifian, J. Lee, A. Harakeh and S. Waslander: Joint 3D Proposal Generation and Object Detection from View Aggregation . arXiv preprint arXiv:1803.08168 (2018)								
3	AVOD		code	42.36 %	57.73 %	40.11 %	0.08 s	Titan X (Pas)

<https://github.com/kujason/avod>

3D Object Detection Results



Outline

- The State of Autonomous Driving in Canada
- Waterloo's Autonomoose Program
- 3D Object Detection with Deep Learning
- Dataset Creation for Deep Learning
- Mitigating the Effects of Canadian Weather



Importance of Datasets

More data is needed to allow for better generalization

- Rule of Thumb for dataset size
 - **5K** for labelled example for acceptable performance
 - **10M** to exceed human performance

Existing Datasets:

From ApolloScapes

Table 1. Total and average number of instances in Kitti, Cityscapes, and our dataset (instance-level). The letters, e, m, and h, indicate easy, moderate, and hard subsets respectively.

Count	Kitti	Cityscapes	Ours (instance)		
total ($\times 10^4$)	ApolloScape				
person	0.6	2.4	s	54.3	
vehicle	3.0	4.1		198.9	
average per image			e	m	h
person	0.8	7.0	1.1	6.2	16.9
vehicle	4.1	11.8	12.7	24.0	38.1
car	-	-	9.7	16.6	24.5
motorcycle	-	-	0.1	0.8	2.5
bicycle	-	-	0.2	1.1	2.4
rider	-	-	0.8	3.3	6.3
truck	-	-	0.8	0.8	1.4
bus	-	-	0.7	1.3	0.9
tricycle	0	0	0.4	0.3	0.2

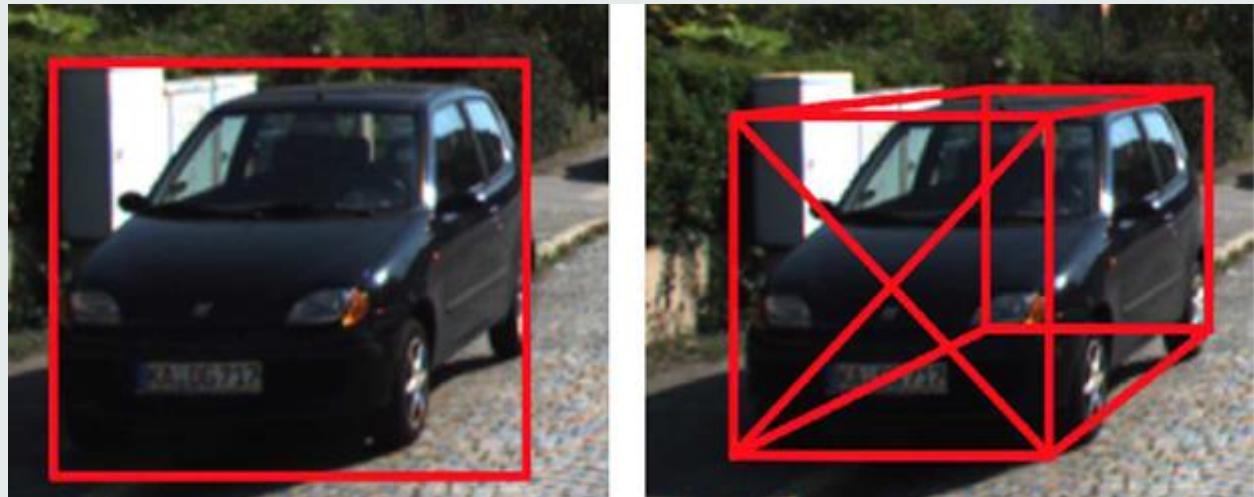
Motivation

Why is it so hard to label in 3D?

- Requires much more spatial reasoning skills than 2D
- Missing parts, occlusion (from Lidar), extent estimation
- Requires high mental imagery
- More pose parameters to estimate
- Consistency across frames

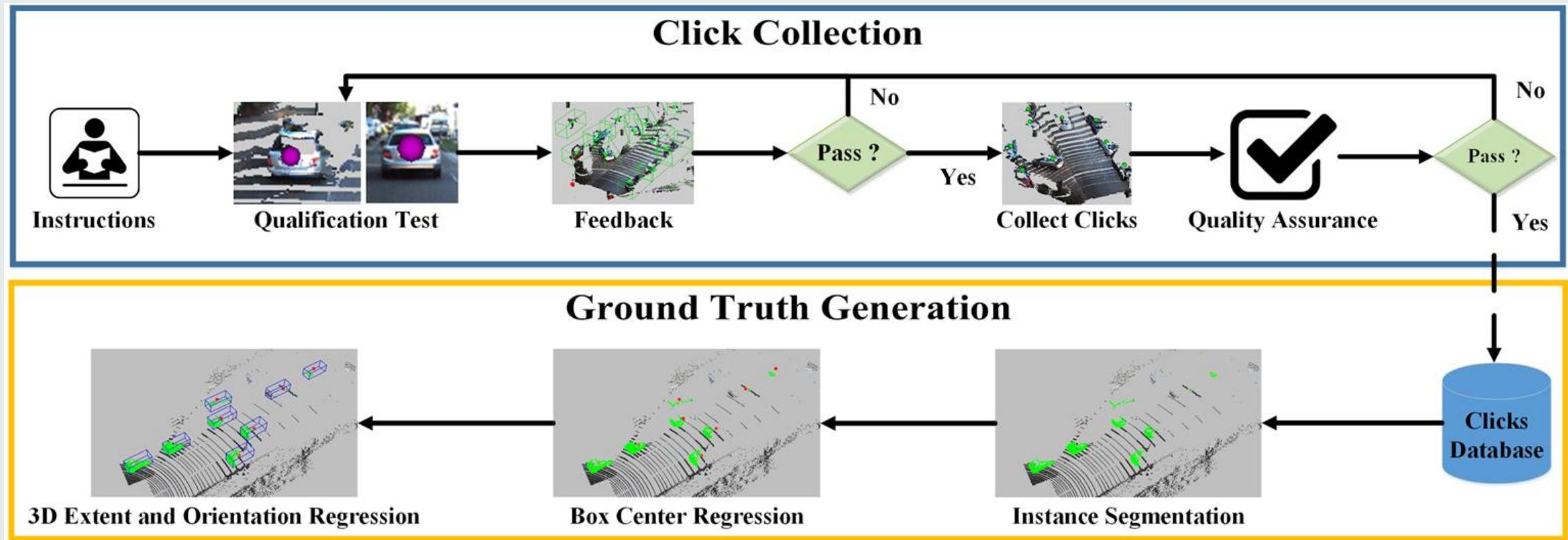
How to make it easier to label?

- Less intensive task for human
 - Reduces cost, time
 - Initialize with pre-trained networks
 - Better human to machine interface (GUI)
- Maintain Quality
 - Verifications/Refinement
 - Quality Assurance (e.g. Voting, Annotator Training)
 - Reduce task switching



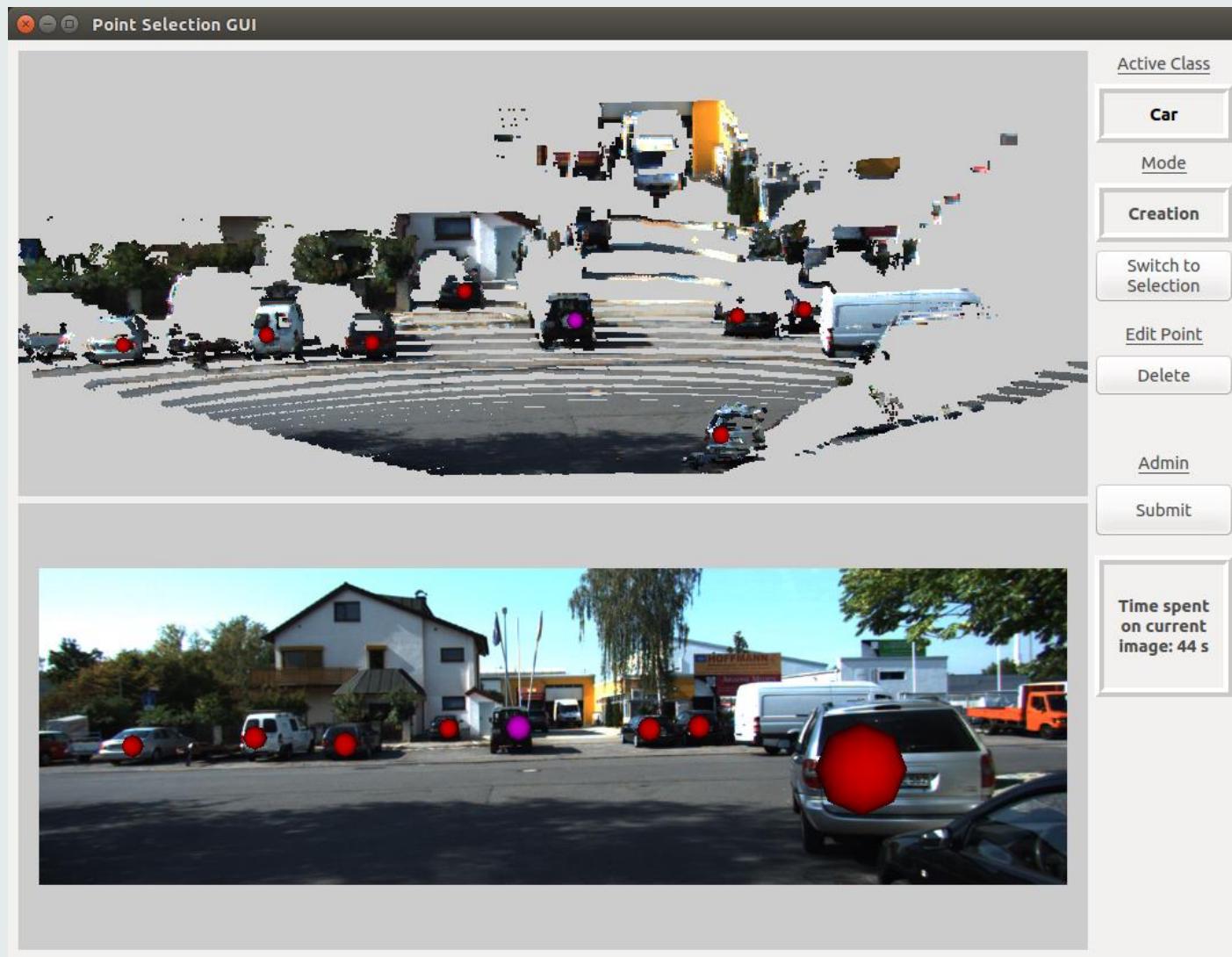
Method

Overall Procedure:



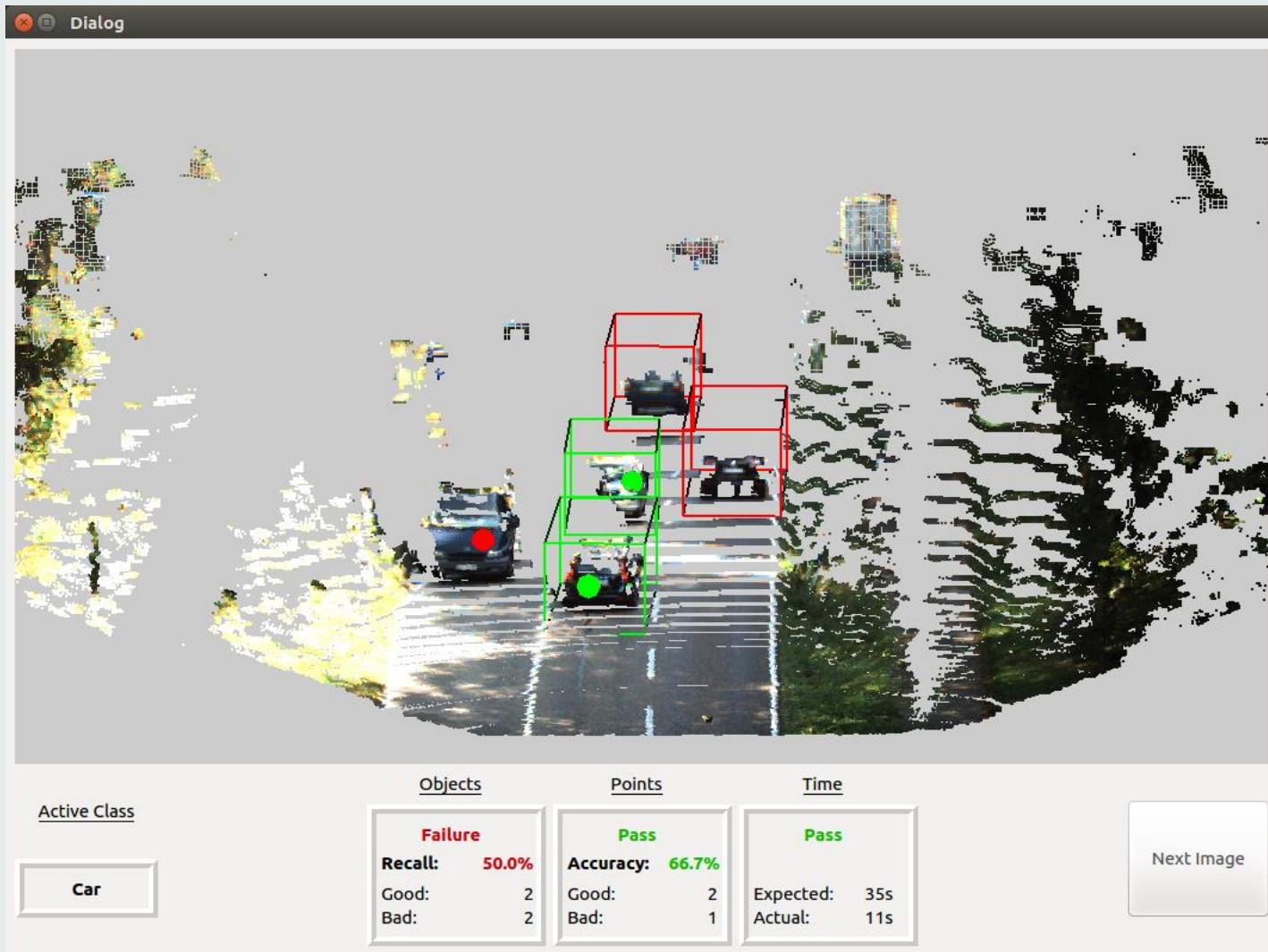
Method

Click Annotation GUI:



Method

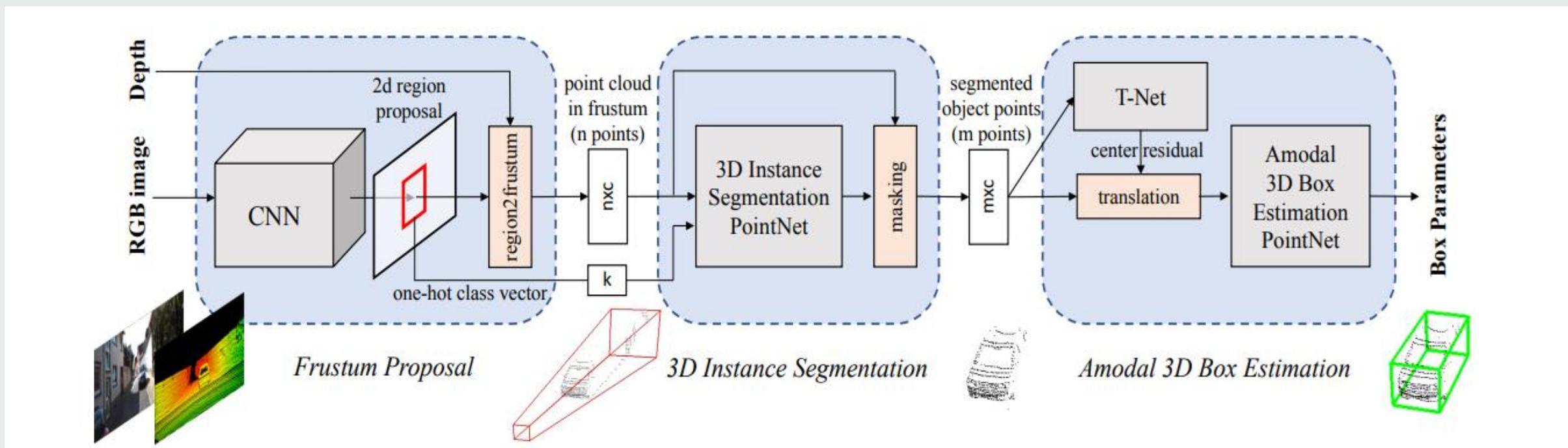
Feedback GUI:



Method

Network Configuration - Base F-Pointnet:

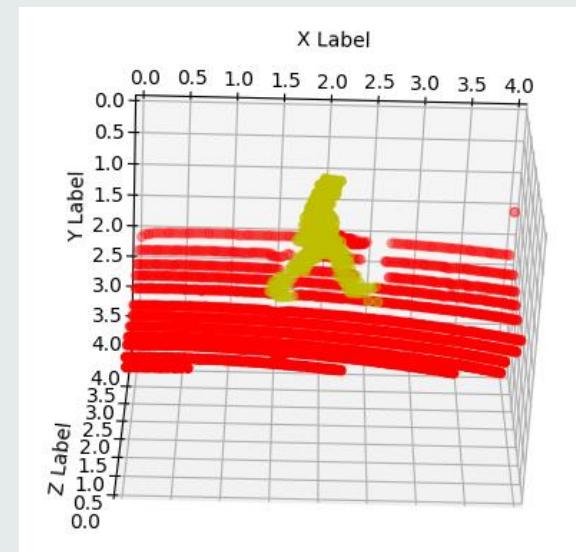
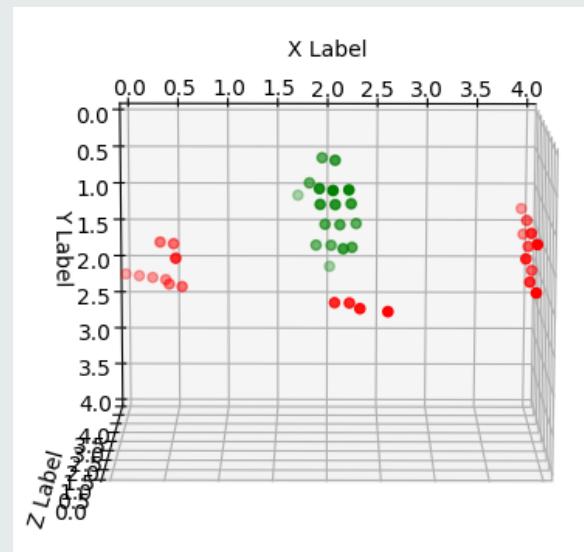
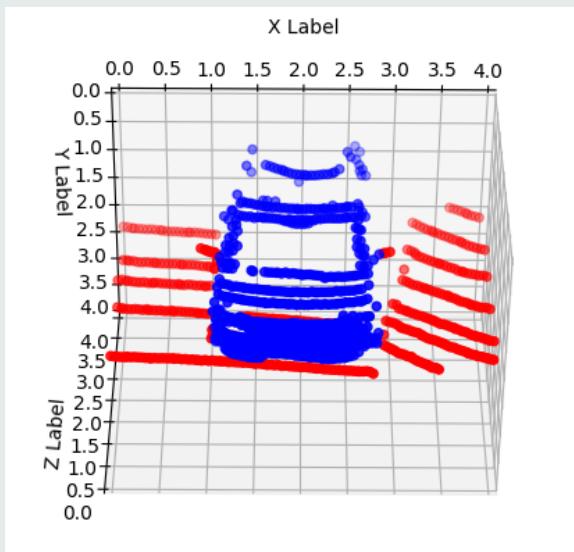
- Dependent on the quality of 2D detections



Method

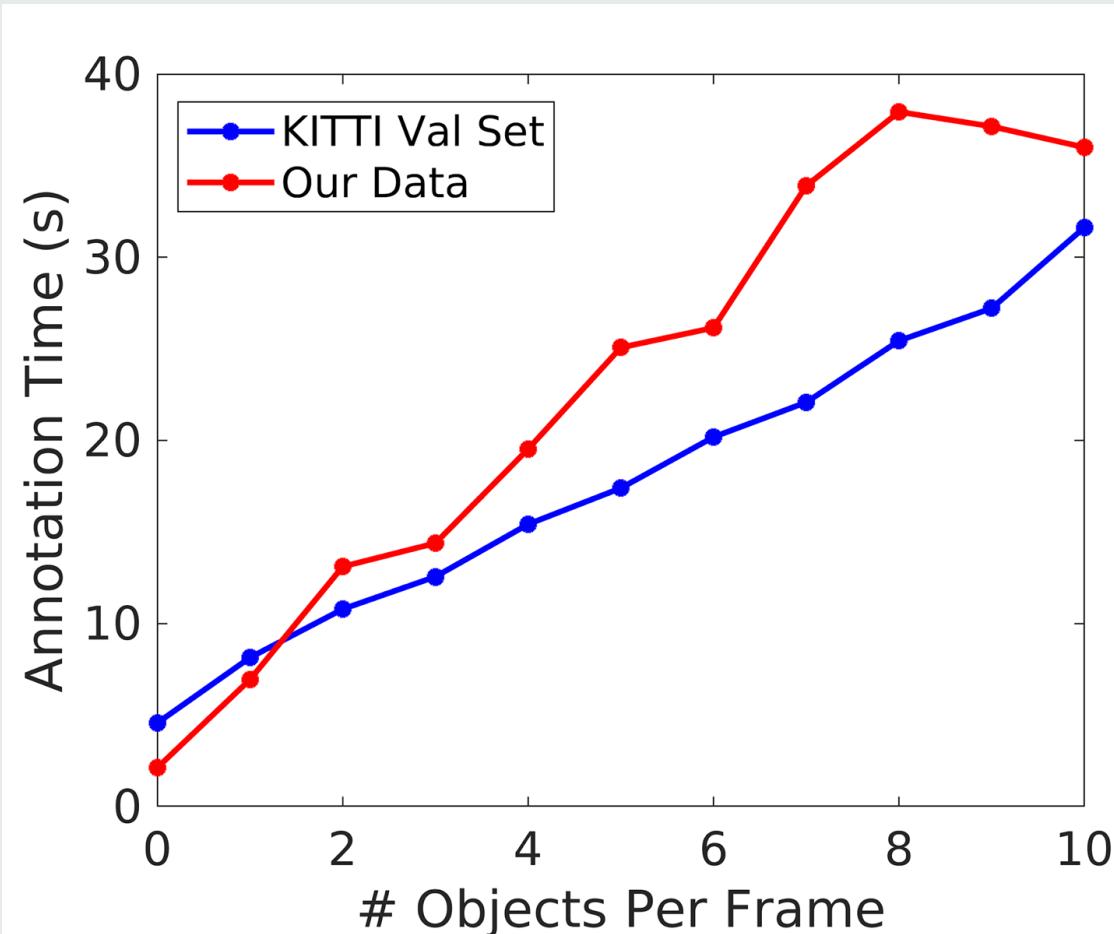
Key Modifications from F-pointnet:

- Instead of Frustum Proposals uses Cubic sampling space
(2D detections vs. Annotations)
- Training with 'simulated' clicks



Results

- Efficient for multi-object label Generation
- 96.5% Recall, 77.06% Precision on Clicks using KITTI GT (medium)



Results

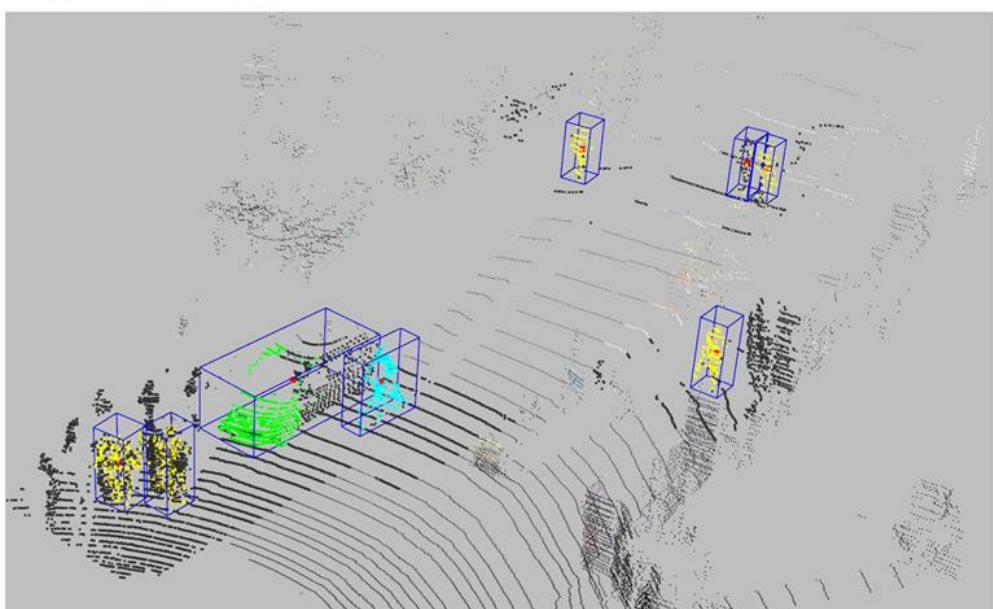
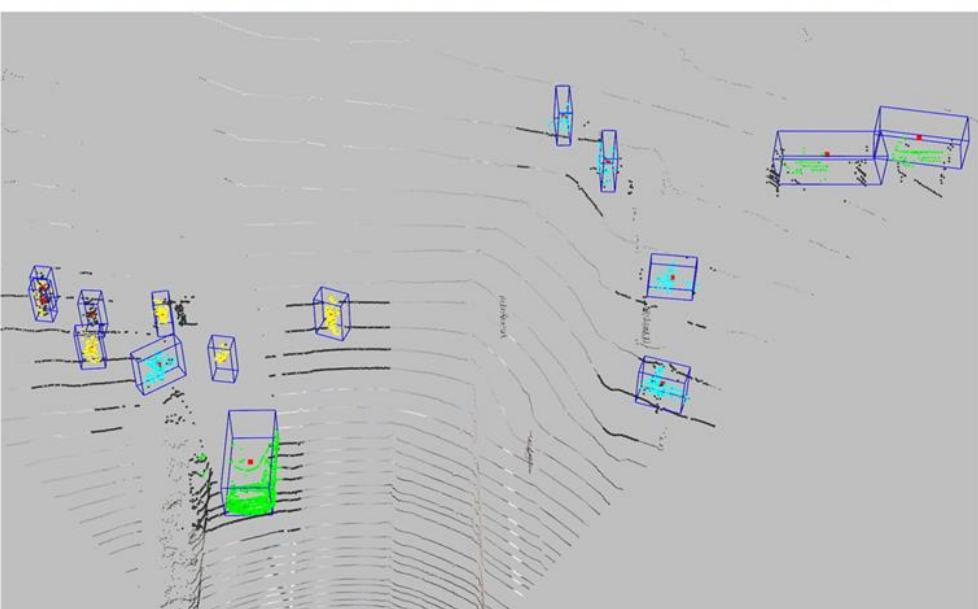
Segmentation and Bounding Box Performances on KITTI Validation:
Evaluated on lower IoU threshold (0.5, 0.25, 0.25)

Class		Number of Instances	i-IoU	Centroid Distance Error [m]	3D Box IoU	Average Precision (3D)
Car		14,318	0.84	±0.23	0.70	88.33
Pedestrian		2,280	0.88	±0.13	0.47	88.73
Cyclist		893	0.82	±0.22	0.56	87.31

TABLE I: Qualitative performance evaluation of the subnetworks in our annotation scheme. All metrics are averages across all instances belonging to a single class.

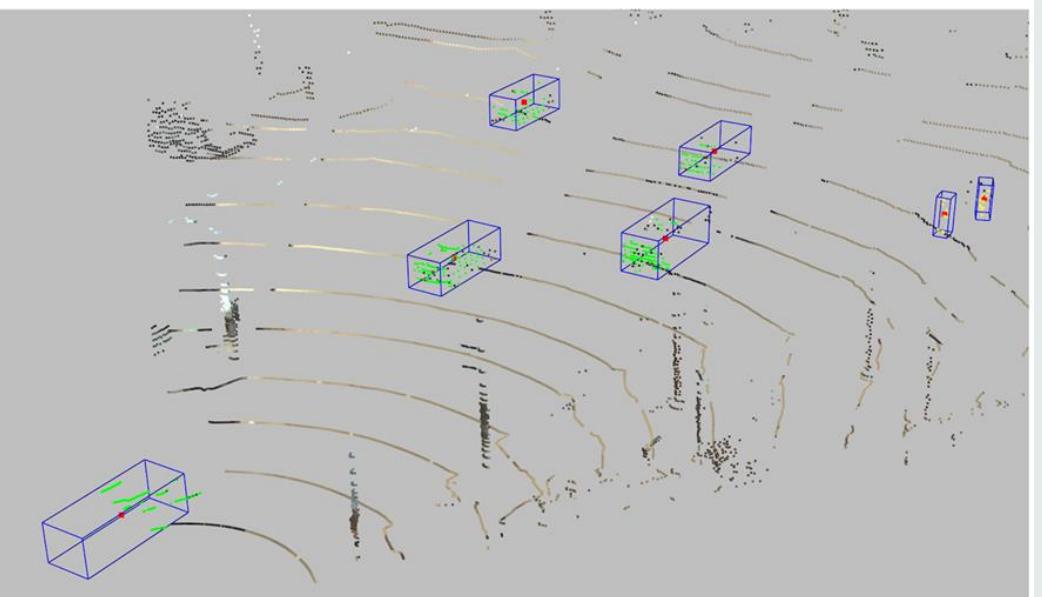
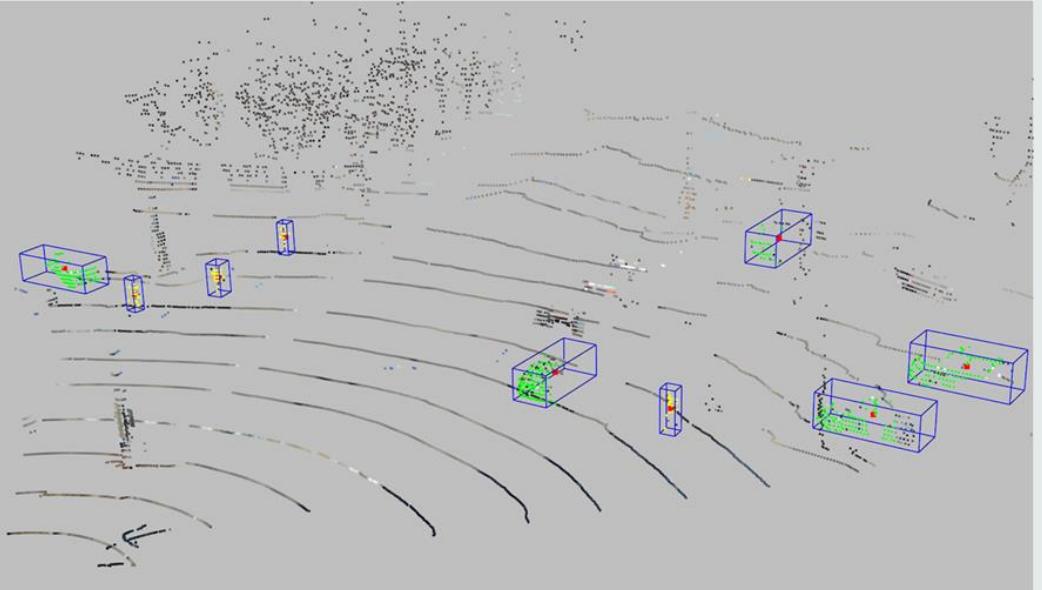
Results

Qualitative Results - KITTI



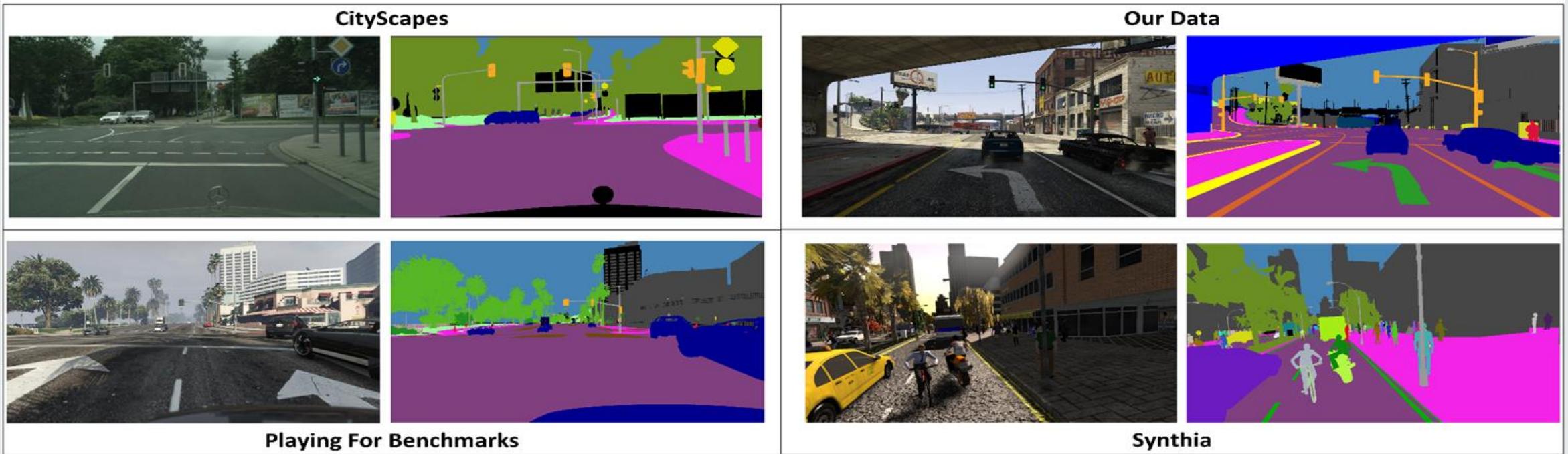
Results

Autonomoose



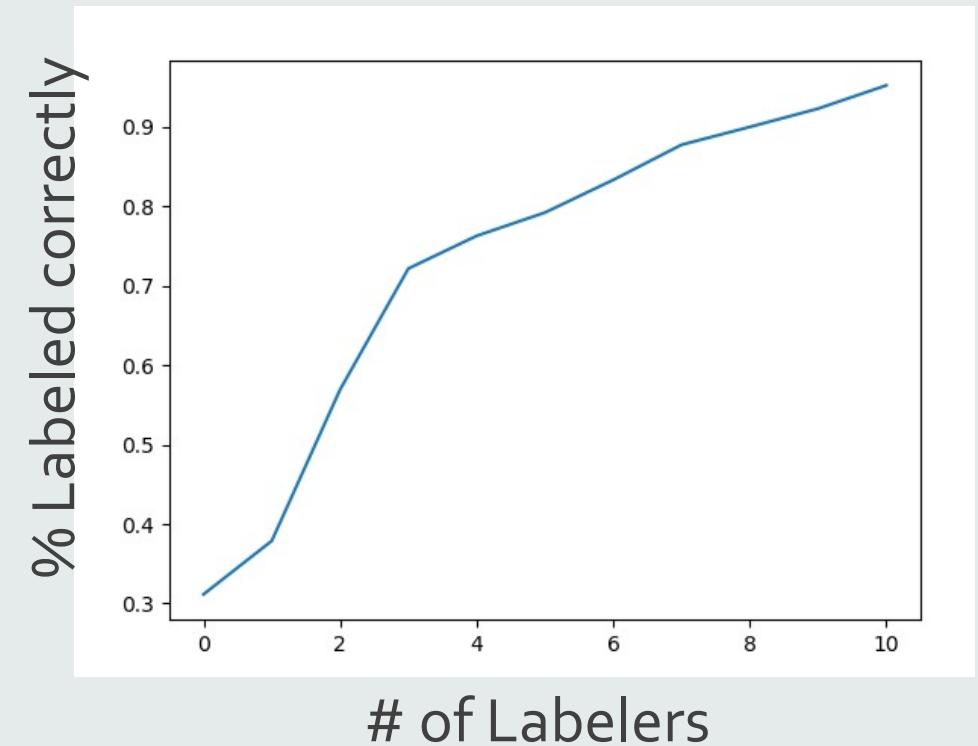
Games vs simulators

- Simulators lack the realistic look.
- Gaming industry invested a great effort to achieve realistic graphics, details, context, and relatively-real physics.
- GTAV contains 11,544 shaders, 96,441 3D-models, and over a million texture.
- Design goals are different. GTAV targets real-time performance as well as securing online players.



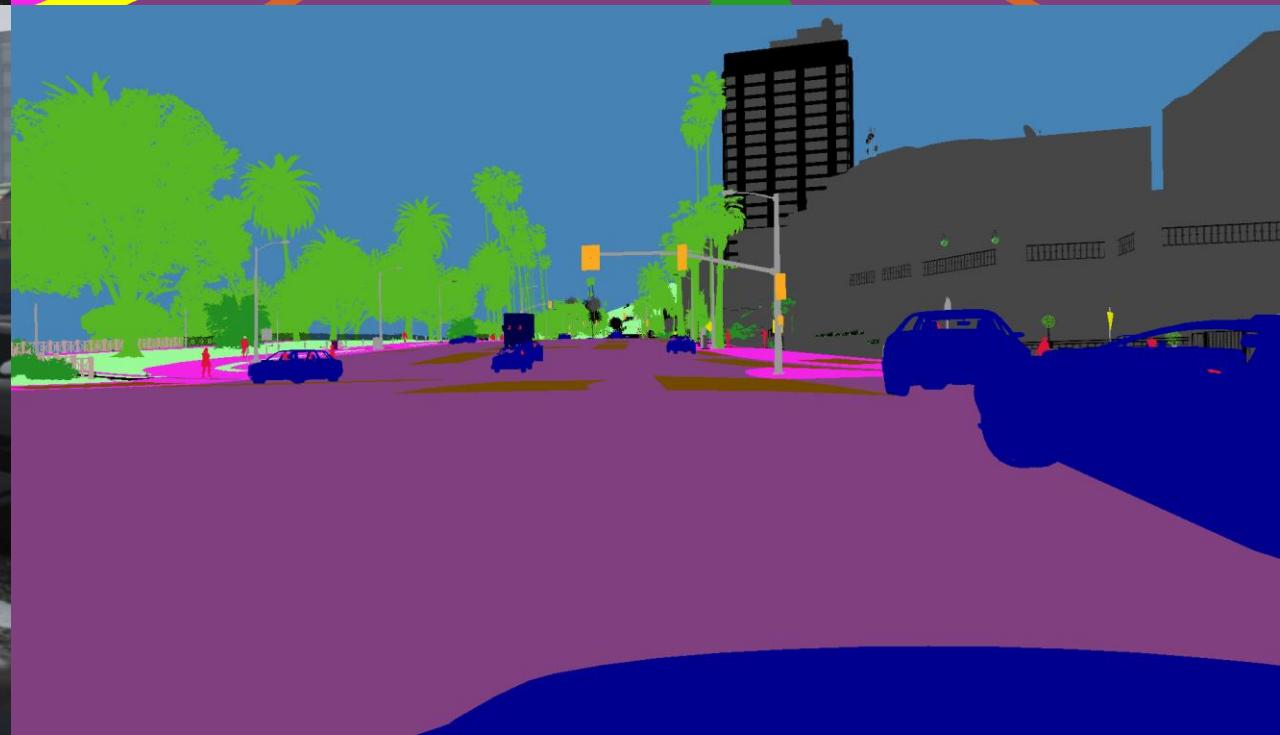
Labeling

- Different textures affect same model
- Created a unique identifier to label each texture
- Dashcam-style scene snapshots imitating the game graphics pipeline
- Amazon Mechanical Turk used to help in labeling.



Results





Outline

- The State of Autonomous Driving in Canada
- Waterloo's Autonomoose Program
- 3D Object Detection with Deep Learning
- Dataset Creation for Deep Learning
- Mitigating the Effects of Canadian Weather



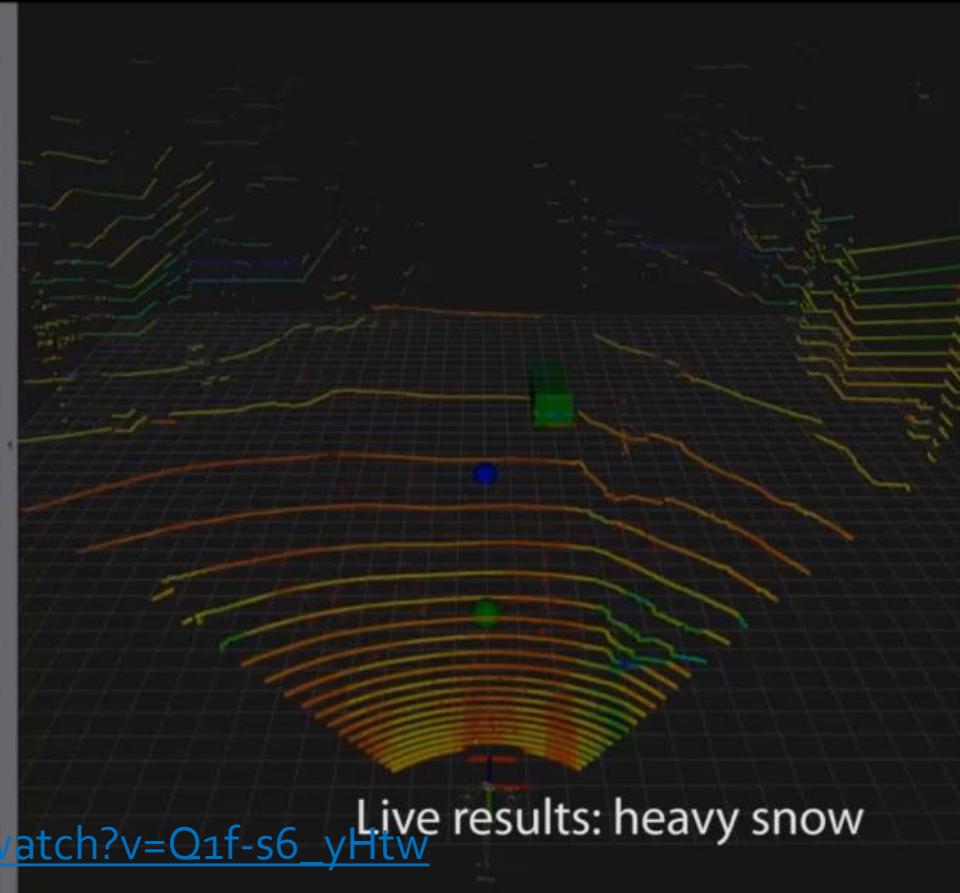
Winter Driving Conditions

- Feature-poor snow cover, reflection, large illumination contrast
- Variable density obstruction on roadways
- Precipitation of many forms, affecting sensors differently



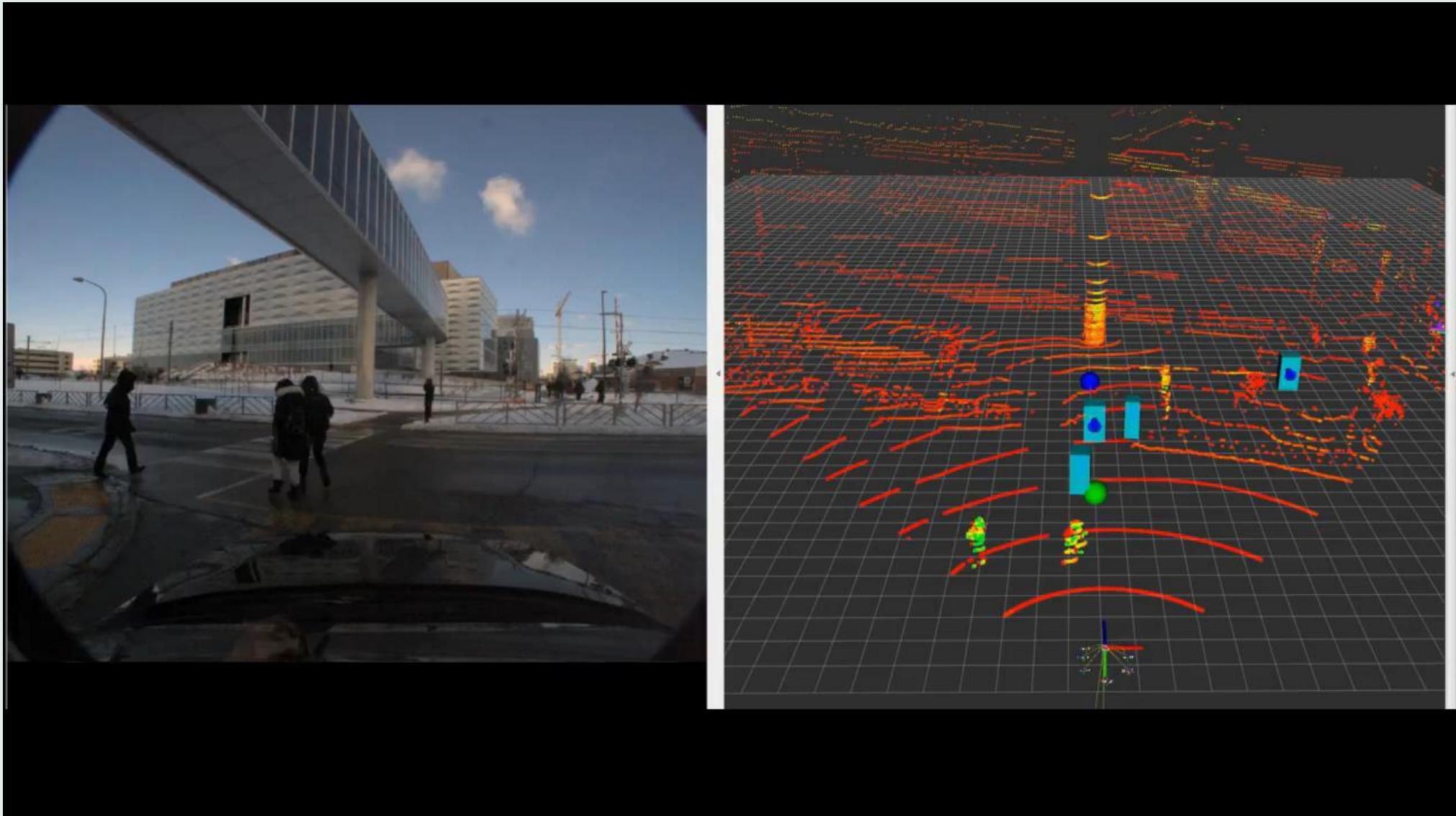
Detection in Snow

AVOD adapted surprisingly well to night and snow conditions, due to consistency of lidar data.



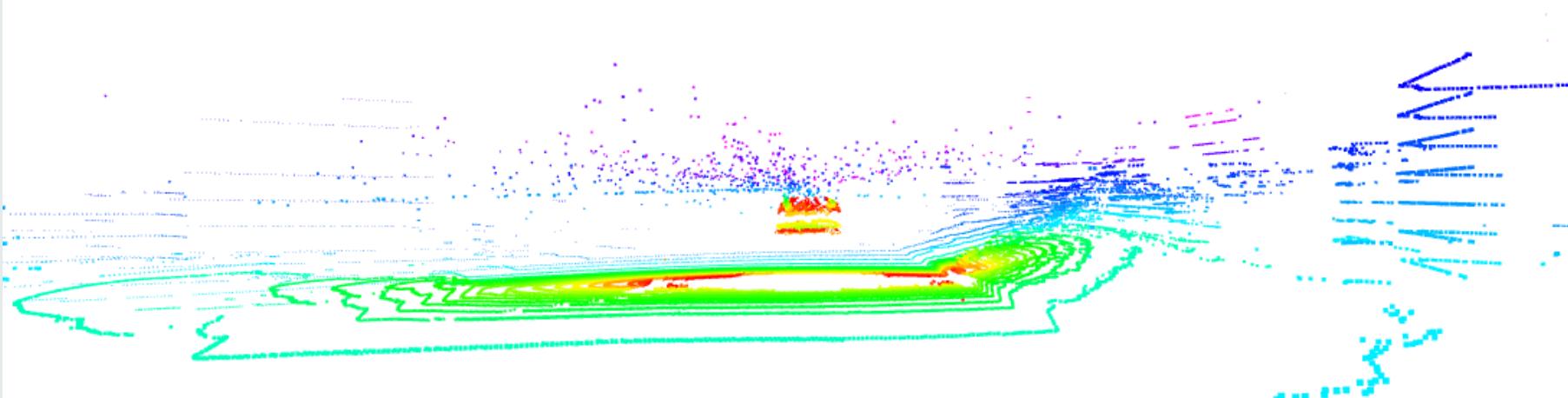
https://www.youtube.com/watch?v=Q1f-s6_yHtw

AVOD on Waterloo Pedestrians

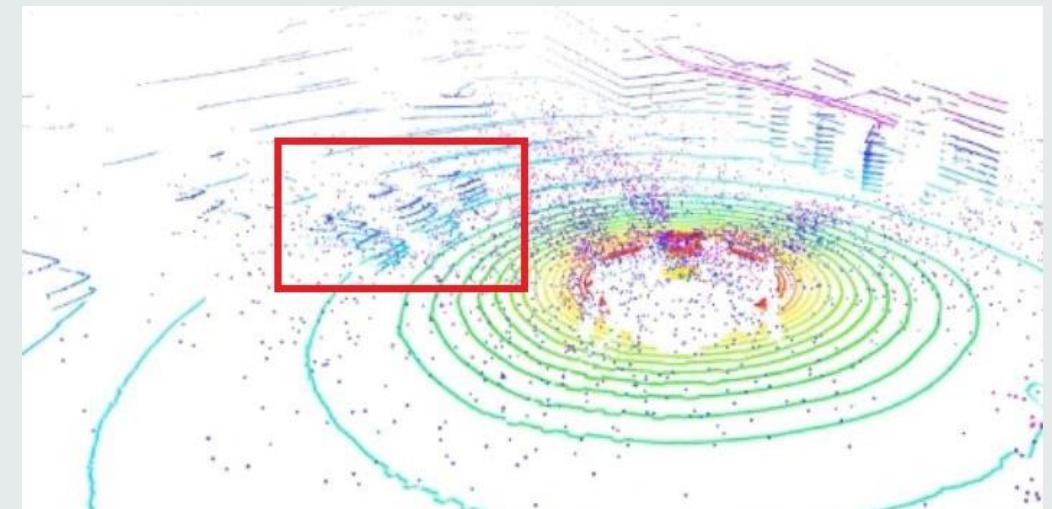


Real-Time Removal of Snow in Lidar Data

- Lidar captures falling snow, causing noisy point clouds

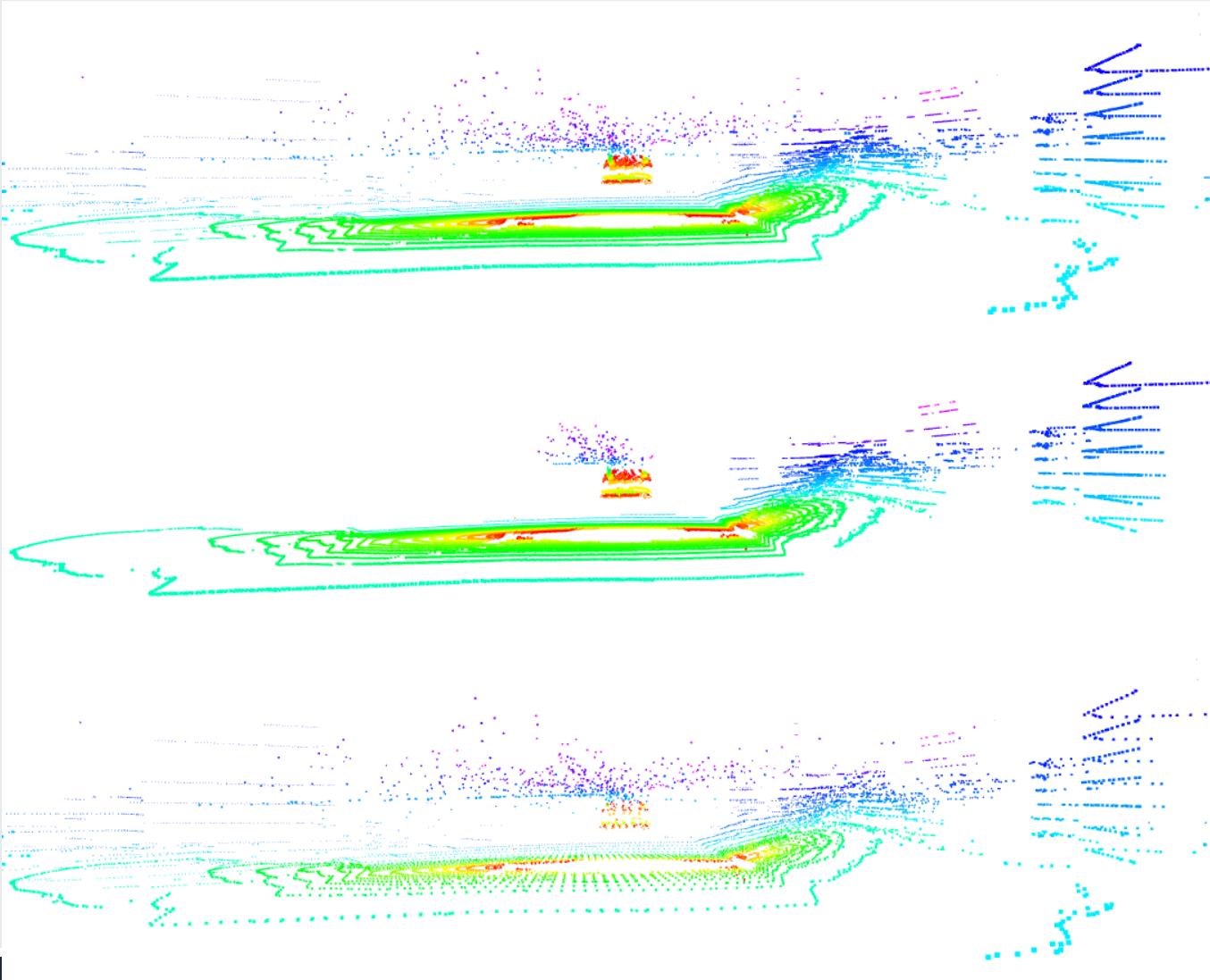


- Noisy point clouds can hinder the ability to perform autonomous driving tasks
(e.g. mapping, localization, object detection and tracking)



Real-Time Removal of Snow in Lidar Data

- Existing Point Cloud Filters



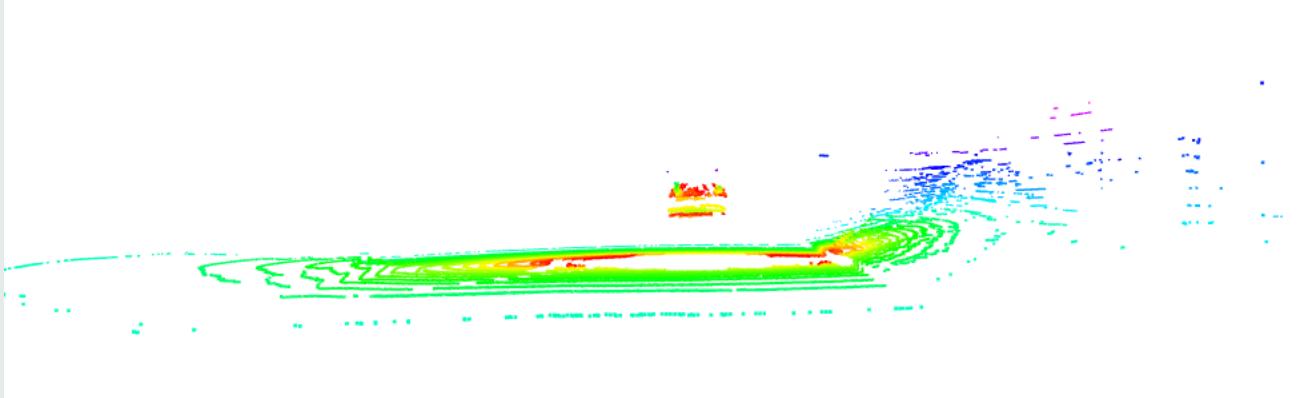
Raw Point Cloud

Statistical Outlier
Filter

Voxel Grid Filter

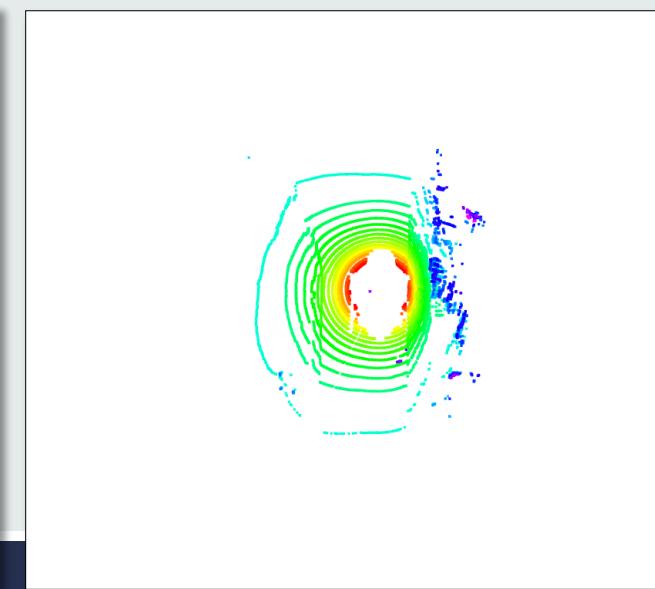
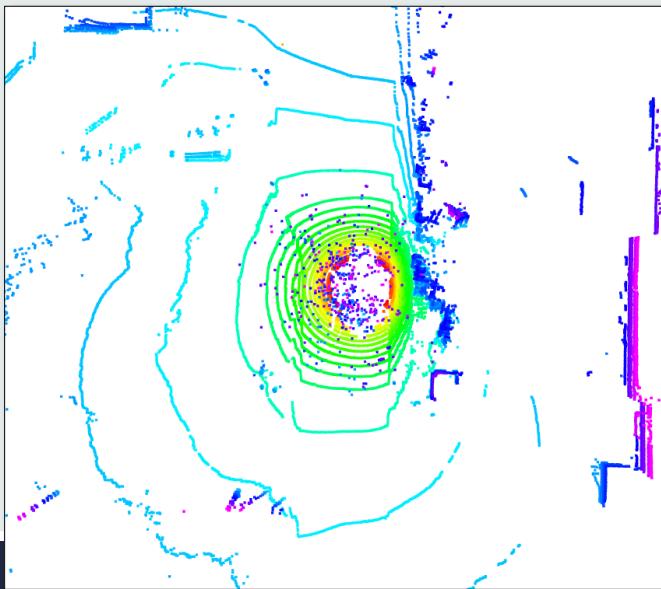
Real-Time Removal of Snow in Lidar Data

- Existing Point Cloud Filters



Radius Outlier
Removal Filter

- Feature removal problem with existing filters

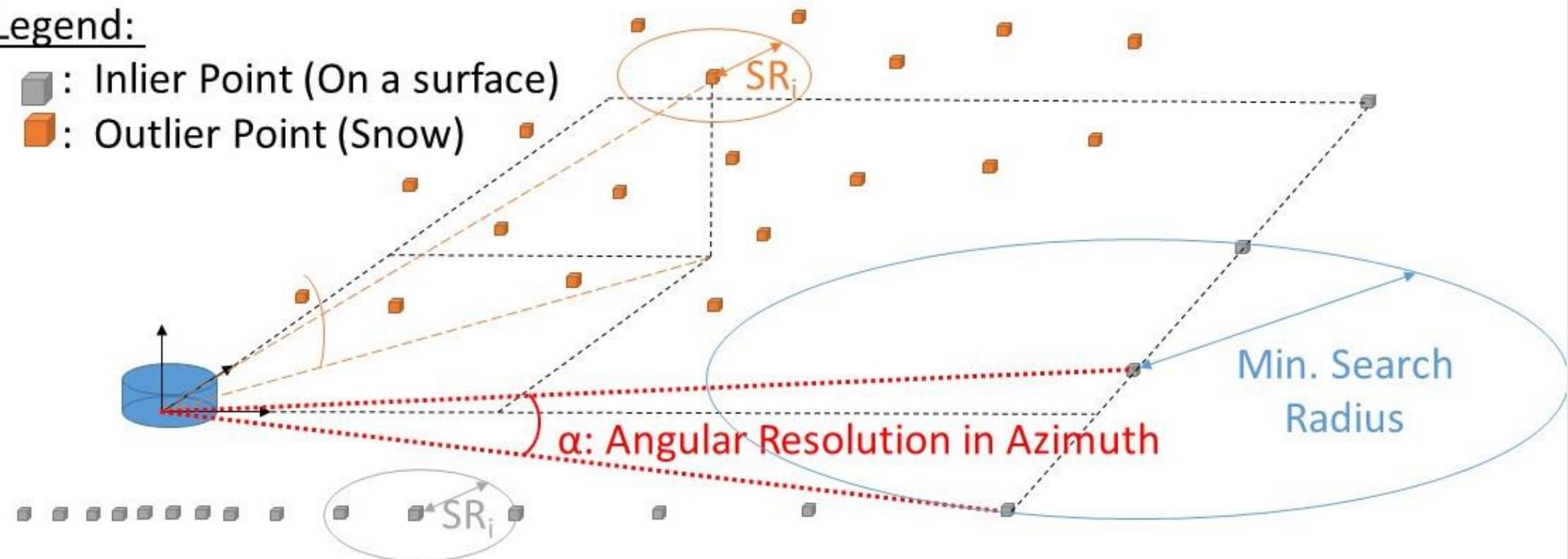


Real-Time Removal of Snow in Lidar Data

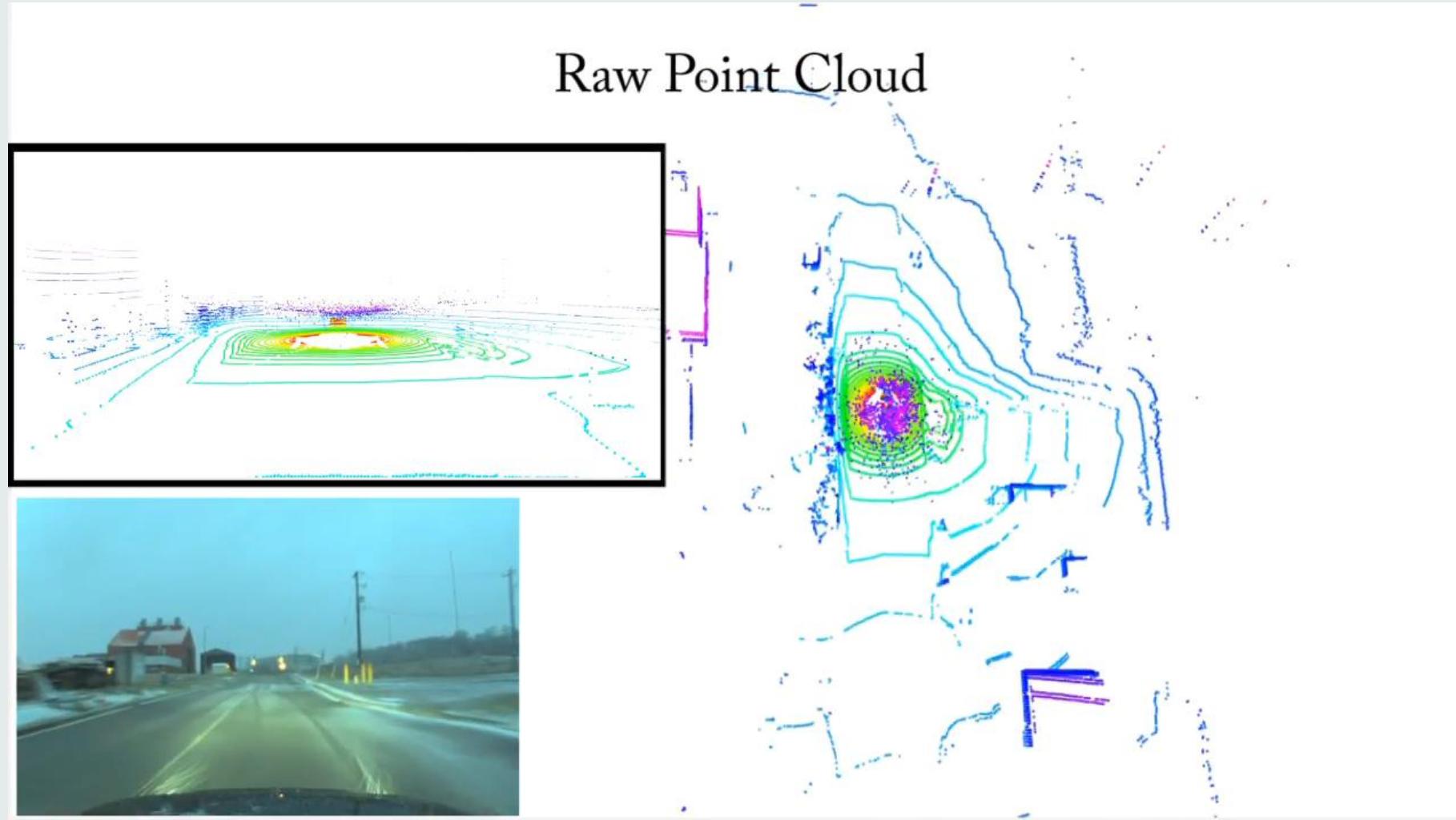
- Our Approach: Dynamic Radius Outlier Removal (DROR) Filter

Legend:

- Inlier Point (On a surface)
- Outlier Point (Snow)



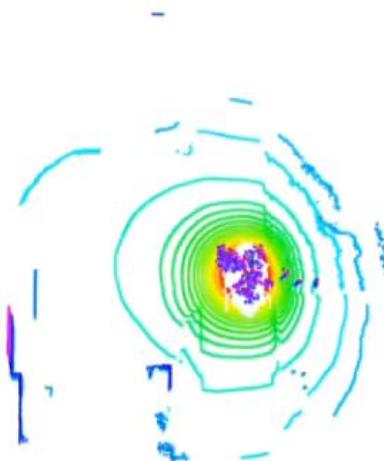
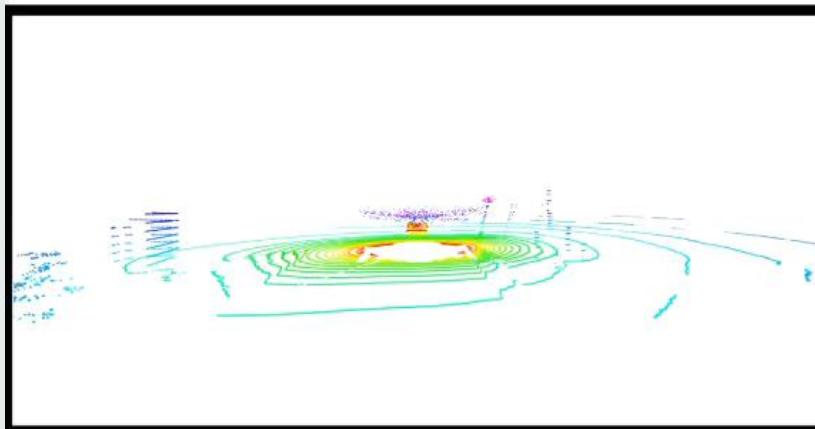
LIDAR Snow Removal



<https://www.youtube.com/watch?v=7zaf4pSe4v4>

LIDAR Snow Removal

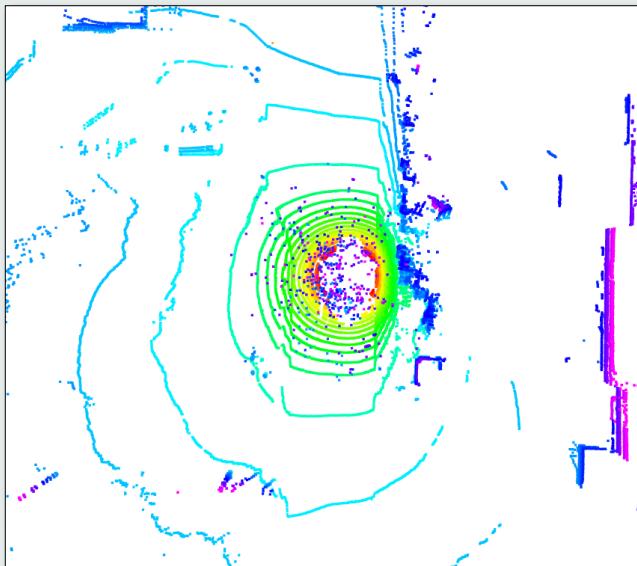
Existing Filters: Statistical Outlier Removal



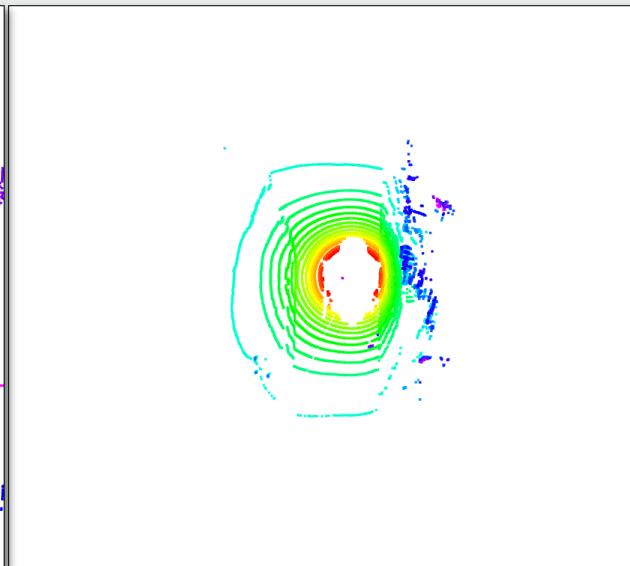
<https://www.youtube.com/watch?v=7zaf4pSe4v4>

Real-Time Removal of Snow in Lidar Data

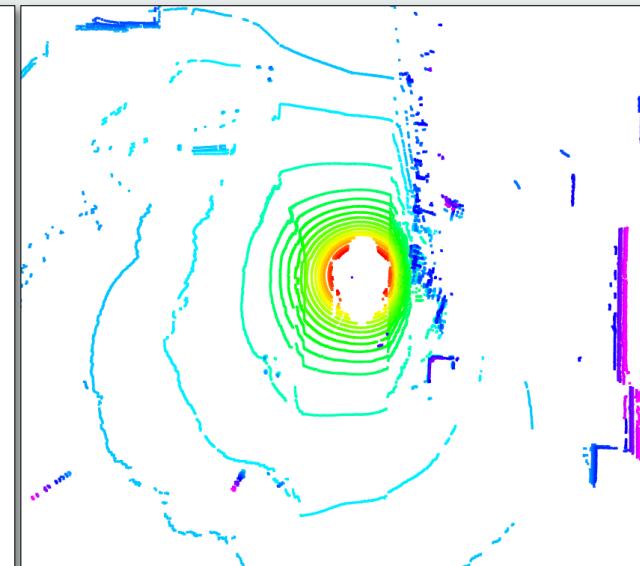
- Results



Raw Point Cloud



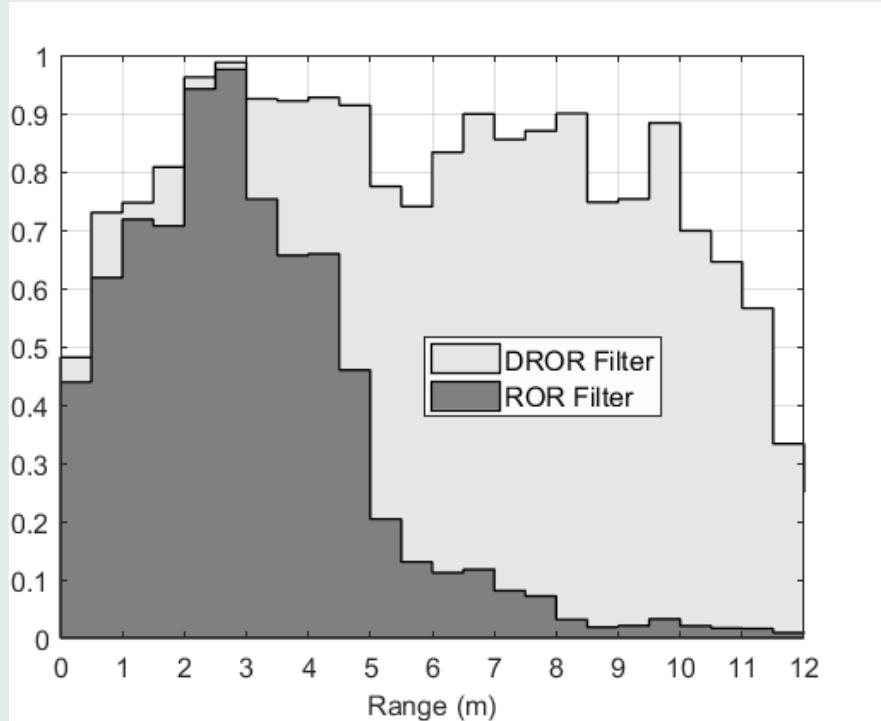
Radius Outlier Removal
(ROR)



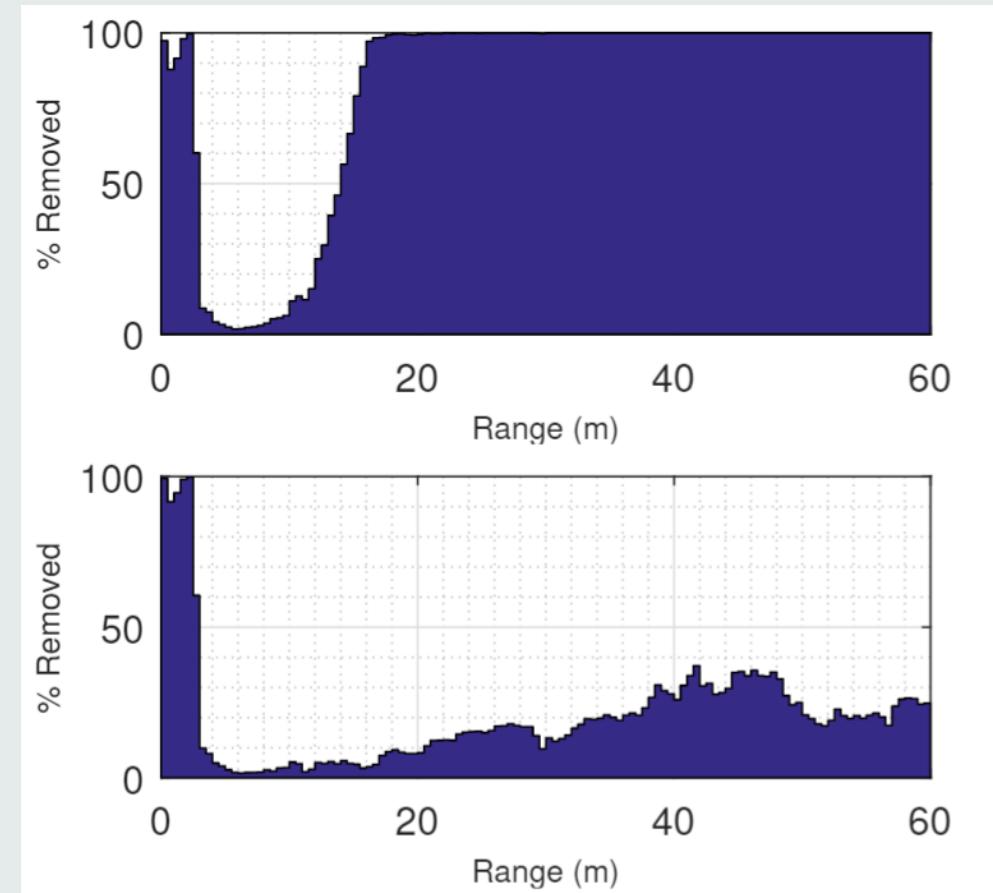
DROR Filter

Real-Time Removal of Snow in Lidar Data

- Results

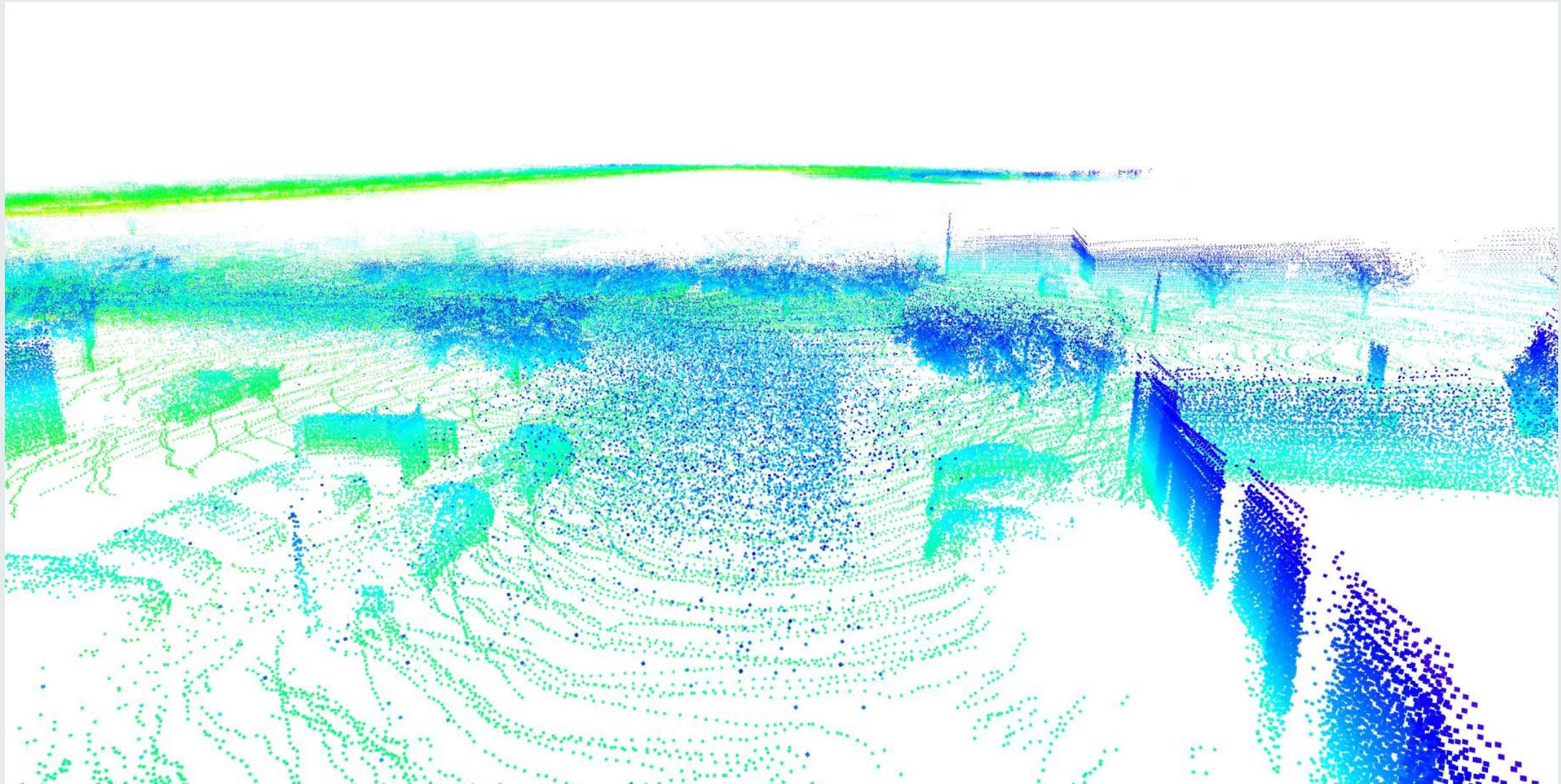


Percentage of points removed which
were labeled as snow

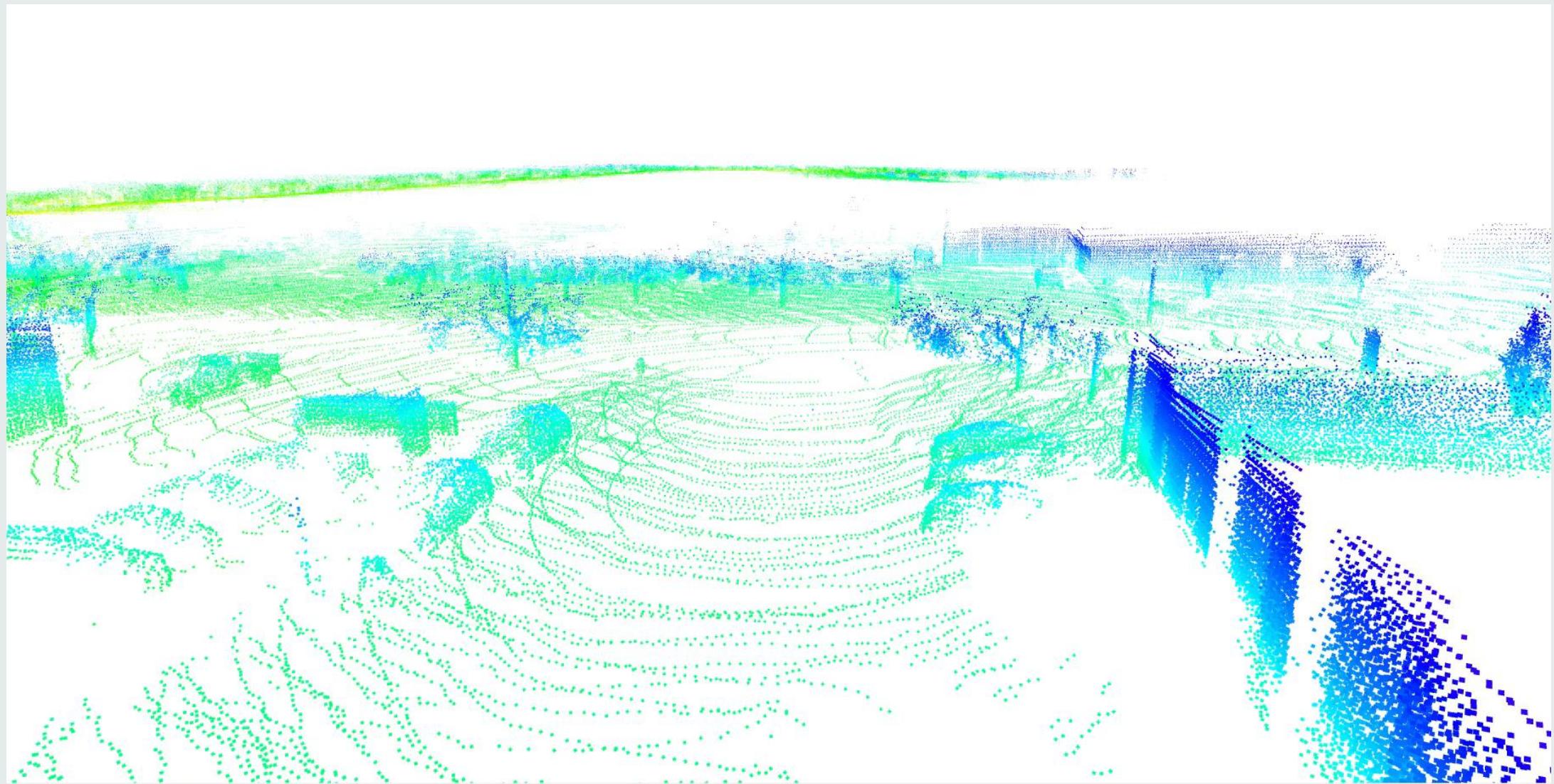


Average Percentage of points removed per
range bin per point cloud for 98 scans (top:
ROR, bottom: DROR)

Mapping with and without snow



Mapping with and without snow



Summary

- Full autonomous driving will soon be available
 - Steady progress toward human level performance
- Canada needs a national policy to permit autonomous vehicles on our roads
 - Ontario's test program hugely successful
 - National standards needed to maintain safe roads and permit growth
- University of Waterloo is at the centre of autonomous vehicles research in Canada
 - Leading the way in all-weather autonomous driving (publicly)
 - Major initiatives in safety assurance, online monitoring, precision planning and control
- Continued funding of science and engineering research essential to grow this position
 - Adoption will be delayed if Canadian problems are not addressed

Thank You



Autonomoose