Towards Fully-automated Driving

Challenges and Potential Solutions

Dr. Gijs Dubbelman
Mobile Perception Systems
EE-SPS/VCA

tue-mps.org

Technische Universiteit
Eindhoven
University of Technology

Where innovation starts

Mobile Perception Systems

tue-mps.org

- **6 PhDs, 1 postdoc, 1 project manager, 2 software engineers**

Automotive Projects

inLane
Lane Navigation Technology

CloudLSVA
Large Scale Video Analysis

VI-DAS

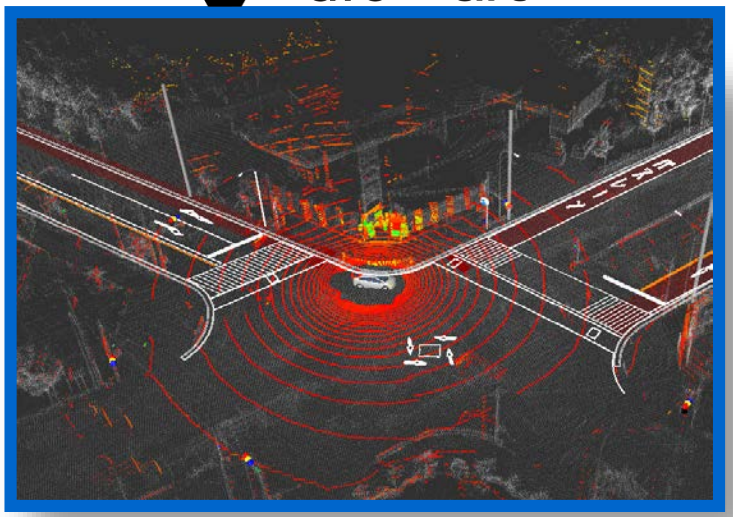
AUTOPILOT

Highly Automated Driving (HAD)

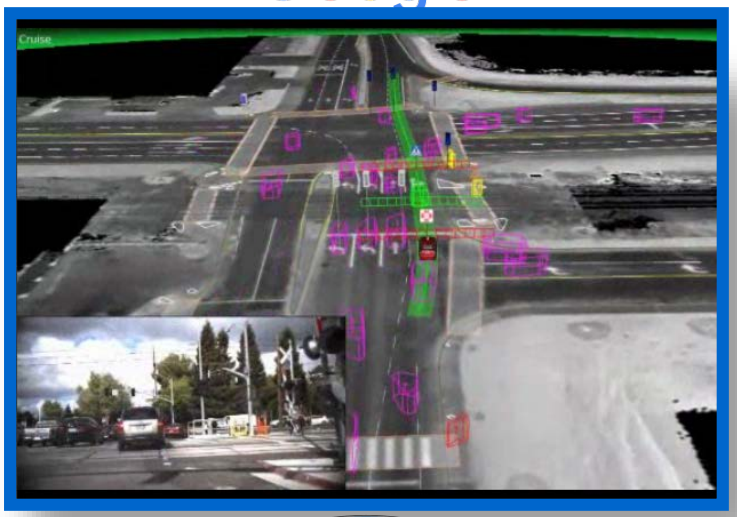
- **SAE level 4 and above**
  - **The driver is no longer the safety fallback**
  - **Including dynamic urban environments**
  - **Including adverse weather conditions**

State-of-the-Art

AutoWare

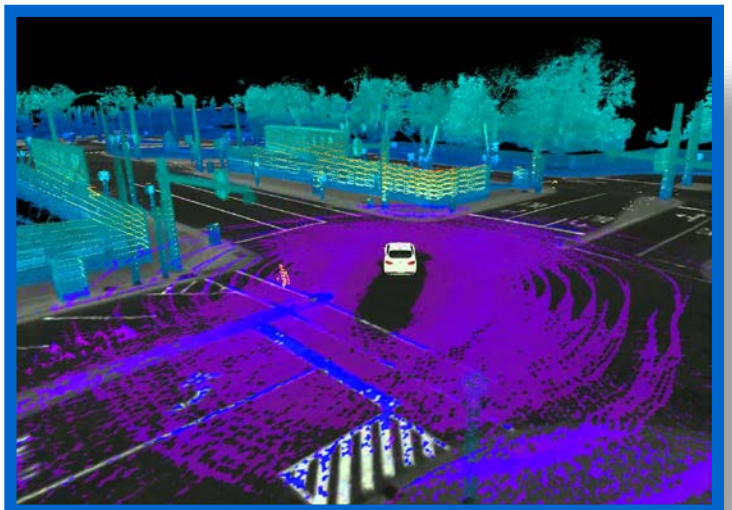


Google



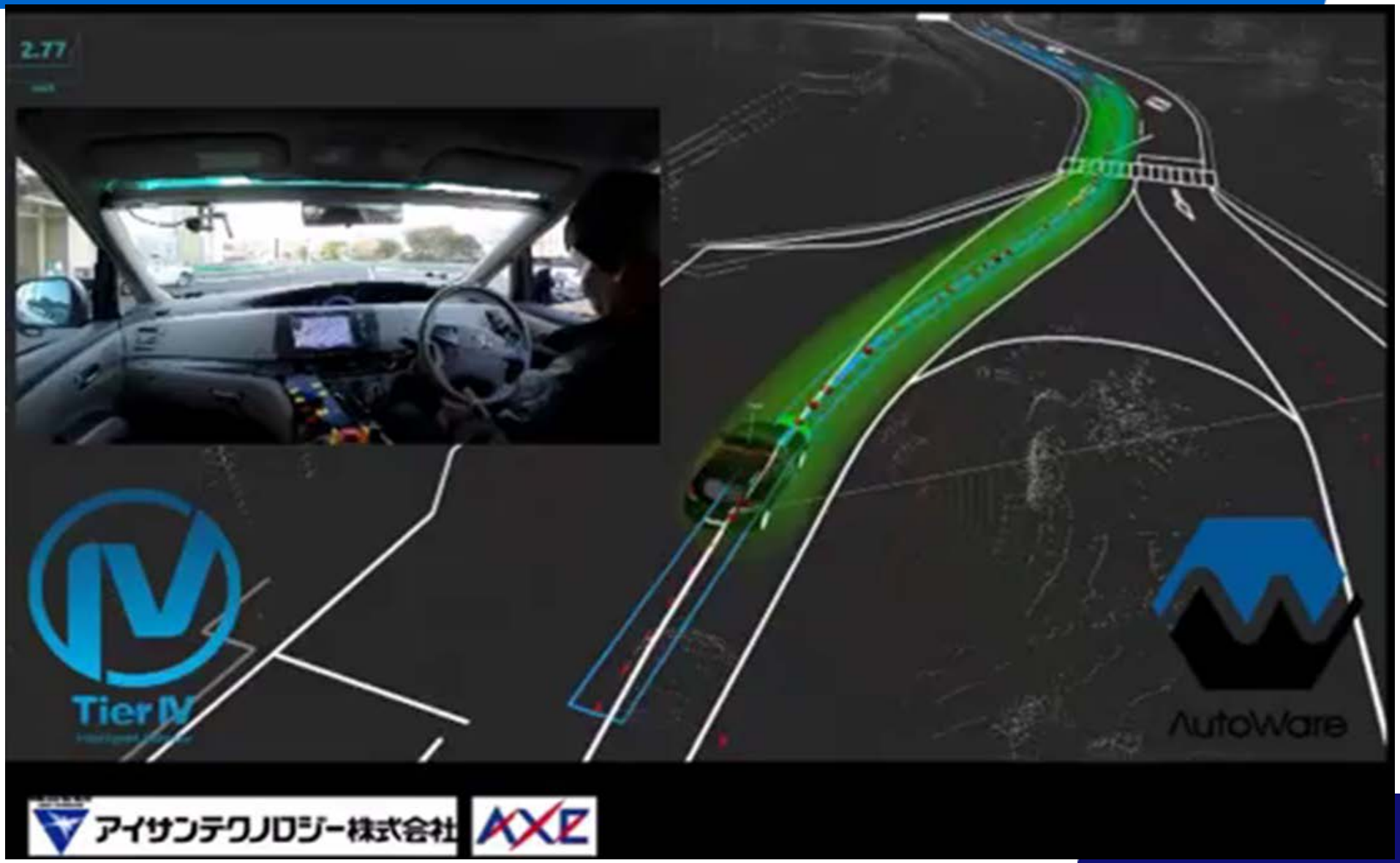
UBER

Ford

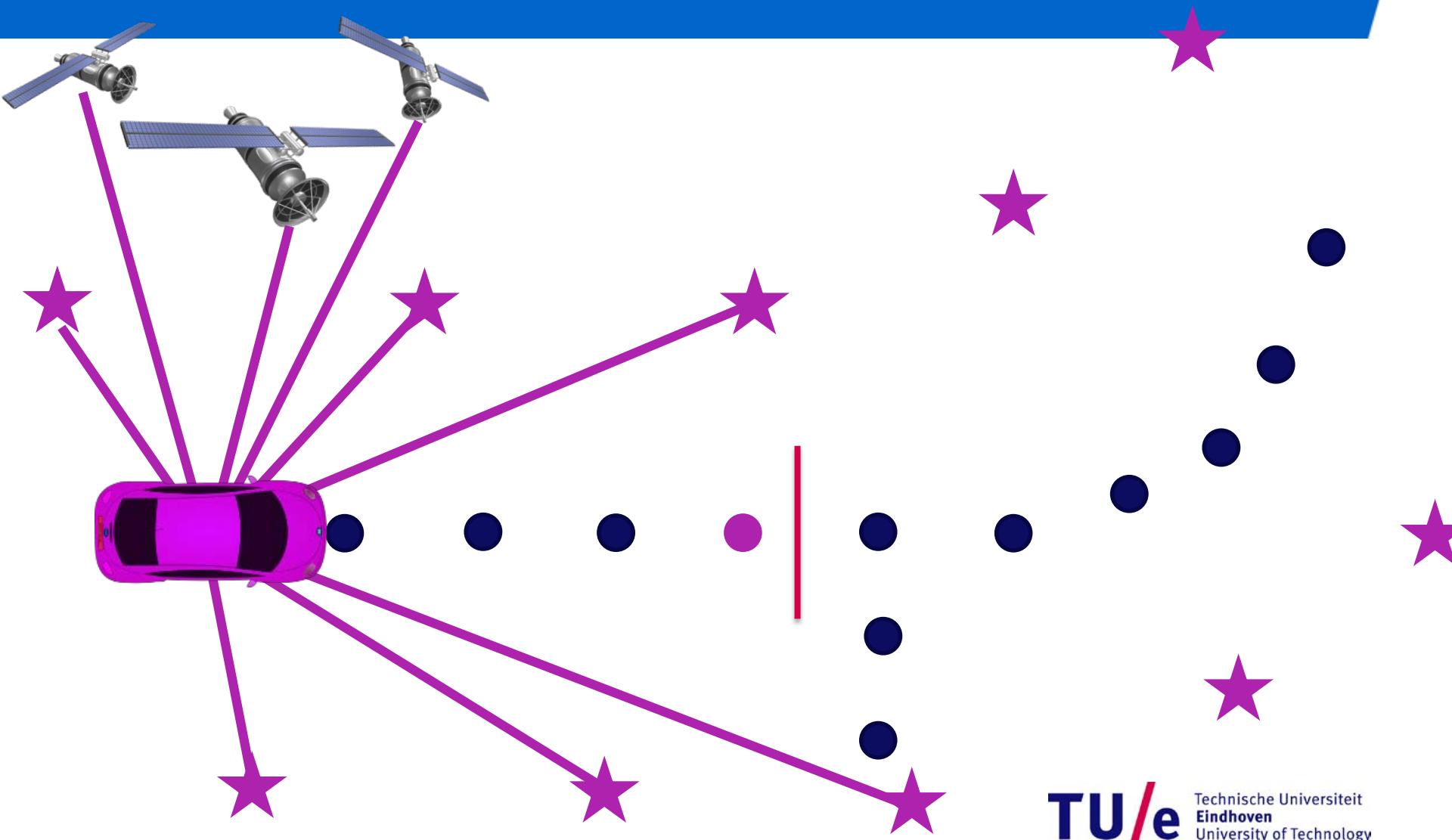


State-of-the-Art Uber

Stare-of-the-Art Autoware



How Does It Work



State-of-the-Art

- **Digital rail road:** nominal vehicle trajectory is pre-defined in the map.
- **Absolute positioning:** RTK-GPS aided by FOG IMU and 3D point-cloud localization
- **Map as a sensor:** road and static object geometry are contained in the map
- **Dynamic objects:** detect and track from vision, RaDAR, and LiDAR

Challenges

- **Scalability:** creating HAD maps for absolute positioning is extremely labor intensive.
- **Timeliness:** frequent updates are required to keep map fresh.
- **Costly hardware:** RTK-GPS, FOGs, and dedicated LiDARS.
- **Limited robustness:** adverse weather and lighting conditions and highly-dynamic traffic scenes

Challenges

SAE Level 4+ & Scalable

?

SAE Level 2
Personally owned
Scalable

SAE Level 4
Robot Taxi
Not scalable

Our Research

1. **Research scalable HAD map technologies**
   - **From absolute to relative positioning**
   - **Distributed Semantic-SLAM techniques to automatically create and update maps**
   - **Crowdsource visual observations**

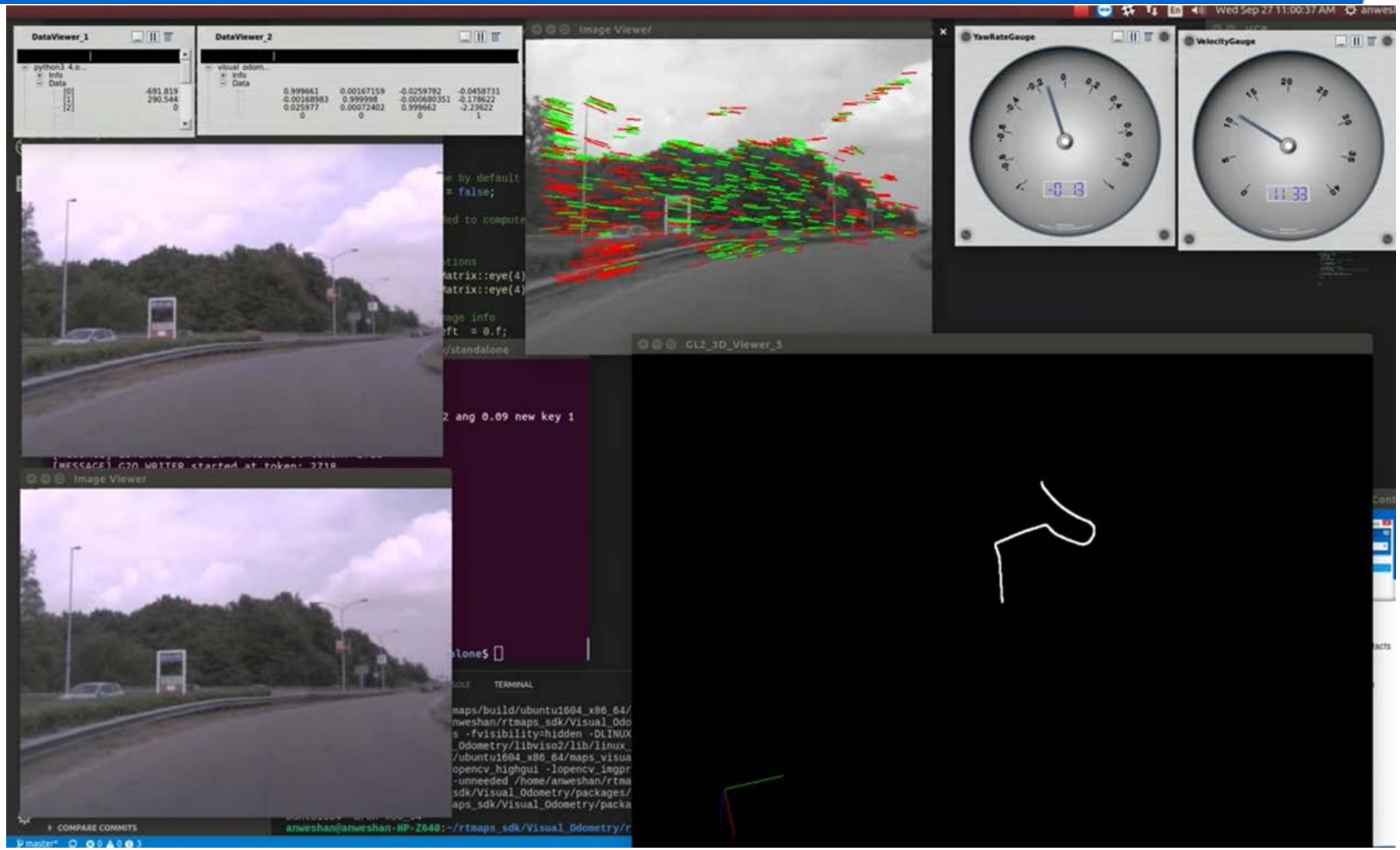
2. **Research AI to make vehicle less dependent on HAD map**
   - **Deep learning and real-time inference**
   - **From pattern recognition to reasoning**

Our Research

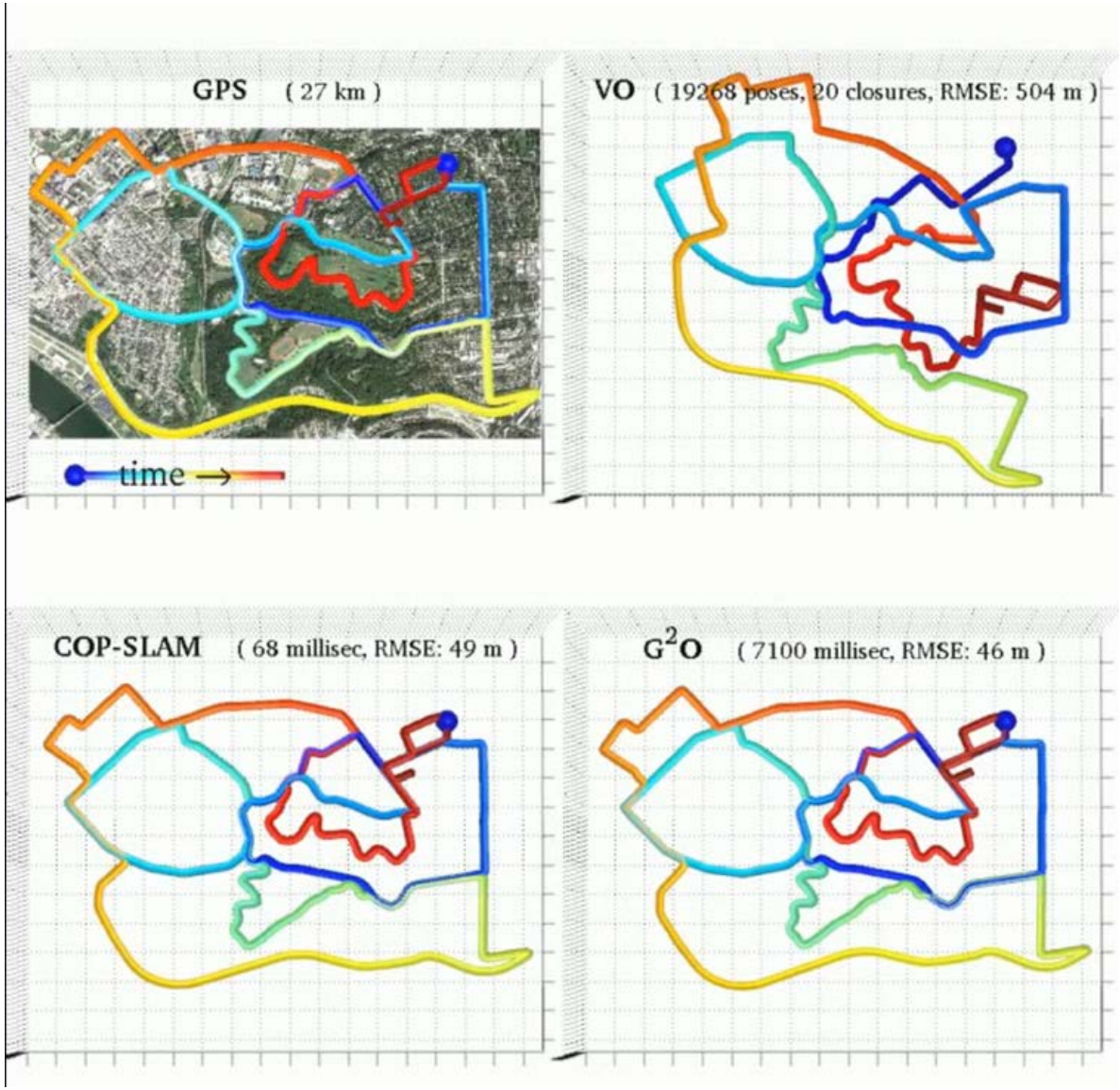
Research scalable HAD map technologies

- Stereo Visual Odometry
- Pose-chain based SLAM
- Pose-graph based GNSS-odometry fusion
- Distributed SLAM using crowdsourcing

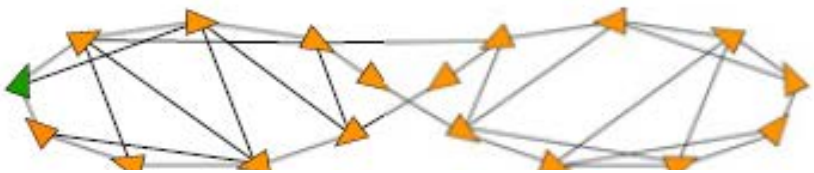
Stereo Visual Odometry

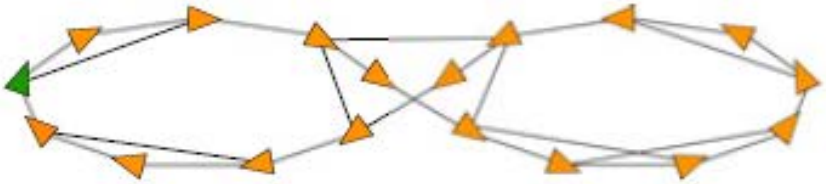


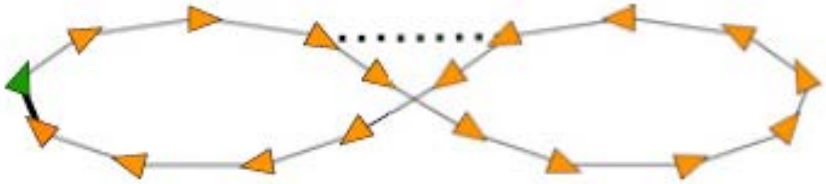
Pose-chain SLAM

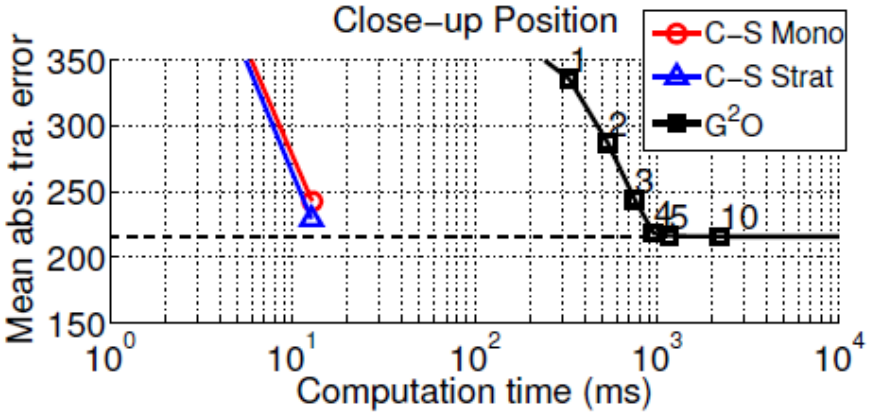


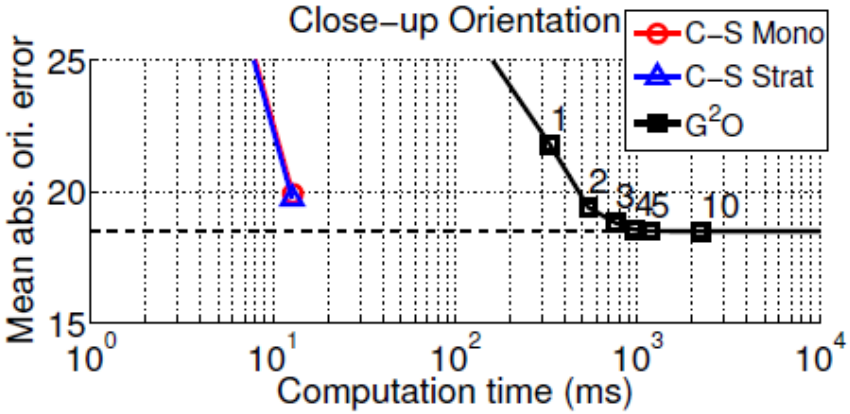
Pose-chain SLAM



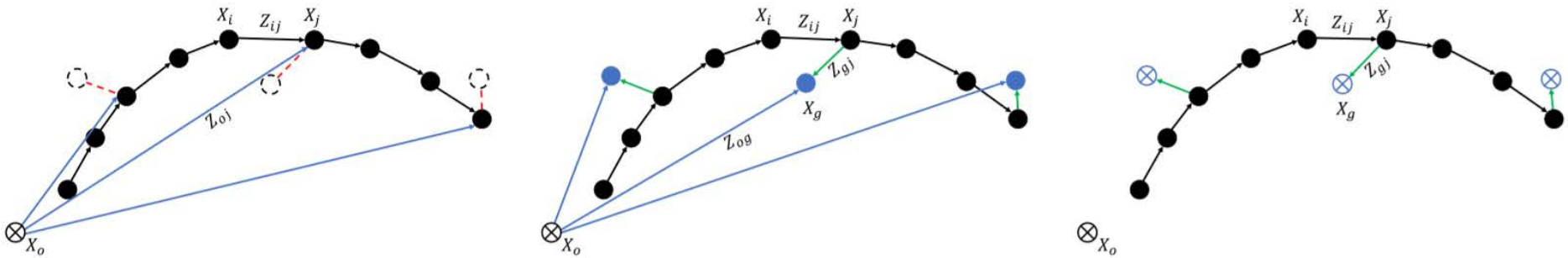


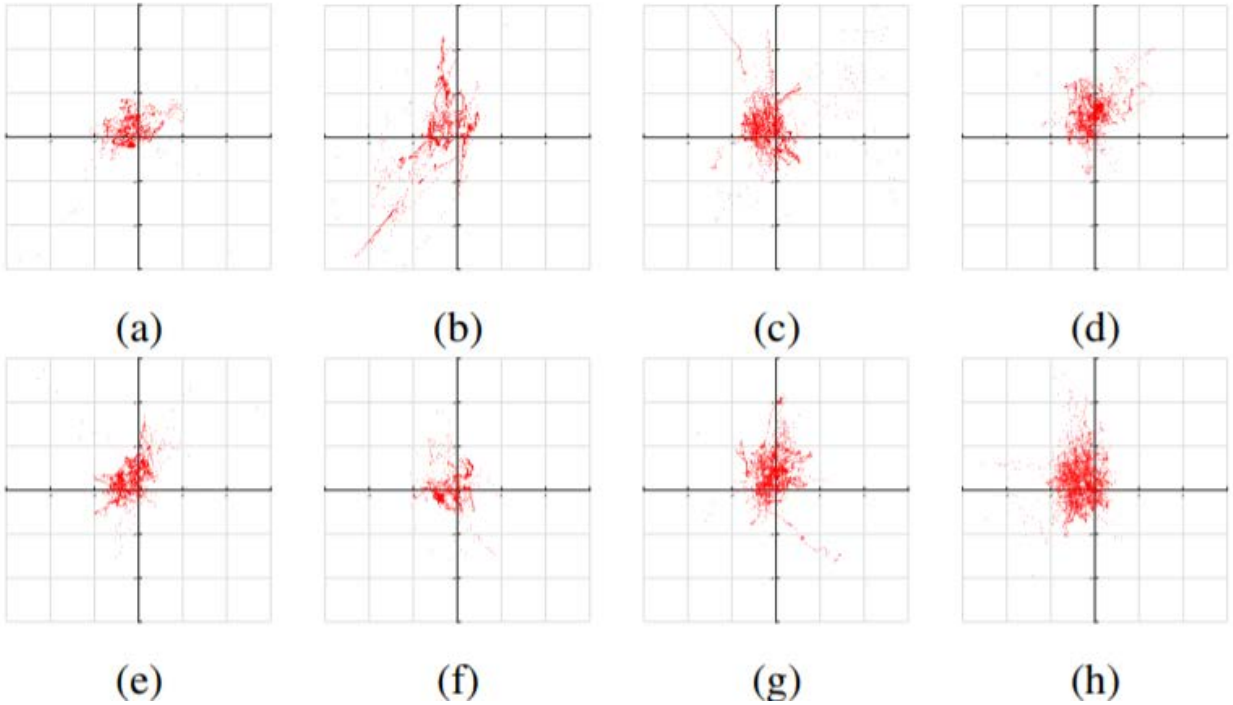




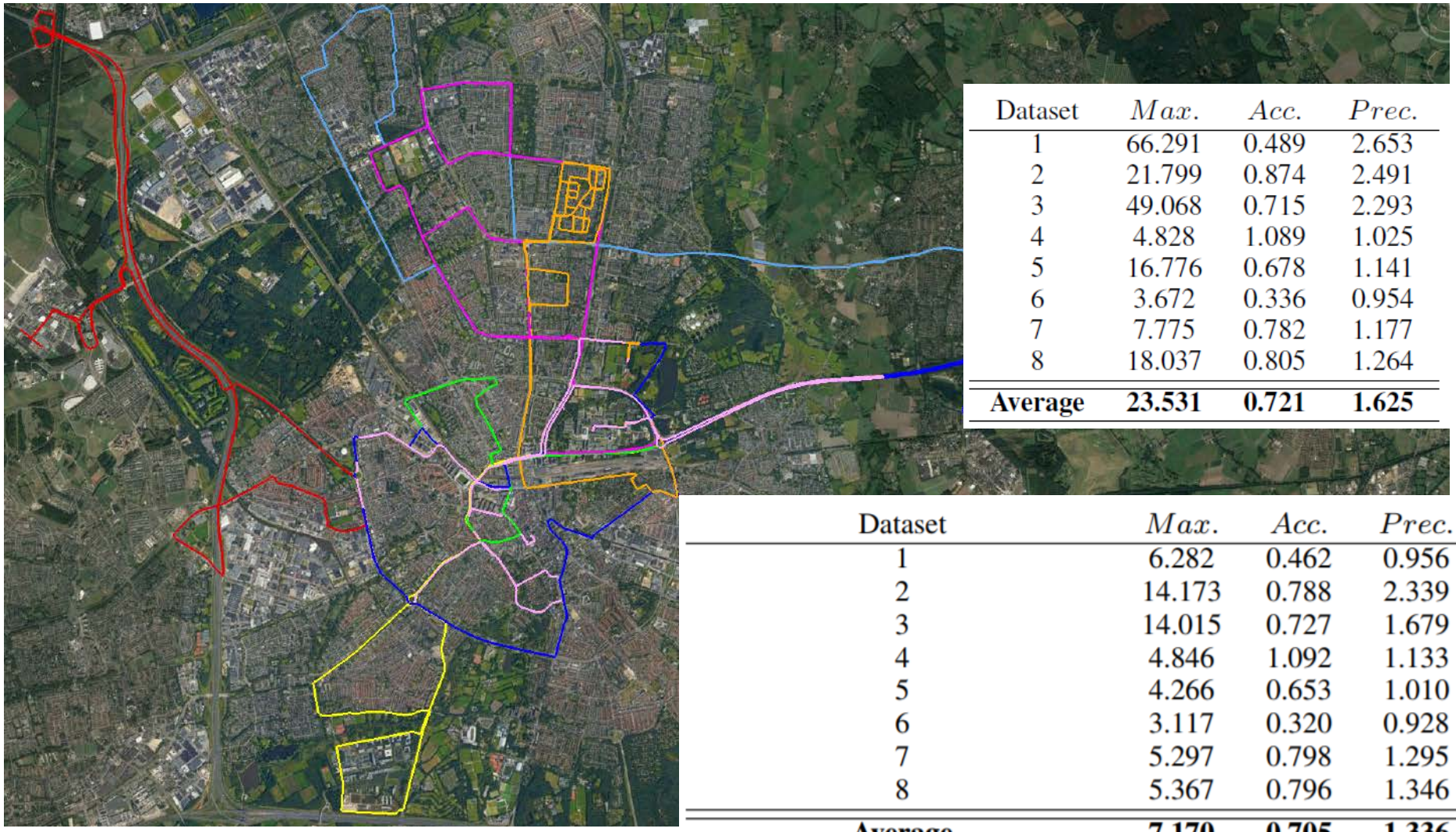


Pose-chain GNNS-Odometry Fusion





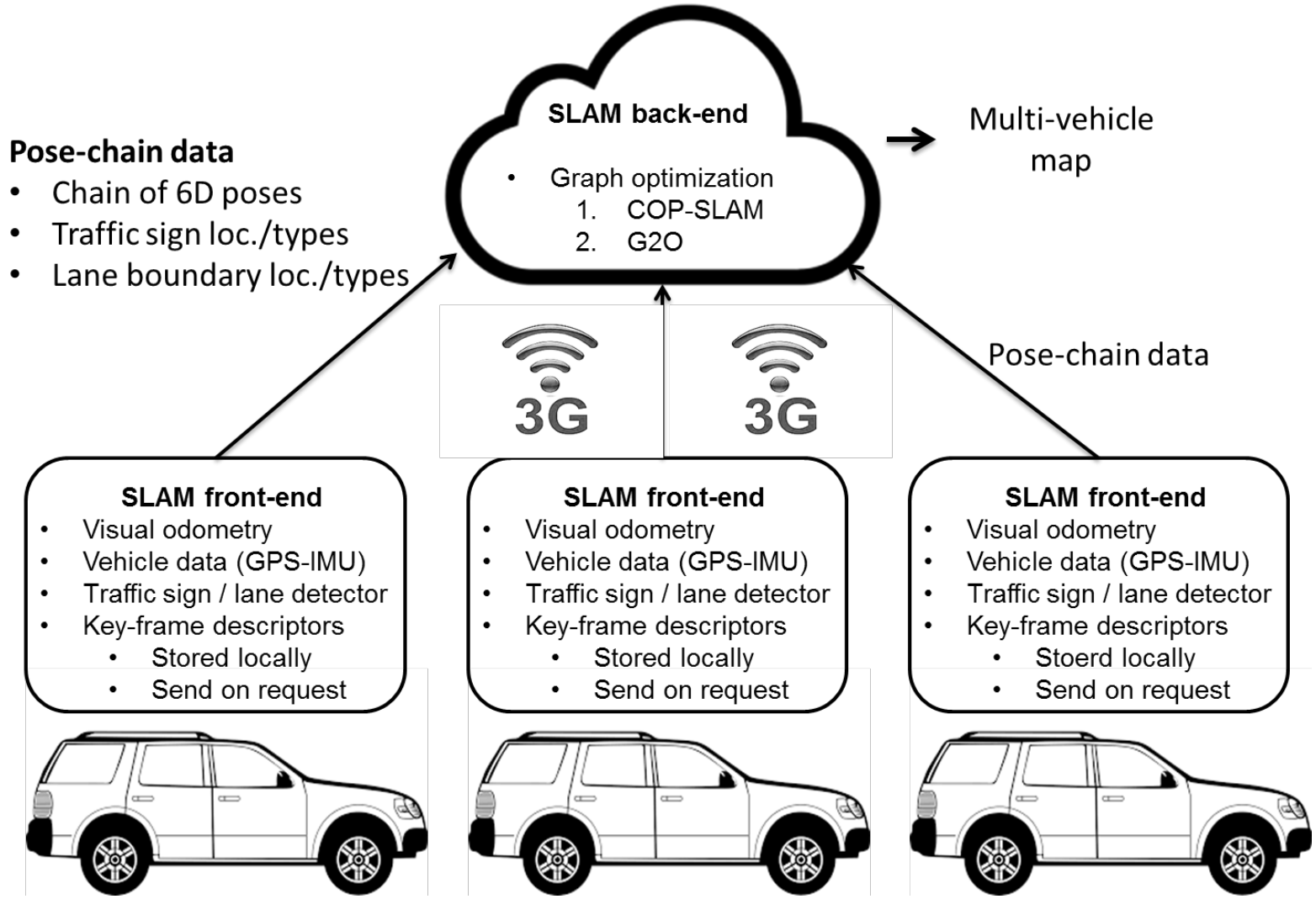
Pose-chain GNNS-Odometry Fusion



Dataset	*Max.*	*Acc.*	*Prec.*
1	66.291	0.489	2.653
2	21.799	0.874	2.491
3	49.068	0.715	2.293
4	4.828	1.089	1.025
5	16.776	0.678	1.141
6	3.672	0.336	0.954
7	7.775	0.782	1.177
8	18.037	0.805	1.264
Average	**23.531**	**0.721**	**1.625**

Dataset	*Max.*	*Acc.*	*Prec.*
1	6.282	0.462	0.956
2	14.173	0.788	2.339
3	14.015	0.727	1.679
4	4.846	1.092	1.133
5	4.266	0.653	1.010
6	3.117	0.320	0.928
7	5.297	0.798	1.295
8	5.367	0.796	1.346
Average	**7.170**	**0.705**	**1.336**
Improvement w.r.t. GNSS (%)	**69.528**	**2.282**	**17.794**

Distributed SLAM using Crowdsourcing



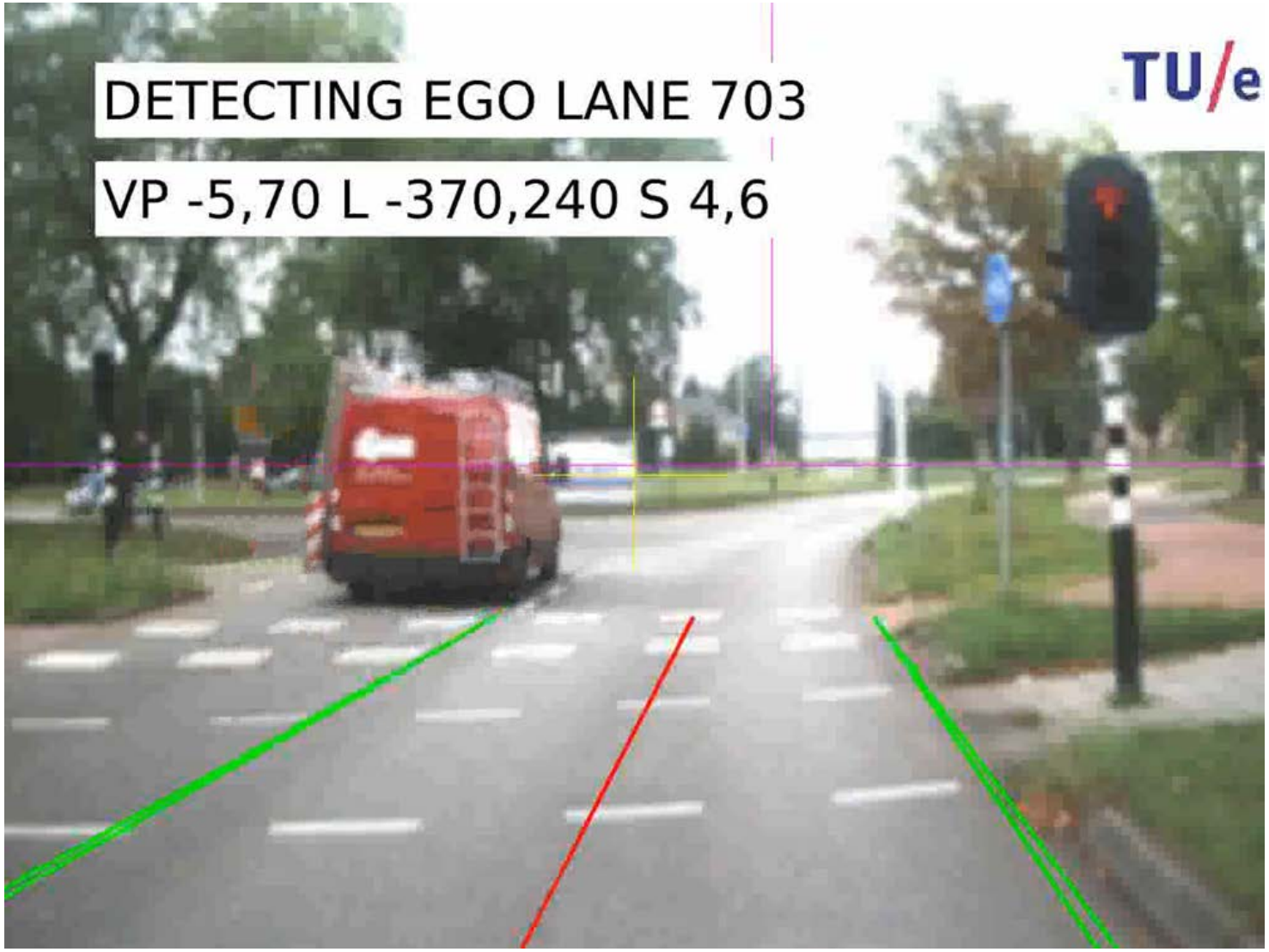
Our Research

Research AI to make vehicle less dependent on HAD map

- **Probabilistic Computer Vision**
  - **Ego-lane boundary detection**
  - **Stereo obstacle detection**

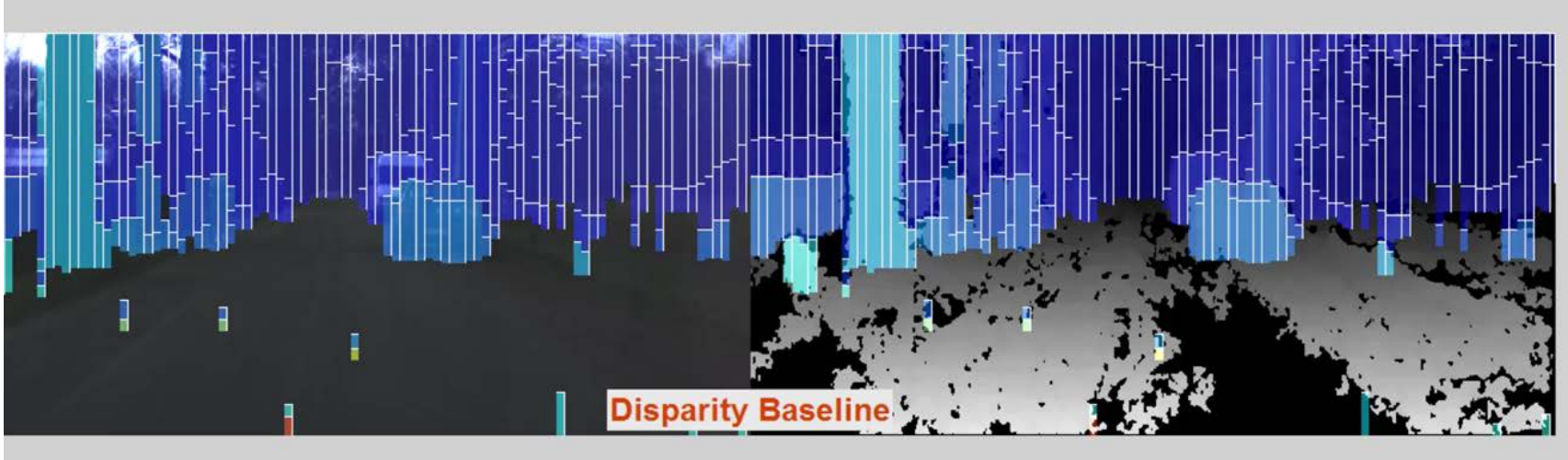
- **Semantic Scene Understanding**
  - **Hierarchical F-CNNs**
  - **End-to-end learning of semantic occupancy grids**

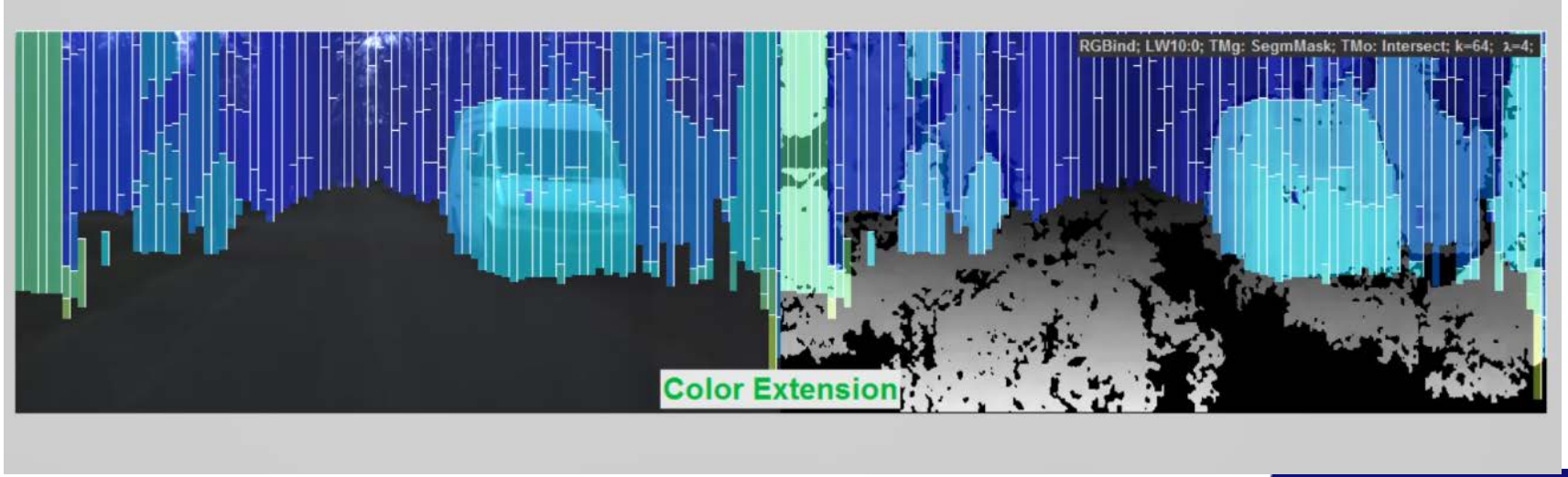
Ego-lane Boundary Detection



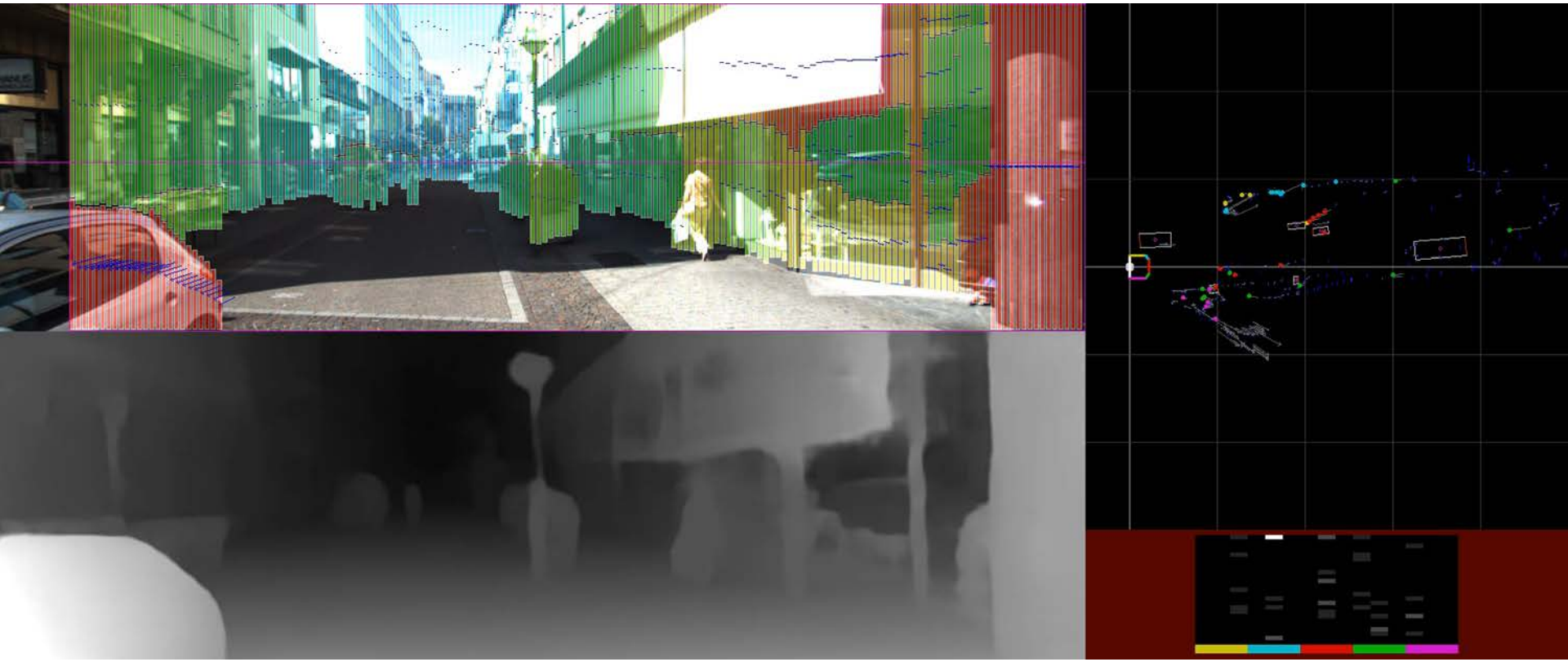
Stereo Obstacle Detection

Stereo Obstacle Detection



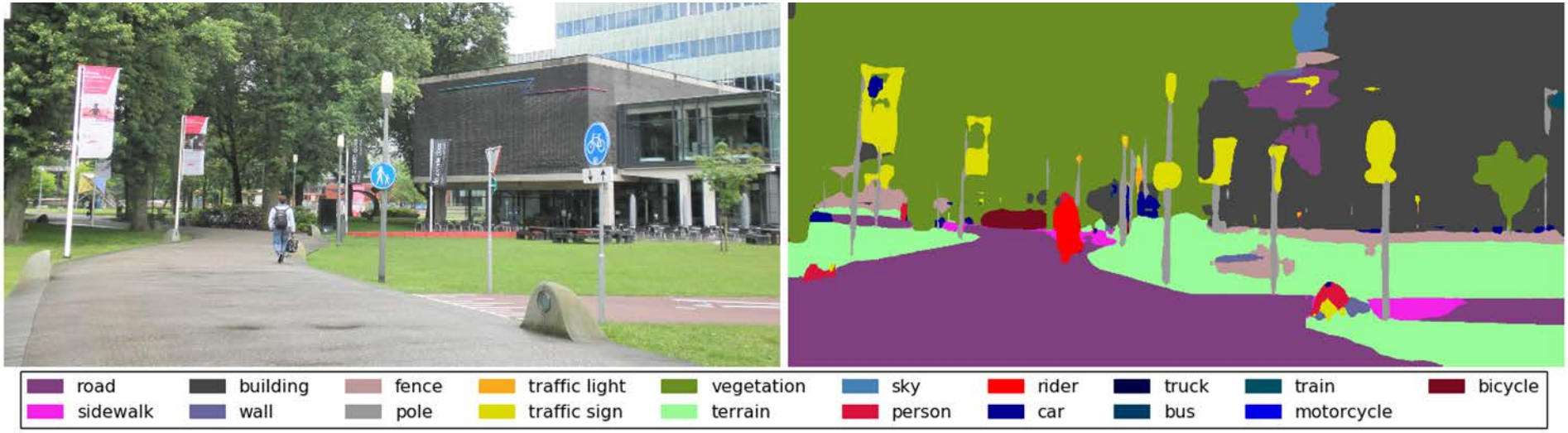


Stereo Obstacle Detection



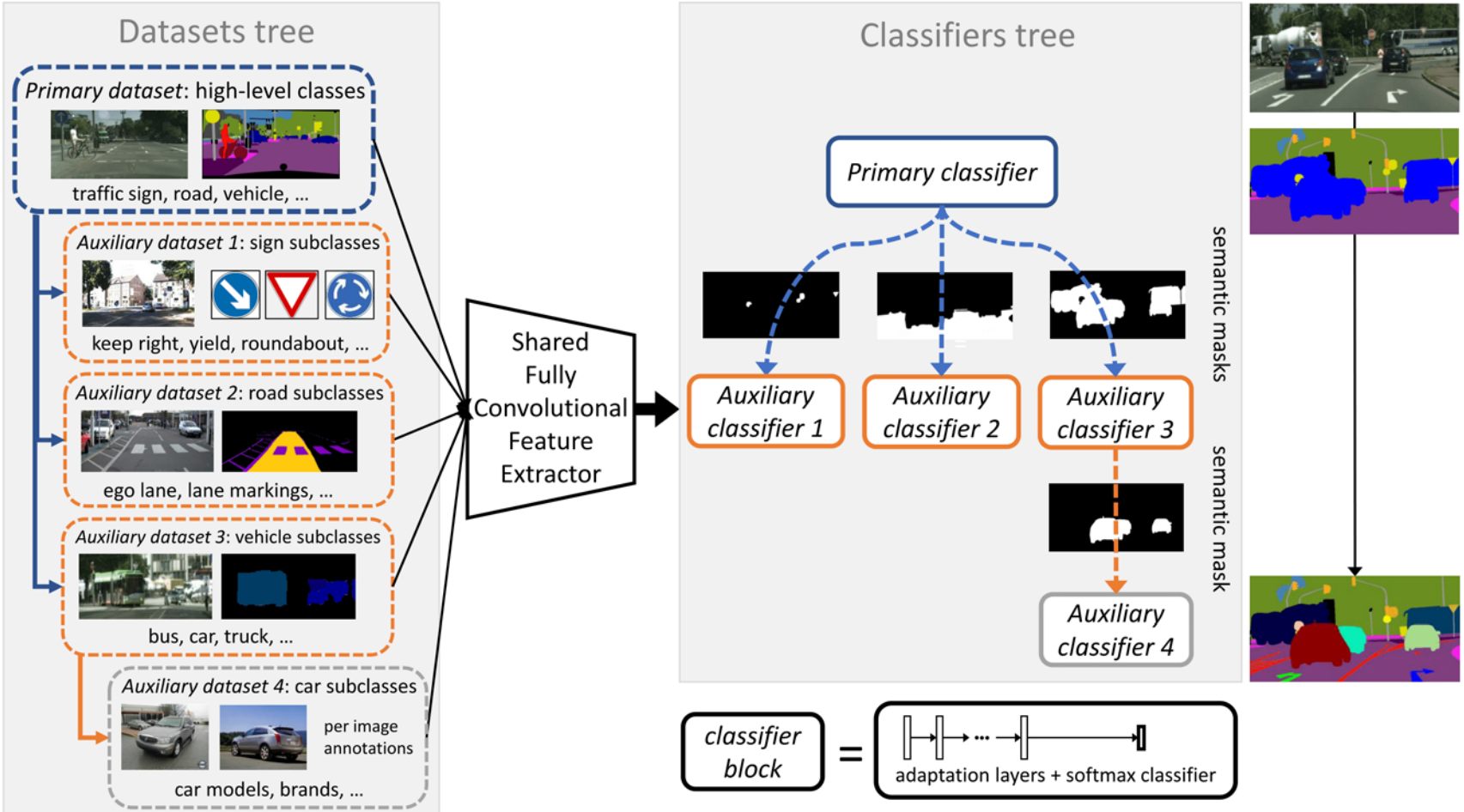
Semantic Scene Understanding

- **Aim: reuse computation for multiple classifiers**



- **Our task: Semantic Scene Segmentation**
- **Goal: Extend the number of classes, without an extra labeling effort and by maximally reusing feature computation layers**

Approach: Hierarchical Network



Recent results

Training on 3-level hierarchy with 3 datasets → inference results for 108 classes

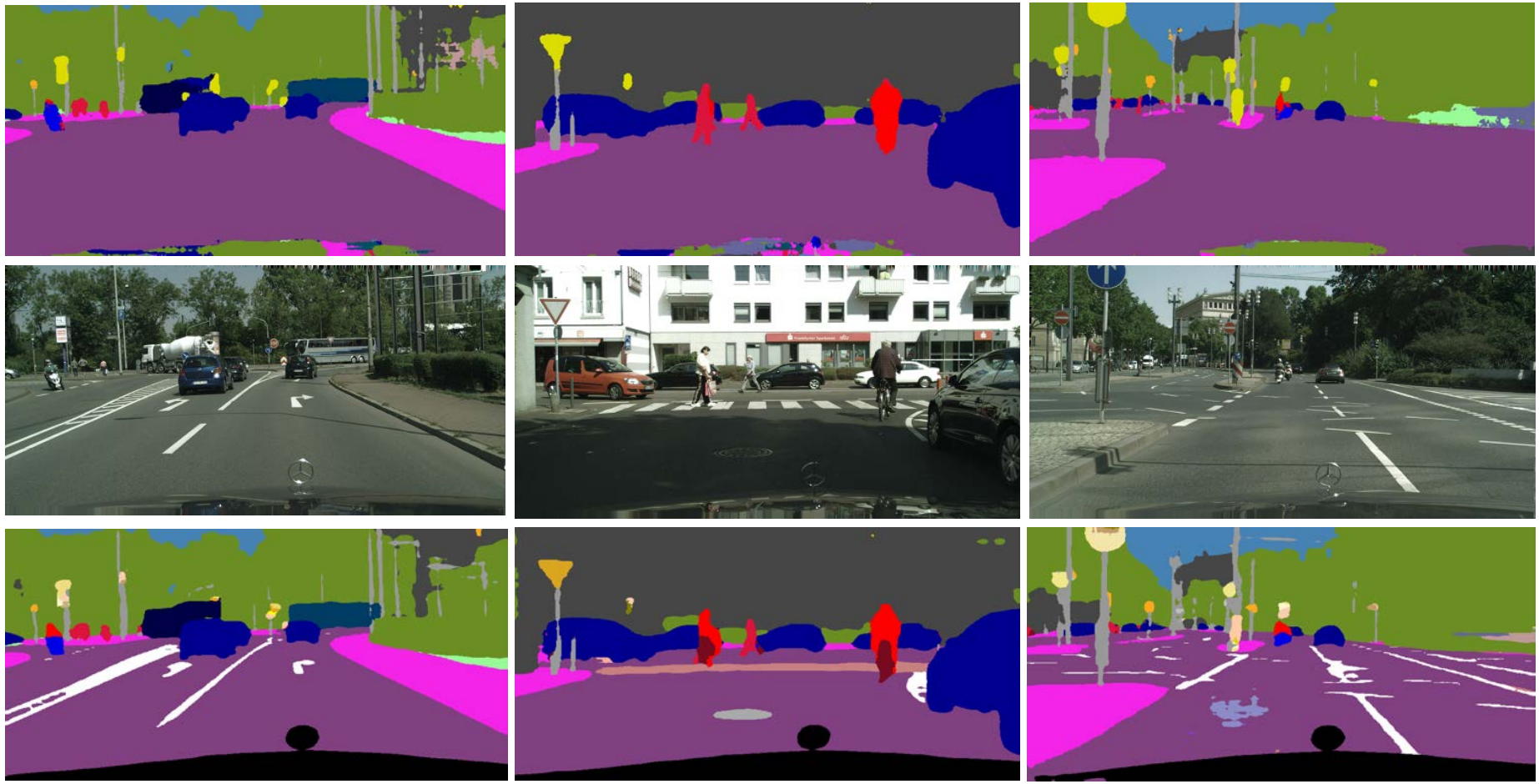
Cityscapes

L2 results

Mapillary Vistas

L3 results

Scene Parsing of108 classes in real-tine



Technische Universiteit Eindhoven University of Technology

Quantitative Comparison

- **Goal: Train traffic sign sub-classes on GTSDB and test on Cityscapes**
- **Compare a flat classifier with our hierarchical classifiers**

Classifier type	Flat				Flat				Hierarchical			
Trained on	$ET_{c,s}$				CT_c and GT_s				CT_c and $GT_{c,s}$			
Tested on	CV_c	GV_c	GV_s	EV_s	CV_c	GV_c	GV_s	EV_s	CV_c	GV_c	GV_s	EV_s
Mean Accuracy (%)	69.4	59.2	7.8	23.0	76.3	98.1	52.1	2.0	75.0	96.0	69.6	39.8
Mean IoU (%)	60.4	-	5.4	12.7	67.0	-	42.4	-	65.2	-	57.1	31.7

The hierarchical classifier approach does better than a flat classifier approach, even when the flat classifier is trained on the target dataset

- **Nvidia TitanX GPU**
- **Tensorflow**

- **52 ms @ 512 x 706**
- **~20 fps**

End-to-end Occupancy Grids

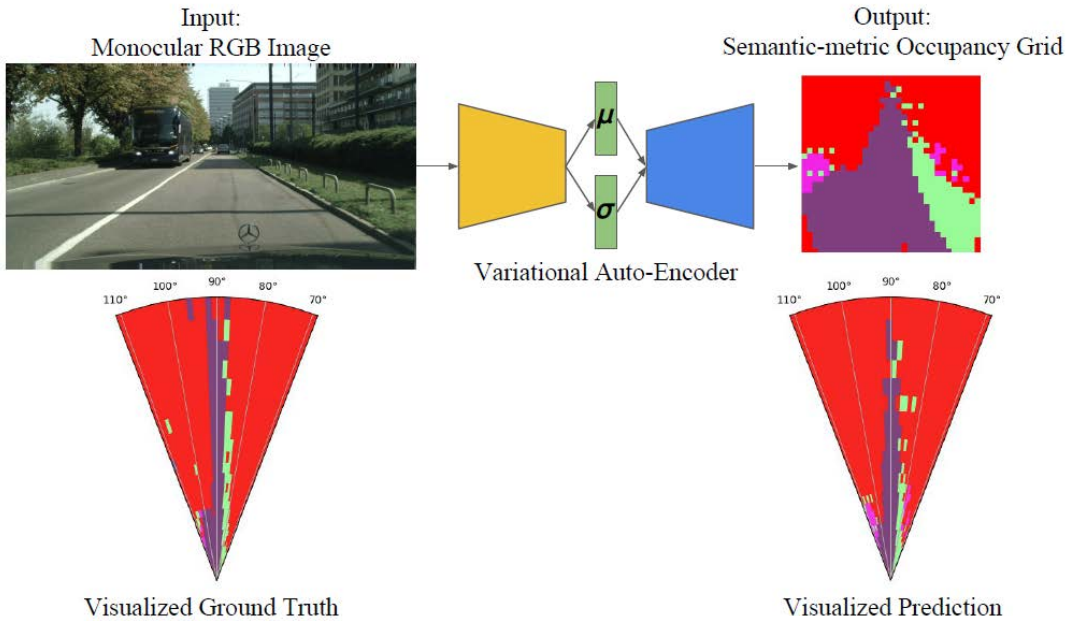


TABLE I

QUANTIFIED PERFORMANCE FOR DIFFERENT MAPPING METHODS.

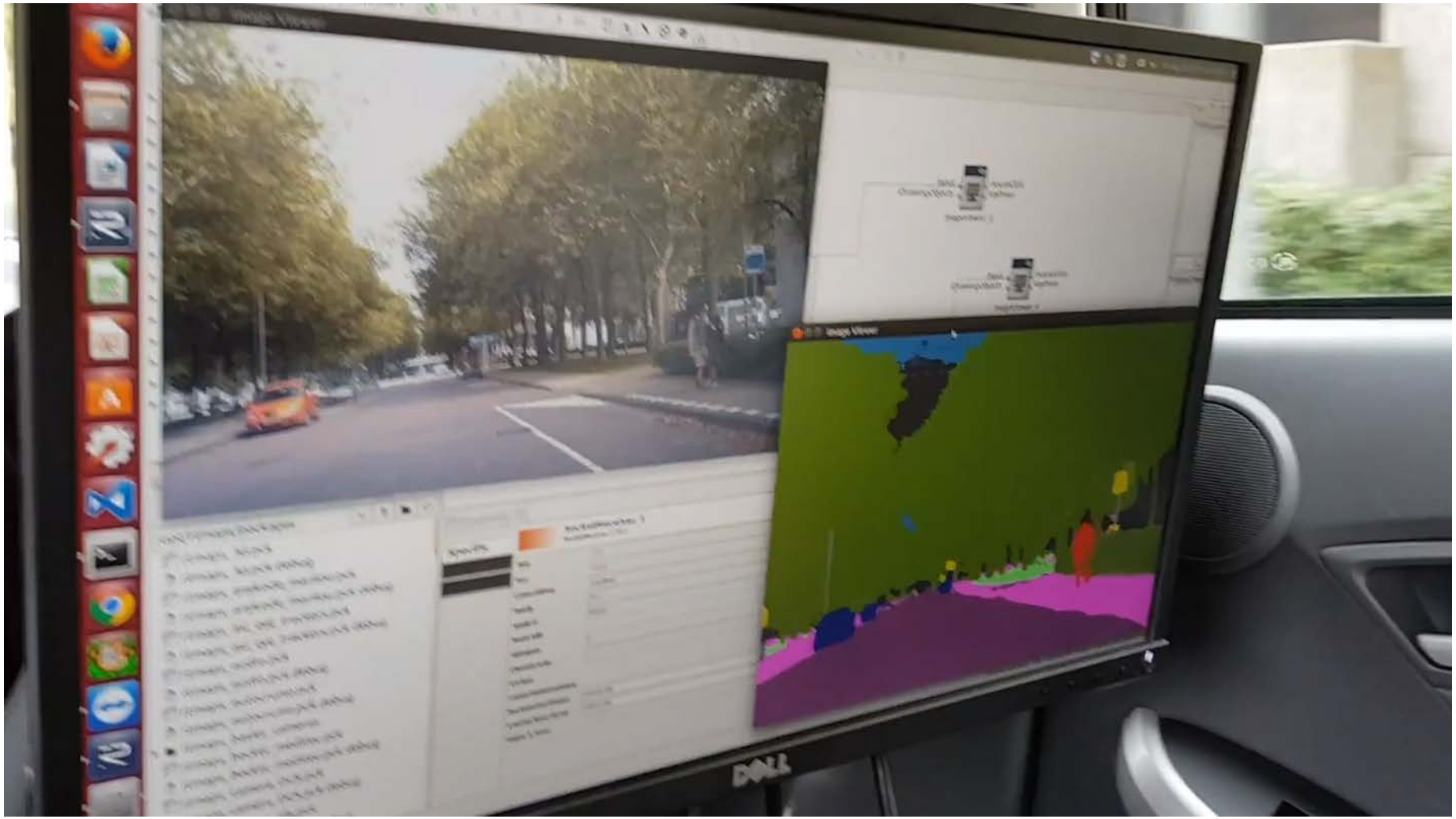
Method	mean IoU	f.w. IoU	disparity	camera calibration	RGB image
With Disp.	80.8	93.2	✓	✓	✓
Fixed-plane Assum.	49.2	63.4	-	✓	✓
Ours	**59.0**	**76.3**	-	-	✓

TABLE II

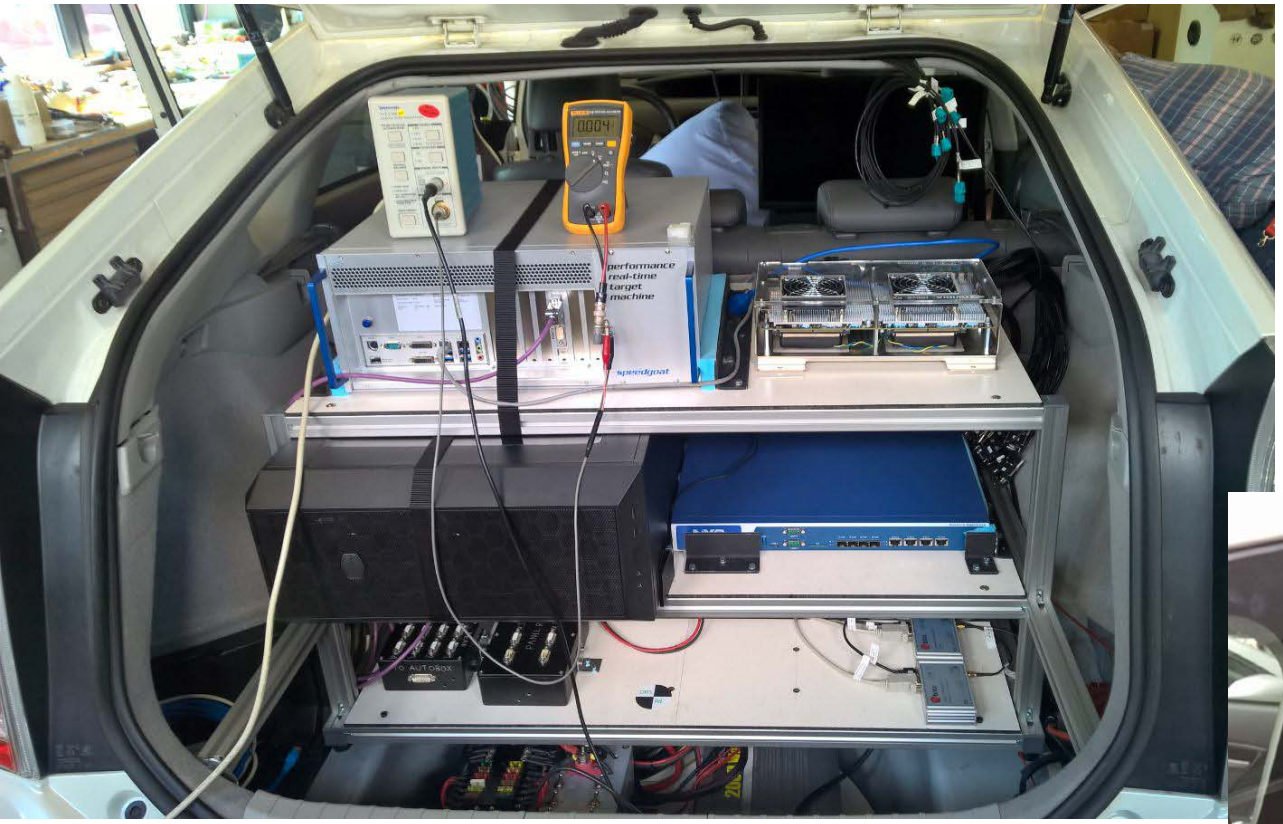
ROBUSTNESS EVALUATION W.R.T. VEHICLE LOCAL DYNAMICS.

Perturbation	mean IoU	f.w. IoU
No perturbation	59.0	76.3
± 1.5° pitch	54.5	71.8
± 3.0° pitch	45.1	61.5
± 5° roll	54.1	71.9
± 10° roll	45.8	63.6

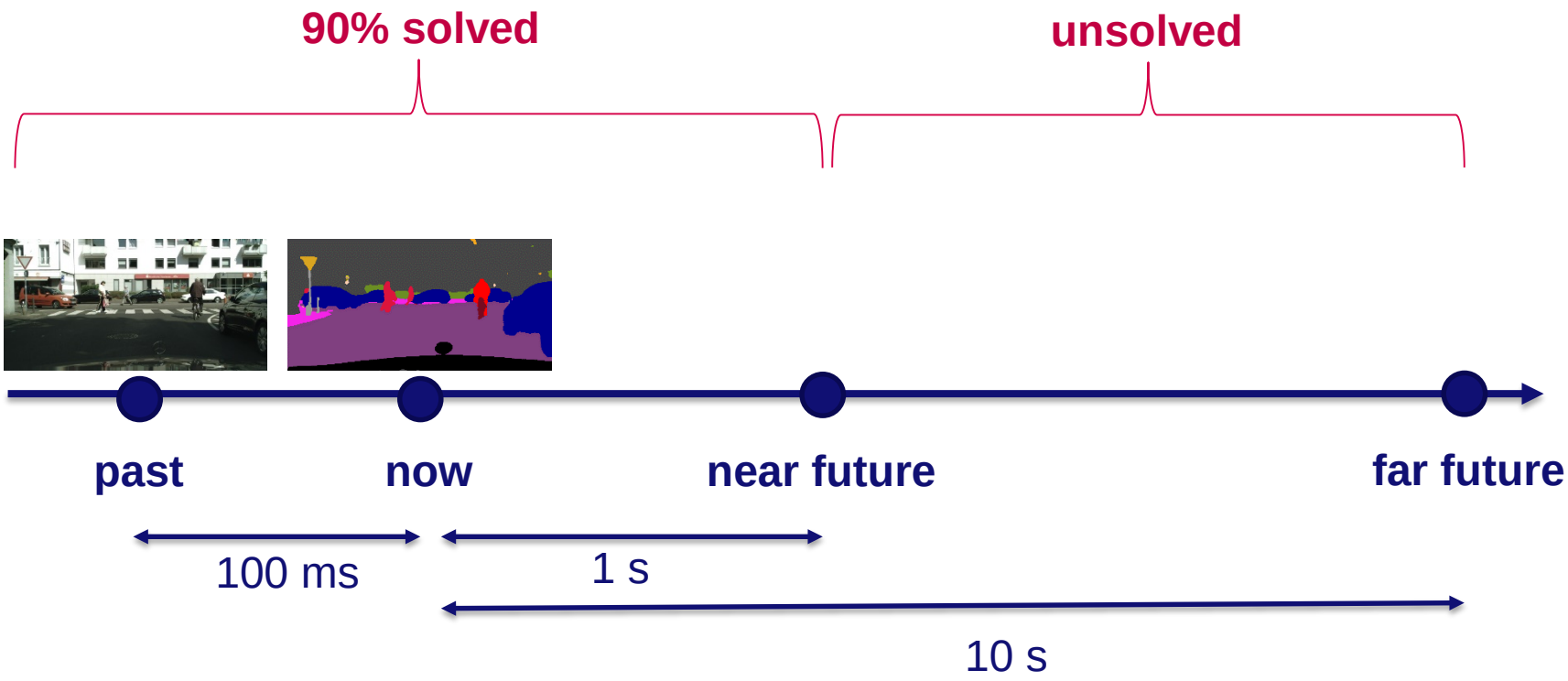
TU/e MPS Research Vehicle



Vehicle Integration



Future Challenges



Anticipatory systems: Anticipate on possible future events from sensory data from the past

Future Research Challenges

Extending the prediction horizon for the behavior of road users in highly dynamic urban traffic situations from 1 second to several seconds.

- **Too complex for rule-based approaches**
- **Too few training samples, due to combinatorics, for learning-based approaches**

Requires novel approaches for spatiotemporal and inter-object reasoning

AI Paradox

- **Majority of adults:**
  - **Cannot beat the Chess world champion**
  - **Can drive a car safely**
- **Artificial Intelligence**
  - **Can beat the Chess world champion**
  - **Cannot drive a car safely**

Driving a car safely is more challenging than beating the Chess world champion

Conclusion

- **Current approaches for automated driving either do not reach SAE level 4 or are not scalable**
- **We require scalable HAD mapping technology**
  - **Crowdsourcing with Distributed SLAM**
- **We require improved AI that can anticipate**
  - **From pattern recognition to reasoning**

Questions