NVIDIA

NVIDIA DEEP LEARNING DAY 2017

ADAS & AD

October 31th, 2017

PROSENSE

Drive for Excellence!

ADAS : Advanced Driver Assistance System
AD : Autonomous Driving

ProSense : www.prosensetek.com

SEJONG
MS AUTOTECH CO.,LTD. 엠에스 오토텍
SEYANG Mechatronics Co.,Ltd
TAEWOOL
CVS COMPUTER VISION SYSTEMS
KETI
NVIDIA
INNOSIMULATION

개요

- 알고리즘 : ADAS & AD
- 비즈니스 모델

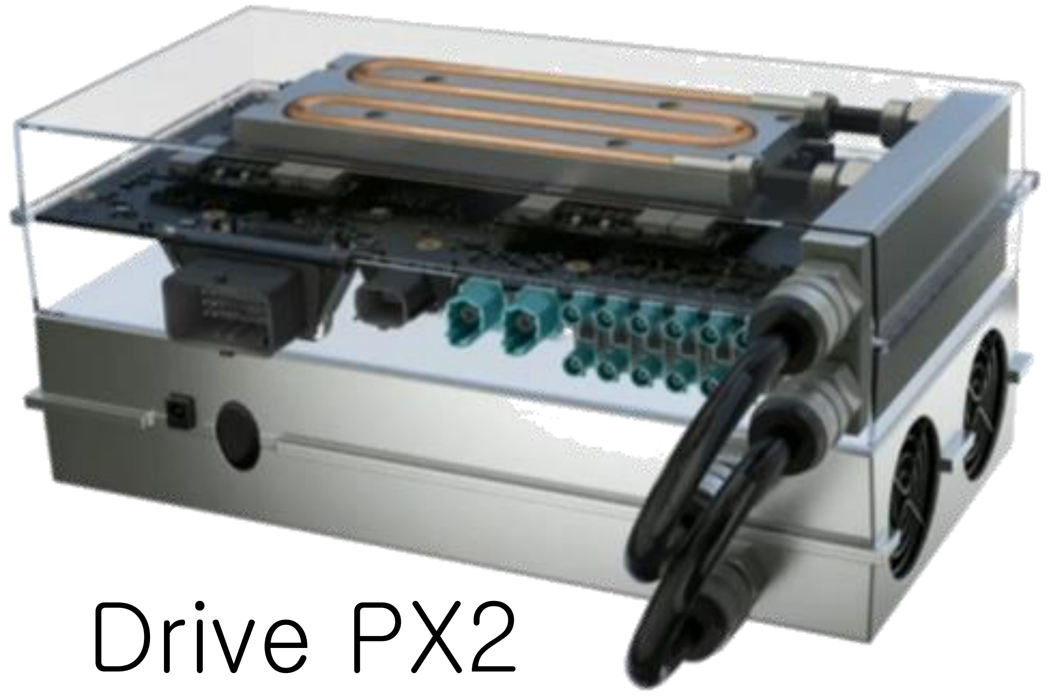
그래픽 기반 자율주행 시뮬레이터

자율주행차 개발의 全 영역(데이터 획득, 데이터 처리, 응용개발, 검증)에 시뮬레이션 방법론을 적용함.

개발 Cycle	**DATA ACQUISITION** Sensor integration Sensor calibration Image capture Storage	**DATA CURATION, LABELING** Manage data lifecycle **DNN TRAINING** Train NN with data **MAP CREATION** Create HD map	**APPLICATION DEVELOPMENT** Compute libraries NN Inferencing System level applications	**SIMULATE, DRIVE, TEST** HIL/SIL simulation Vehicle integration Test
실차량	▪ 센서 데이터 획득 ▪ 센서융합 ▪ 데이터 저장	▪ Dataset 마련 ▪ 딥러닝 학습 ▪ 지도 제작	▪ AD SDK ▪ CUDA ▪ 알고리즘	▪ H/W : Drive PX2 ▪ S/W : DriveWorks ▪ TensorRT, cuDNN
시뮬레이션	▪ 가상 센서 데이터 ▪ 가상 센서융합	▪ 가상 Dataset 자동생성 ▪ 딥러닝 학습 ▪ 가상 지도 자동생성	▪ AD SDK ▪ CUDA ▪ 알고리즘	▪ H/W : Drive PX2 ▪ S/W : DriveWorks ▪ TensorRT, cuDNN

Test Platform : Real-time Perception, Self-driving

An SUV for real-time perception and a self-driving car including control algorithms



Drive PX2

Multi-Camera

Multi-Radar

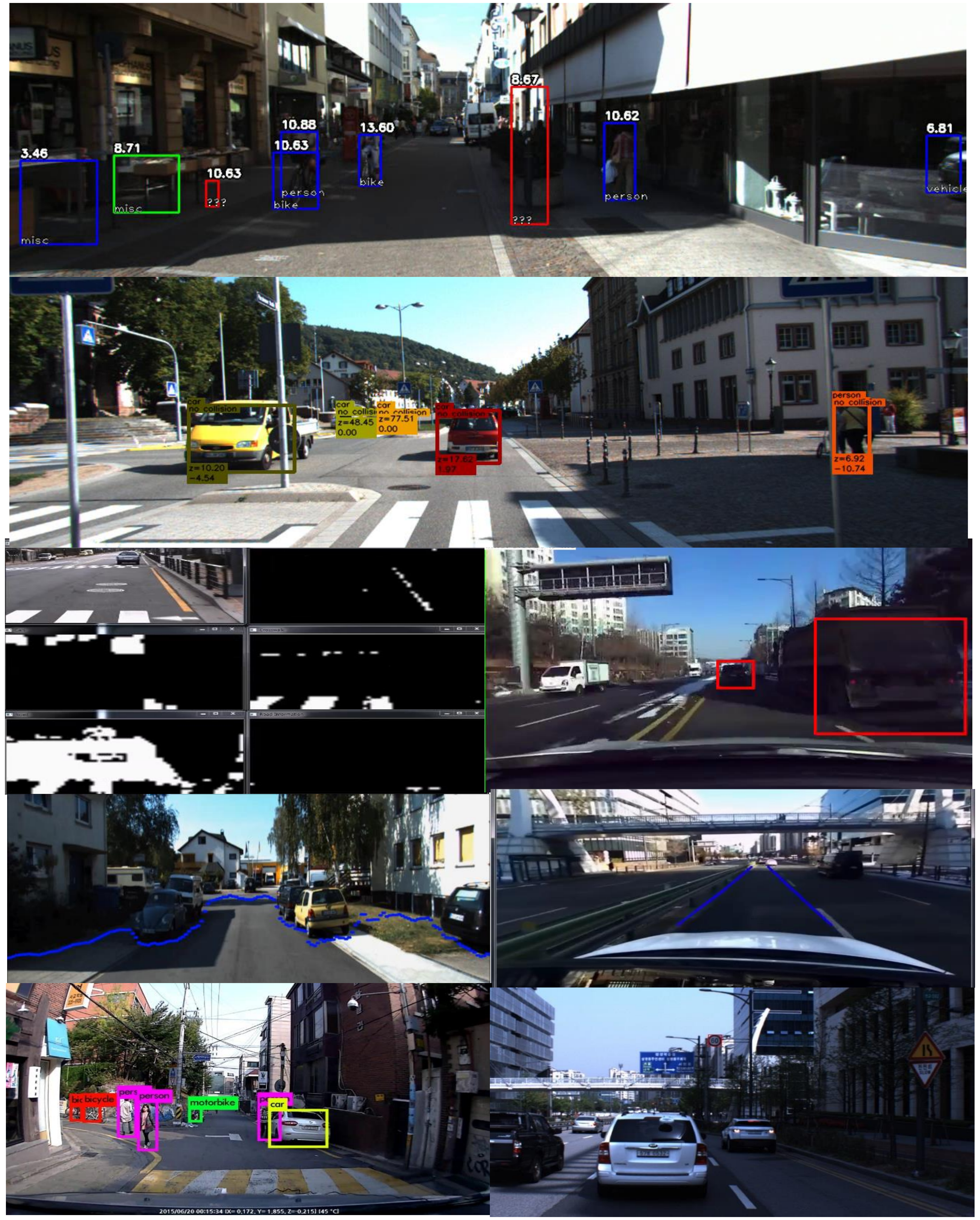
ADAS Directions : Algorithms Licensing

- 3D Lane & Road, Radar Fusion
- High Accuracy AEB : Depth/VD/PD
- Vision Cruise : High Curvature
- Robustness

AD (Autonomous Driving) Directions

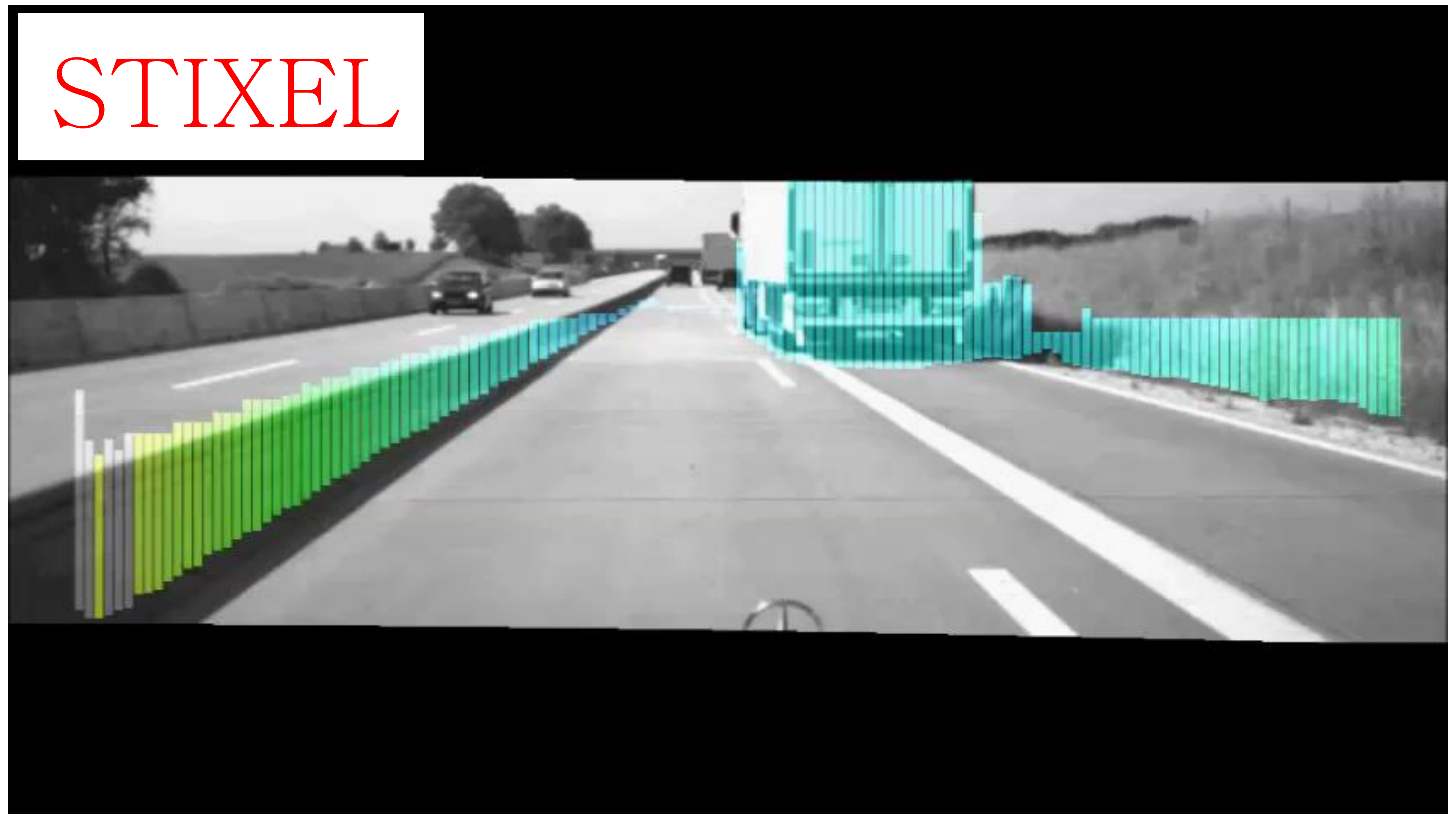
- Perception : Real Time High Accuracy
- Better Control by better perception
- Service : Last Mile Mobility
- Simulation based AD Development

Algorithms : VD, PD, LD, 6D, Tracking, CNN-OD, Stixel, SVM



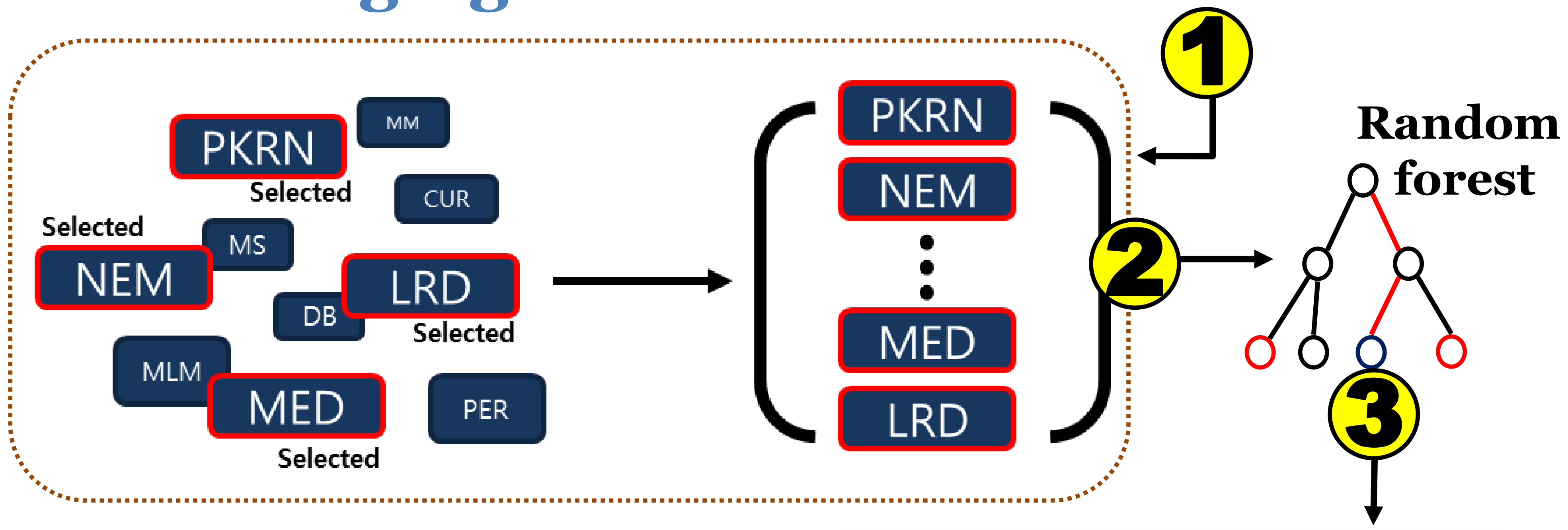
Stereo Matching & Stixel

- Real-time depth map estimation
  - 20ms (avg.) with GTX 1080
  - 200~300ms with Jetson TX1
- STIXEL (Stick + Cell)
  - Used to detect and to represent obstacles

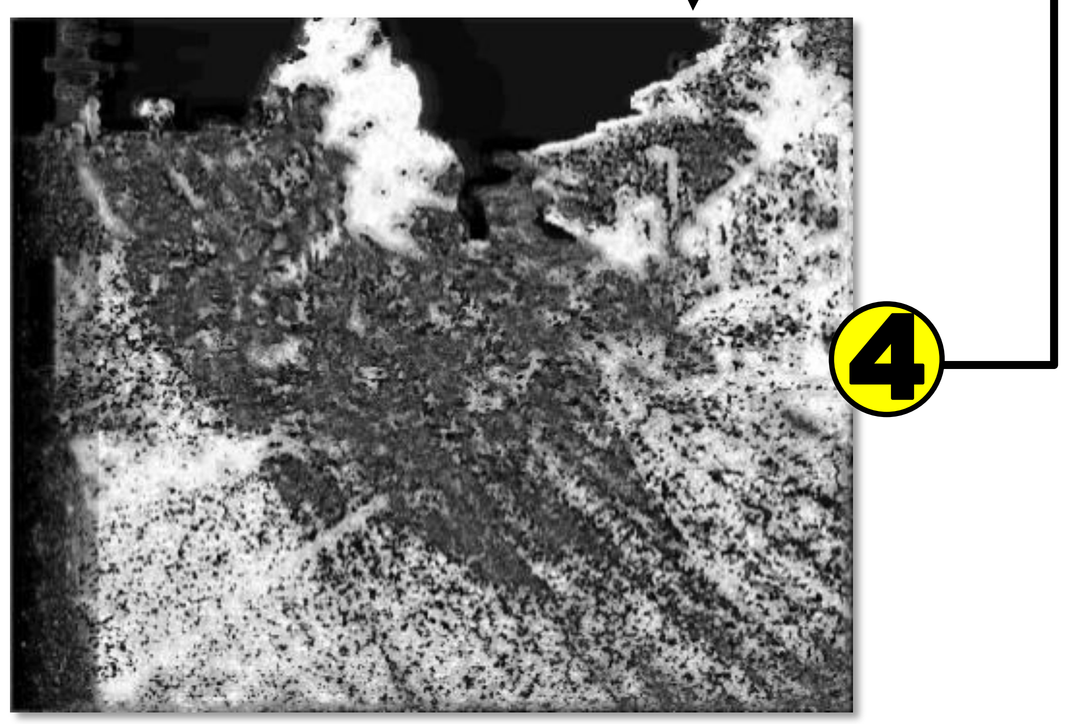


Robust Stereo Matching

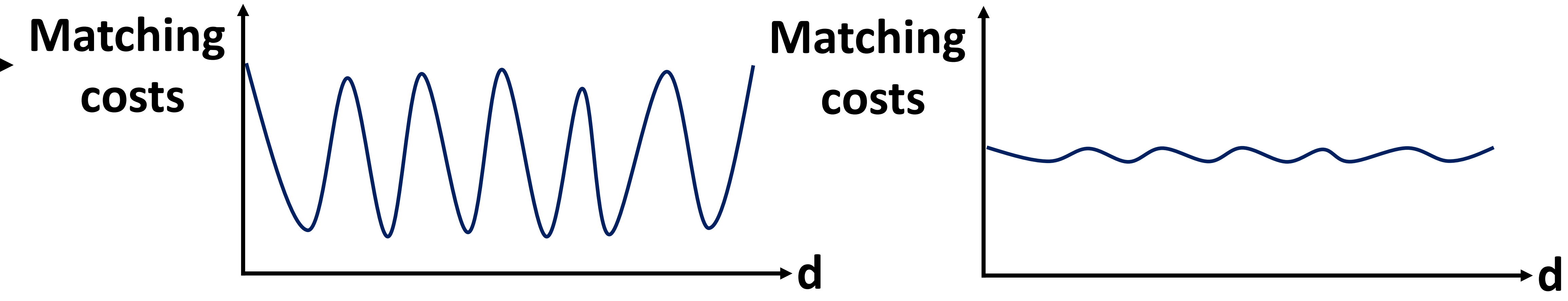
Challenging outdoor environment



Learning-based confidence measure selection and training

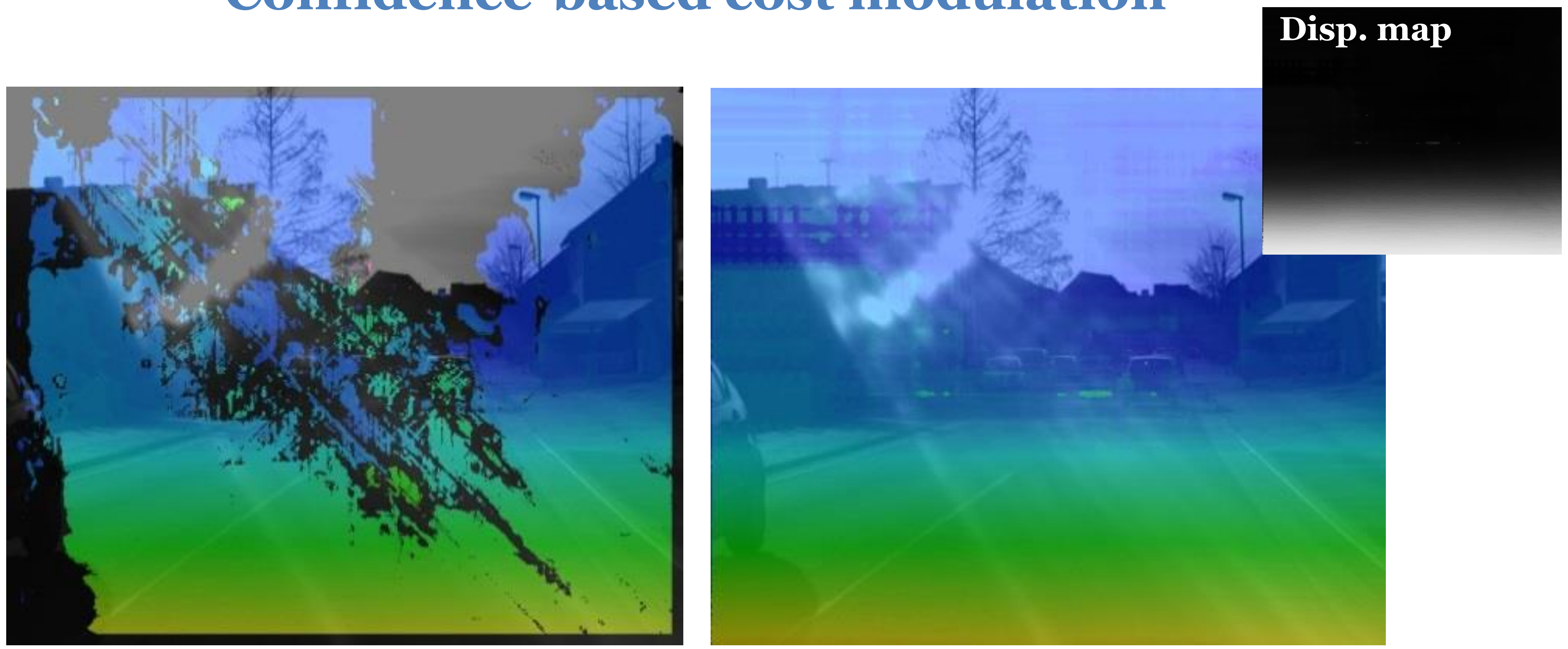


Confidence prediction



$$\hat{C}(\mathbf{p},d) = Q(\mathbf{p})\,C(\mathbf{p},d) + (1-Q(\mathbf{p}))\sum_{k\in D}\frac{C(\mathbf{p},k)}{|D|}$$

Confidence-based cost modulation



Semi-global matching

Proposed method

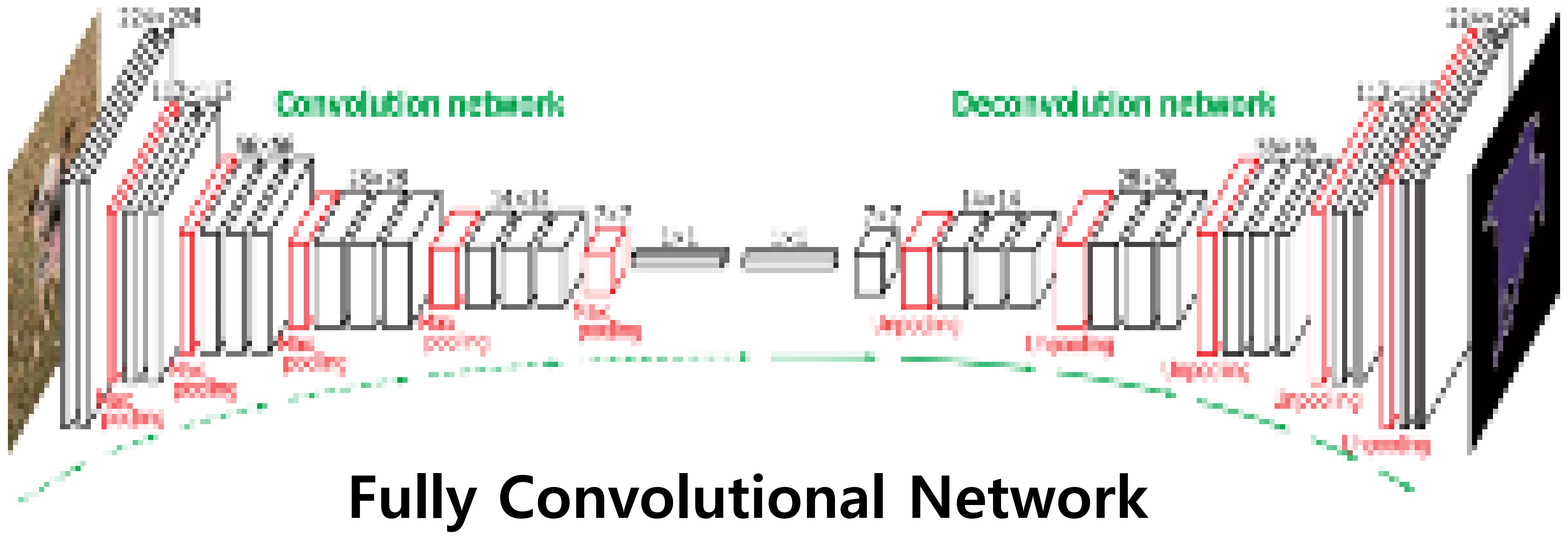
Min-Gyu Park and Kuk-Jin Yoon, "Leveraging Stereo Matching with Learning-based Confidence Measures," *IEEE Conf. on Computer Vision and Pattern Recognition (**CVPR**)*, 2015

Lane Detection

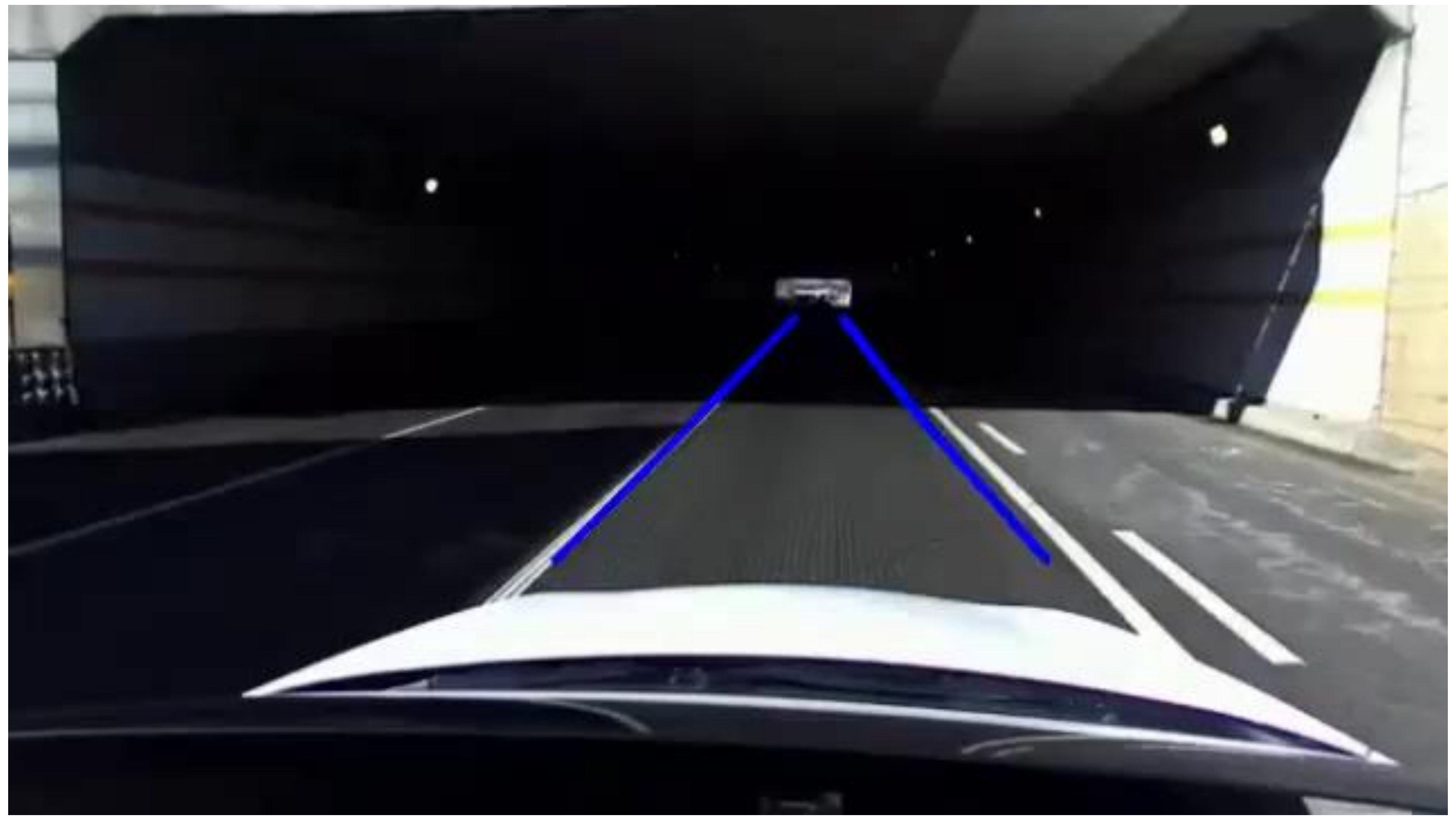
- Classical lane detection approach (300fps with 3.00GHz CPU)

Input image　IPM image　Filter & Threshold　Detect lanes　Result

- Deep learning-based lane detection
  - Dataset : 1,500 images
  - Train dataset : 60,000 images (augmentation)



Fully Convolutional Network



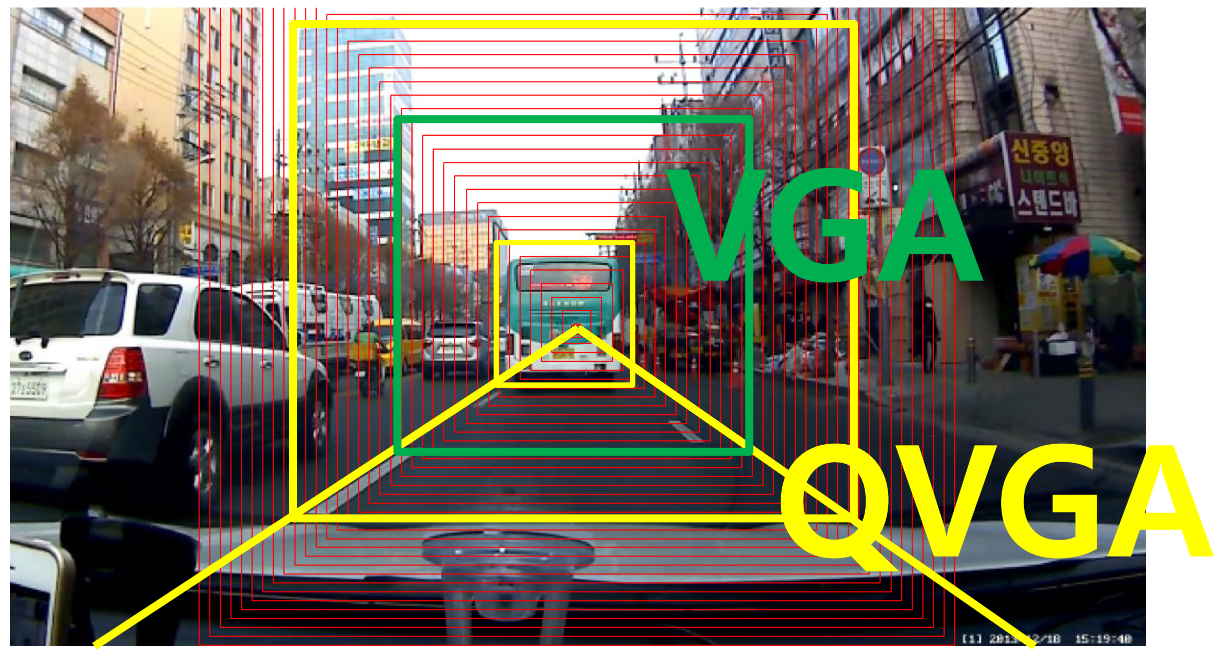
10ms (100fps) with Jetson TX1

Vehicle Detection

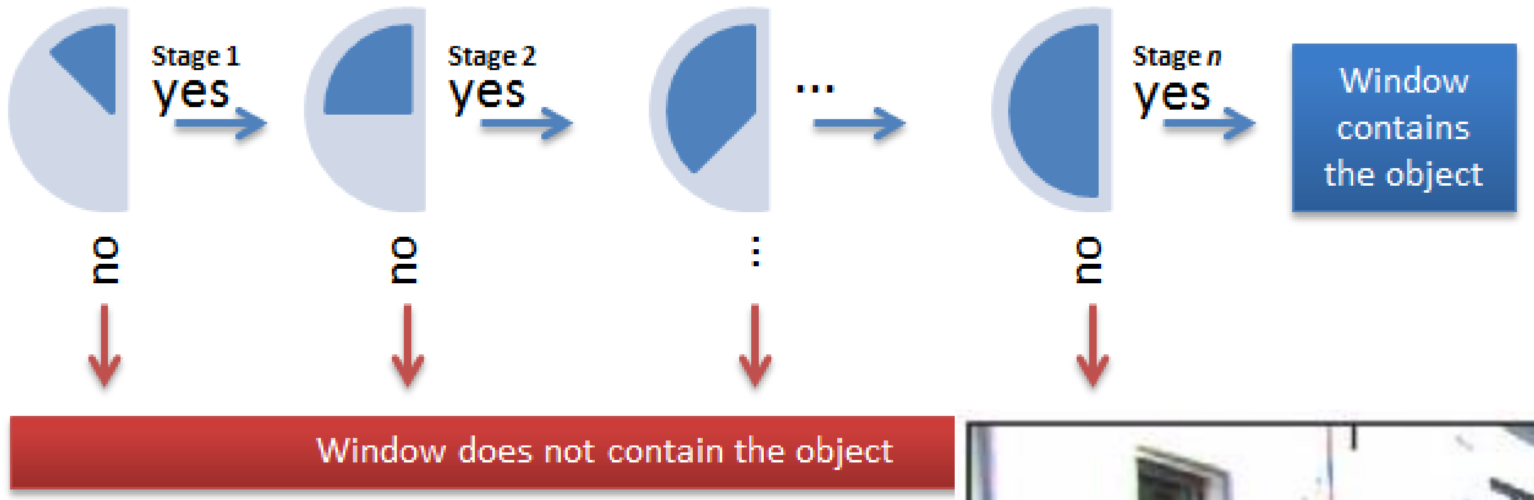
- Real-time vehicle detection
  - Under development
    - HoG + Kernel SVM
    - ACF + Boosting (aggregated channel features)
  - 15ms with Jetson TX1

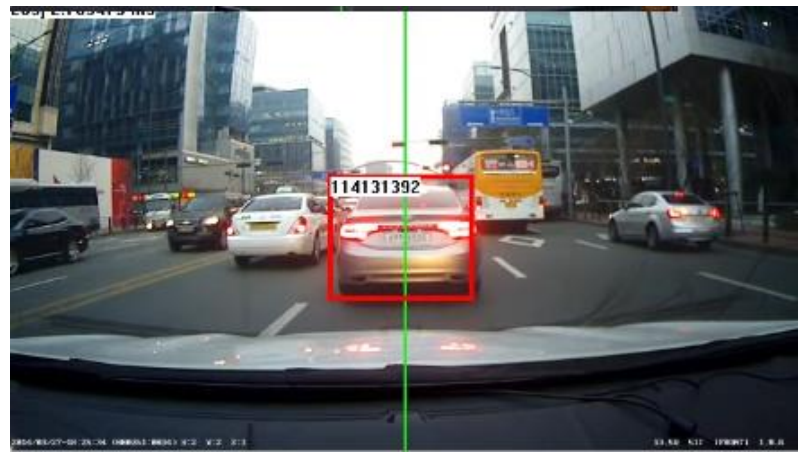
- Expandable to PD (pedestrian detection)

Feature extraction



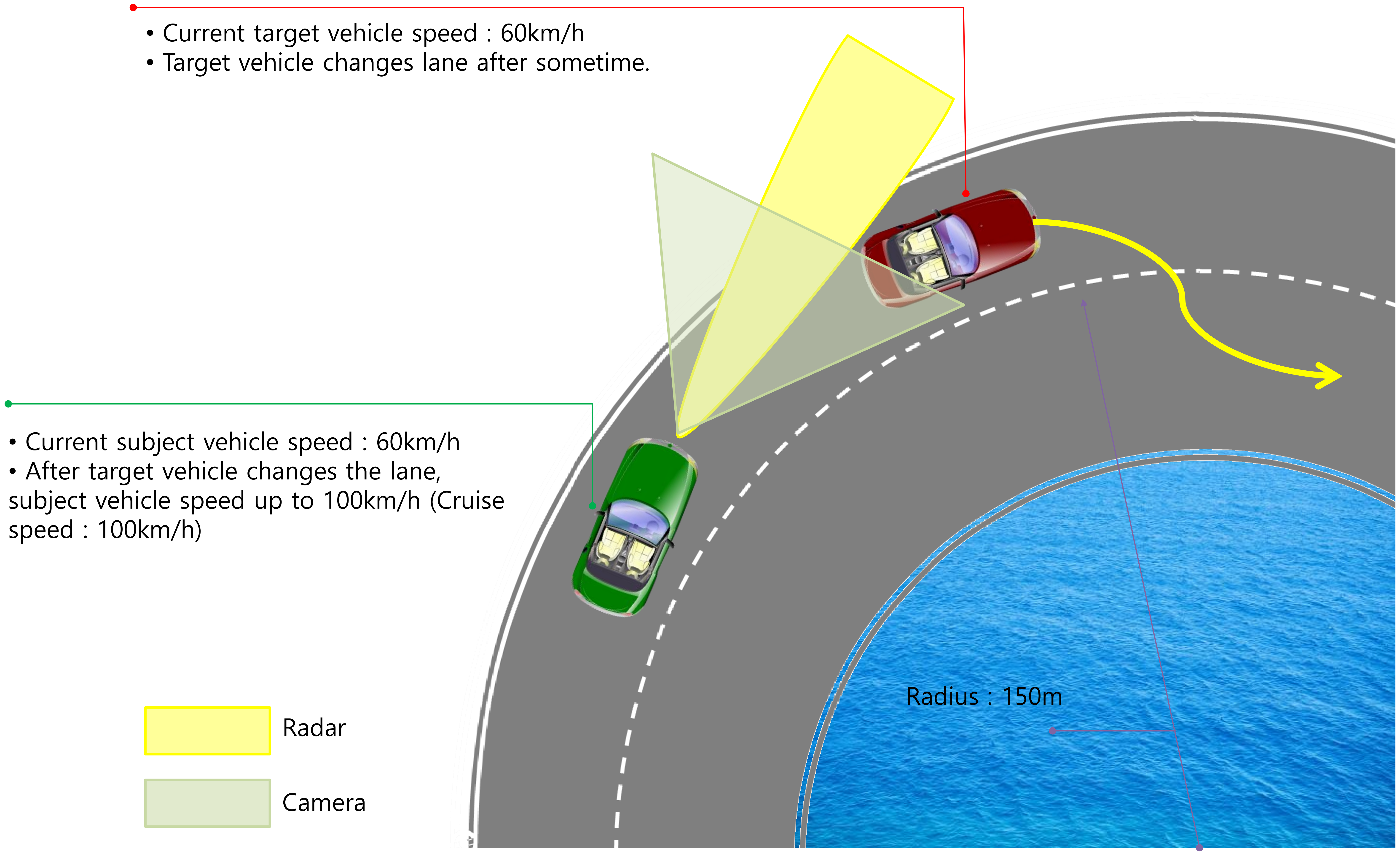
ROI definition





ADAS current project : Vision based Cruise Control

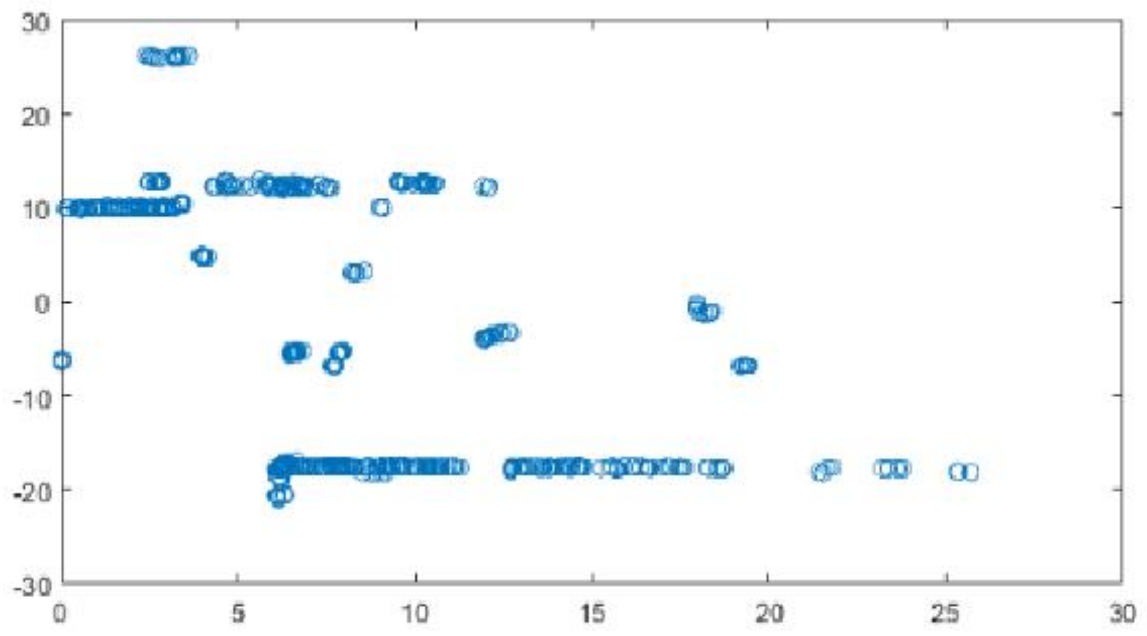
When the blue vehicle changes lane, the red vehicle should drive steadily without abrupt acceleration

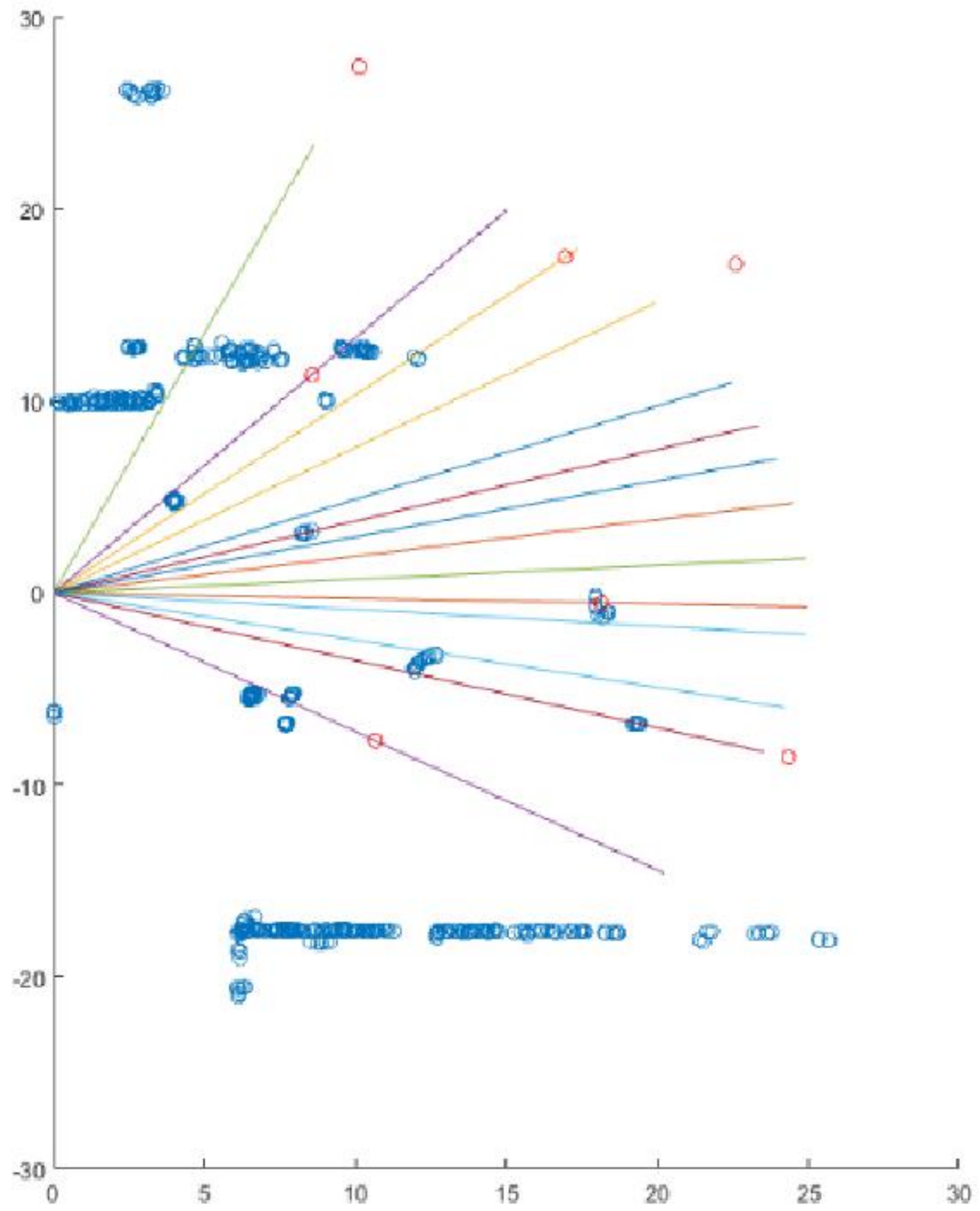


Adaptive Cruise Control

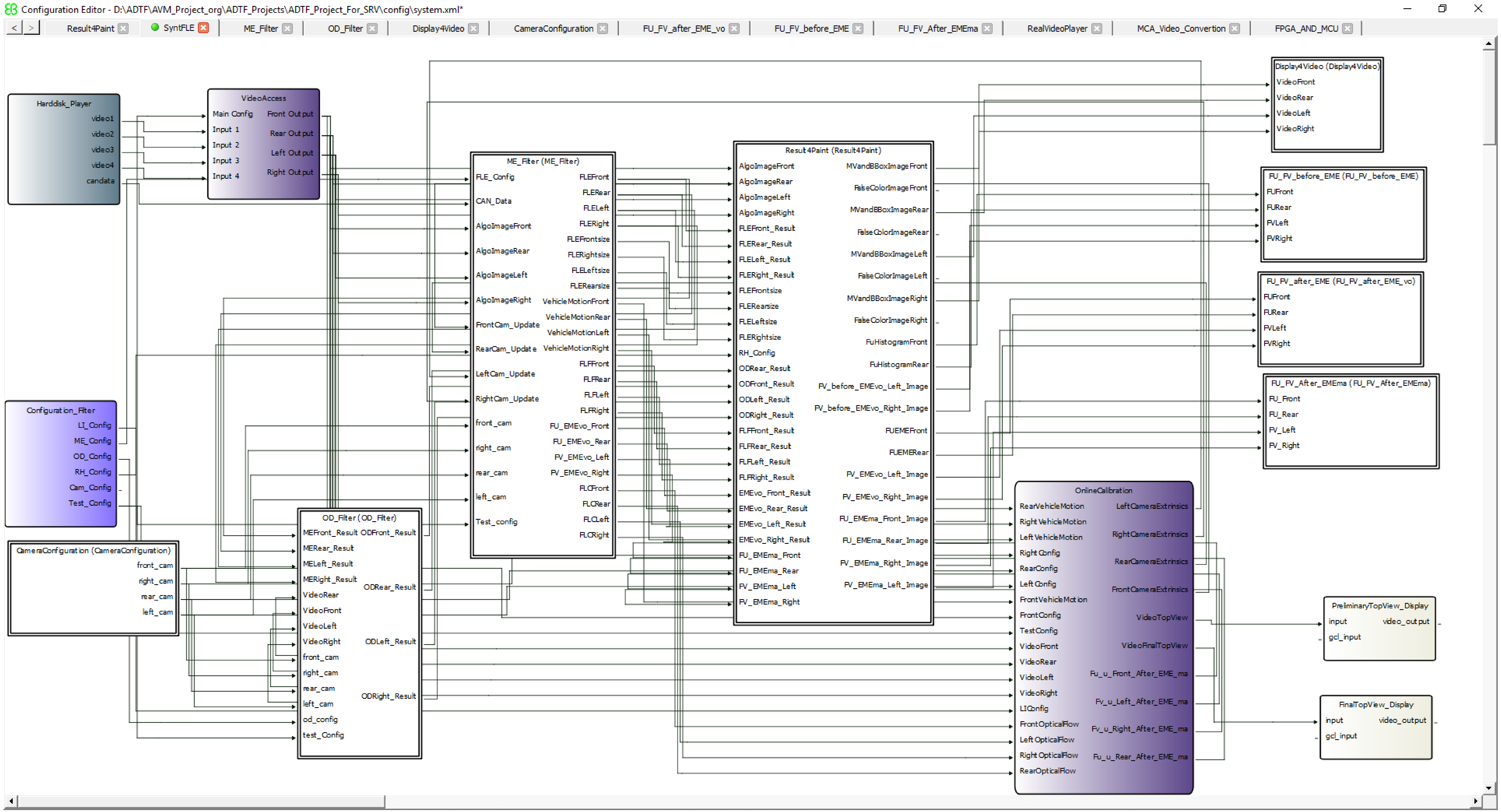
Implementation details

- Fusion: Radar + Camera
- Lane monitoring by Camera
  - Ego-lane detection
  - Medium curvature
- Proprietary Radar algorithm
- Calibration Radar + Camera





ADTF Architecture for Surround View Monitoring System

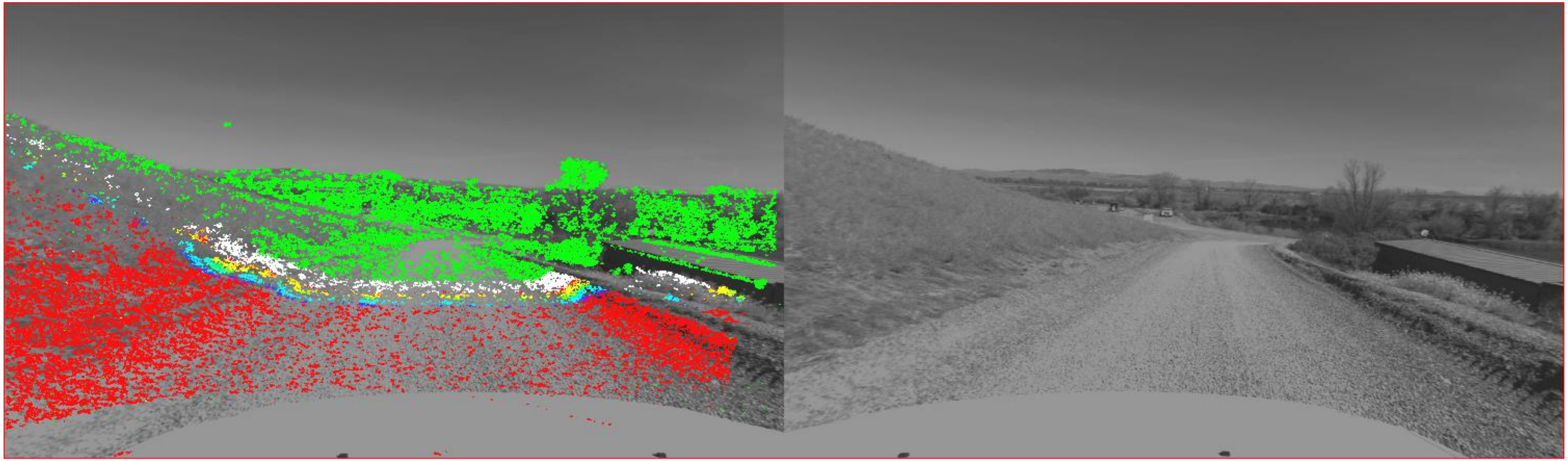


Key competences

- Photogrammetry algorithms
- 3D Reconstruction algorithms
- ADAS algorithms
- System engineering
- Mature development processes (Scrum, Kanban, Waterfall)
- CEVA MM-3101 & XM4 optimization
- CUDA & OpenCL intensive computation
- FPGA programming

Stereo vision algorithm

- High performance – 20 fps (Zynq 7020, HD resolution)
- Low latency – 50 ms (Zynq 7020 , HD resolution)
- Sparse disparity map – 60000-110000 features (HD resolution)



Forward Collision Avoidance

Functions

- Stereo – SGM, AD-Census
- Stixels & CNN-based obstacle detection
- Collision estimation (tracking & trajectory & time to collision)
- Free space to drive

Forward Collision Avoidance

CNN Object detection

- Classes – car, pedestrian, cyclist
- CNN – YOLOv2
- Training set – 90110 images
- mAP – 0.7-0.8
- Performance
  - Platform – nVidia Drive PX2
  - 20 fps
  - Resolution – 512x288 px

Traffic sign detection

Implementation details

- SSD-M with MobileNet features
- High performance – 20 fps on Jetson
- High accuracy – 0.74 mAP
- European, Russian, South Korean traf
    - ~30000 images
- Classifying speed limits
    - ~17000 images
- Low model size – 22 Mb
- Detection performance – TX2: 35 ms, PX2: 24 ms
- Classification performance – TX2: 6 ms, PX2: 4 ms
- Accuracy F1-score – 0.89

그래픽 기반 자율주행 시뮬레이터 (1/3)

자율주행을 위한 그래픽 기반 가상 Dataset의 효용성이 검증됨.
또한 실제로 구현이 어려운 상황의 경우 그래픽 기반 가상 Dataset의 효용성은 더욱 증대됨.

2016 IEEE Conference on Computer Vision and Pattern Recognition

The SYNTHIA Dataset: A Large Collection of Synthetic Images for Semantic Segmentation of Urban Scenes

German Ros[†‡], Laura Sellart[†], Joanna Materzynska[§], David Vazquez[†], Antonio M. Lopez[†‡]

[†]Computer Vision Center
Edifici O,
Campus UAB
Barcelona, Spain

[‡]Computer Science Dept.
Universitat Autonoma de Barcelona
Campus UAB
Barcelona, Spain

[§] Faculty of Mathematics,
University of Vienna
Oskar-Morgenstern-Platz
Vienna, Austria

{gros, laura.sellart, dvazquez, antonio}@cvc.uab.es, a1248555@unet.univie.ac.at

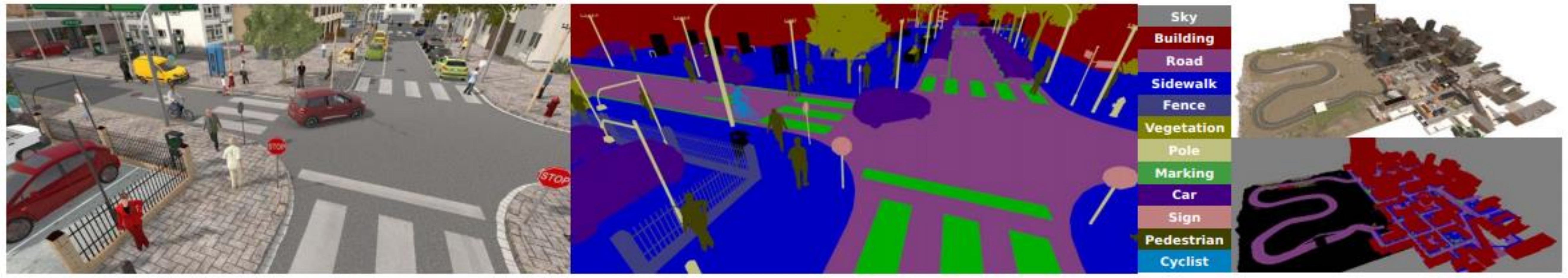


Figure. 1. The SYNTHIA Dataset. A sample frame (Left) with its semantic labels (center) and a general view of the city (right).

2016 IEEE Conference on Computer Vision and Pattern Recognition

Virtual Worlds as Proxy for Multi-Object Tracking Analysis

Adrien Gaidon[1*] Qiao Wang[2*] Yohann Cabon[1] Eleonora Vig[1†]

[1]Computer Vision group, Xerox Research Center Europe, France
[2]School of Electrical, Computer, and Energy Engineering and School of Arts, Media, and Engineering, Arizona State University, USA
{adrien.gaidon,yohann.cabon}@xrce.xerox.com qiao.wang@asu.edu eleonora.vig@dlr.de
http://www.xrce.xerox.com/Research-Development/Computer-Vision/Proxy-Virtual-Worlds

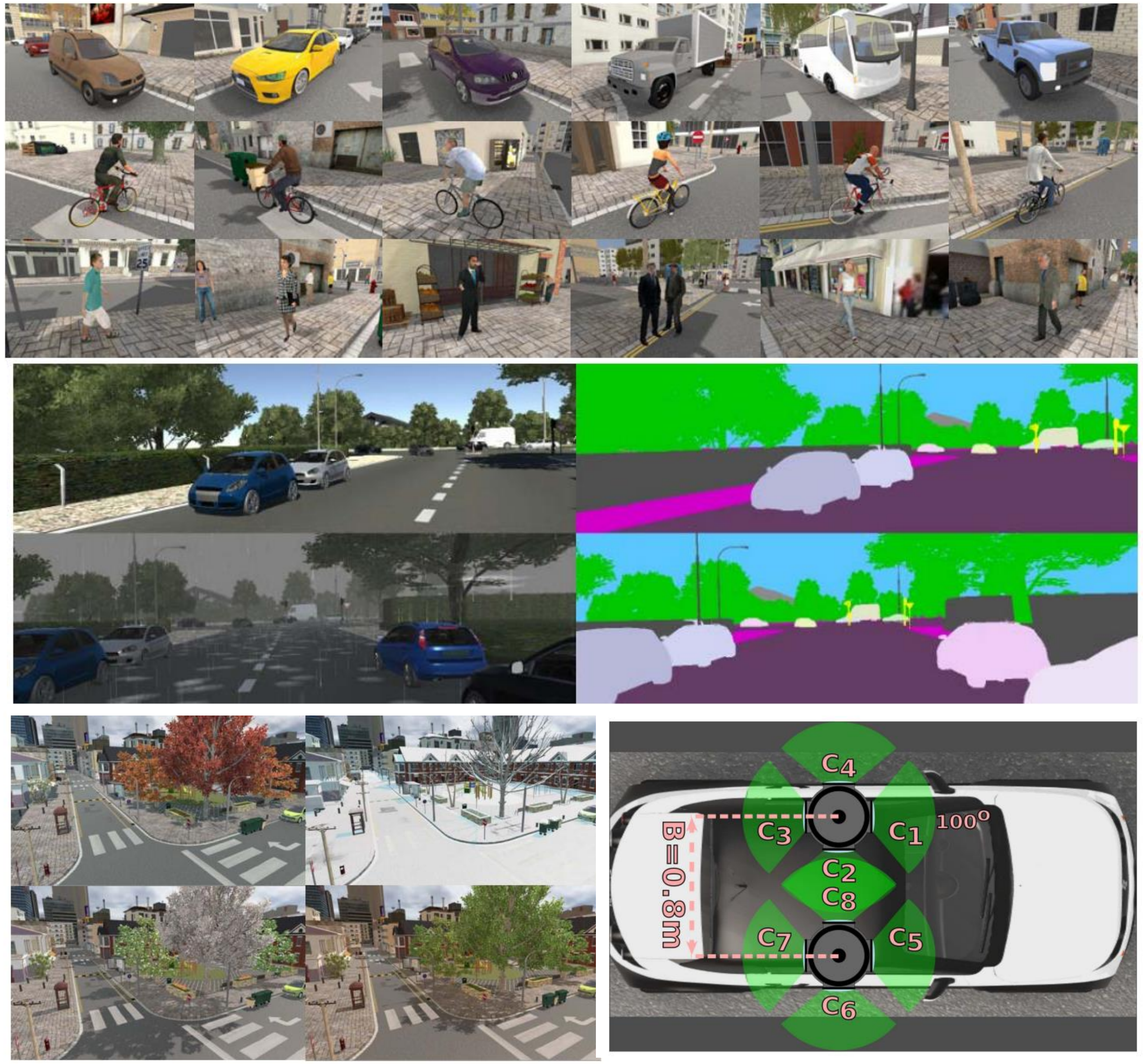
Abstract

Modern computer vision algorithms typically require expensive data acquisition and accurate manual labeling. In this work, we instead leverage the recent progress in computer graphics to generate fully labeled, dynamic, and photo-realistic proxy virtual worlds. We propose an efficient real-to-virtual world cloning method, and validate our approach by building and publicly releasing a new video dataset, called "Virtual KITTI" [1], automatically labeled with accurate ground truth for object detection, tracking, scene and instance segmentation, depth, and optical flow. We provide quantitative experimental evidence suggesting that (i) modern deep learning algorithms pre-trained on real data behave similarly in real and virtual worlds, and (ii) pre-training on virtual data improves performance. As the gap between real and virtual worlds is small, virtual worlds enable measuring the impact of various weather and imaging conditions on recognition performance, all other things being equal. We show these factors may affect drastically otherwise high-performing deep models for tracking.

Figure 1: Top: a frame of a video from the KITTI multi-object tracking benchmark [1]. Middle: the corresponding rendered frame of the synthetic clone from our Virtual KITTI dataset with

그래픽 기반 자율주행 시뮬레이터 (2/3)

그래픽 기반 가상 Dataset을 통한 학습과 실영상 기반 학습을 병행한 경우 더욱 우수한 인식 성능



< 그래픽 객체 예시 및 가상 카메라 배치>

Table 2. Results of training a T-Net and a FCN on *SYNTHIA-Rand* and evaluating it on state-of-the-art datasets of driving scenes.

Method	Training	Validation	sky	building	road	sidewalk	fence	vegetat.	pole	car	sign	pedest.	cyclist	**per-class**	**global**
T-Net [30]	*SYNTHIA-Rand* (A)	CamVid (V)	66	85	86	67	0	27	55	79	3	75	46	48.9	79.7
	SYNTHIA-Rand (A)	KITTI (V)	73	78	92	27	0	10	0	64	0	72	14	39.0	61.9
	SYNTHIA-Rand (A)	U-LabelMe (V)	20	59	92	13	0	22	38	89	1	64	23	38.3	53.4
	SYNTHIA-Rand (A)	CBCL (V)	74	71	87	25	0	35	21	68	2	42	36	41.8	66.0
FCN [20]	*SYNTHIA-Rand* (A)	CamVid (V)	78	66	86	72	12	79	17	91	43	78	68	62.5	74.9
	SYNTHIA-Rand (A)	KITTI (V)	56	65	59	26	17	65	32	52	42	73	40	47.1	62.7
	SYNTHIA-Rand (A)	U-LabelMe (V)	31	63	68	40	23	65	39	85	18	71	46	50.0	59.1
	SYNTHIA-Rand (A)	CBCL (V)	71	59	73	32	26	81	40	78	31	63	72	56.9	68.2

Table 3. Comparison of training a T-Net and FCN on real images only and the effect of extending training sets with *SYNTHIA-Rand*.

Method	Training	Validation	sky	building	road	sidewalk	fence	vegetation	pole	car	sign	pedestrian	cyclist	**per-class**	**global**
T-Net [30]	Camvid (T)	CamVid (V)	99	65	95	52	7	79	5	80	3	26	6	46.3	81.9
	Camvid (T) + *SYNTHIA-Rand* (A)	CamVid (V)	98	**90**	91	**63**	5	**83**	**9**	**94**	0	**58**	**31**	56.5 (**10.2**)	90.7 (**8.8**)
	KITTI (T)	KITTI (V)	79	83	87	73	0	85	0	69	0	10	0	44.2	80.5
	KITTI (T) + *SYNTHIA-Rand* (A)	KITTI (V)	**89**	**86**	**90**	58	0	72	0	**76**	0	**66**	**29**	51.6 (**7.4**)	80.8 (**0.3**)
	U-LabelMe (T)	U-LabelMe (V)	72	80	75	45	0	62	2	53	0	14	2	36.4	62.4
	U-LabelMe (T) + *SYNTHIA-Rand* (A)	U-LabelMe (V)	69	77	**93**	33	0	62	**11**	**77**	**1**	**67**	**24**	46.7 (**10.3**)	72.1 (**9.7**)
	CBCL (T)	CBCL (V)	62	77	86	41	0	74	5	63	0	7	0	37.9	73.9
	CBCL (T) + *SYNTHIA-Rand* (A)	CBCL (V)	**72**	**82**	**90**	39	0	58	**26**	**70**	**5**	**52**	**39**	48.4 (**10.5**)	75.2 (**1.3**)
FCN [20]	Camvid (T)	CamVid (V)	99	65	98	45	27	54	16	77	11	34	25	52.8	78.4
	Camvid (T) + *SYNTHIA-Rand* (A)	CamVid (V)	97	**70**	98	**66**	**39**	**88**	**41**	**88**	**53**	**75**	**79**	72.1 (**18.3**)	83.6 (**5.2**)
	KITTI (T)	KITTI (V)	75	77	77	64	47	84	18	78	5	1	1	51.5	82.3
	KITTI (T) + *SYNTHIA-Rand* (A)	KITTI (V)	**84**	**81**	**82**	**71**	**60**	**86**	**43**	**83**	**24**	**7**	**32**	59.4 (**7.9**)	80.8 (**-1.5**)
	U-LabelMe (T)	U-LabelMe (V)	93	81	83	57	2	79	41	72	20	71	63	60.1	79.4
	U-LabelMe (T) + *SYNTHIA-Rand* (A)	U-LabelMe (V)	93	72	81	**63**	**10**	76	**46**	79	**49**	**76**	**64**	64.4 (**4.3**)	76.2 (**-3.2**)
	CBCL (T)	CBCL (V)	90	77	90	41	2	80	37	84	10	47	31	53.4	79.7
	CBCL (T) + *SYNTHIA-Rand* (A)	CBCL (V)	82	**78**	74	**56**	1	80	20	78	8	**77**	**35**	53.5 (**0.2**)	75.2 (**-4.5**)

< 학습효과의 성능 비교>

AD current project : Last Mile Mobility

We are now collaborating with Korean mini-bus makers and Korean local governments.

- 전기차 기반 자율주행 버스 아르메(ARMA)
- Lidar 센서 : 환경을 매핑하여 위치 지정, 3D 인식
- GPS RTK : GPS 센서와 기지국 간 통신(차량의 정확한 위치 결정)
- Odometer : 변위 및 바퀴 속도 측정 → 차량 속도 추정, 위치 확인
- 카메라 스테레오 비전 : 도로 환경 분석 및 정보 추출

Arma

Device

Route

http://navya.tech/

• 자율주행 버스 EZ10
• 다중 센서 localization 기술, 장애물 감지 및 회피, 탐색, 경로 계획 및 제어, 연결성(V2V, V2I), 차량 관리, 안전 및 사이버 보안
• 3가지 작동 모드
(Metro 모드 : 모든 역에서 정차, Buses 모드 : 요청 시에 정차, On-demand 모드 : 호출 가능)

3 modes

Hybrid localization

- Bus Rapid Transit(BRT) : 전용 차선에서 운영. 속도, 규칙성, 빈도, 다수의 수송인원
- Buses : 경로 명확. 상업적, 경제적, 생태적 효율성
- Tram-train : 도시 지역의 경전철 트랙과 교외 지역의 철도 트랙에서 작동
- Metro : Roissy-Charles de Gaulle의 공항 터미널, 기차역 및 주차장을 연결하는 전기 메트로 CDGVal 운영

BRT

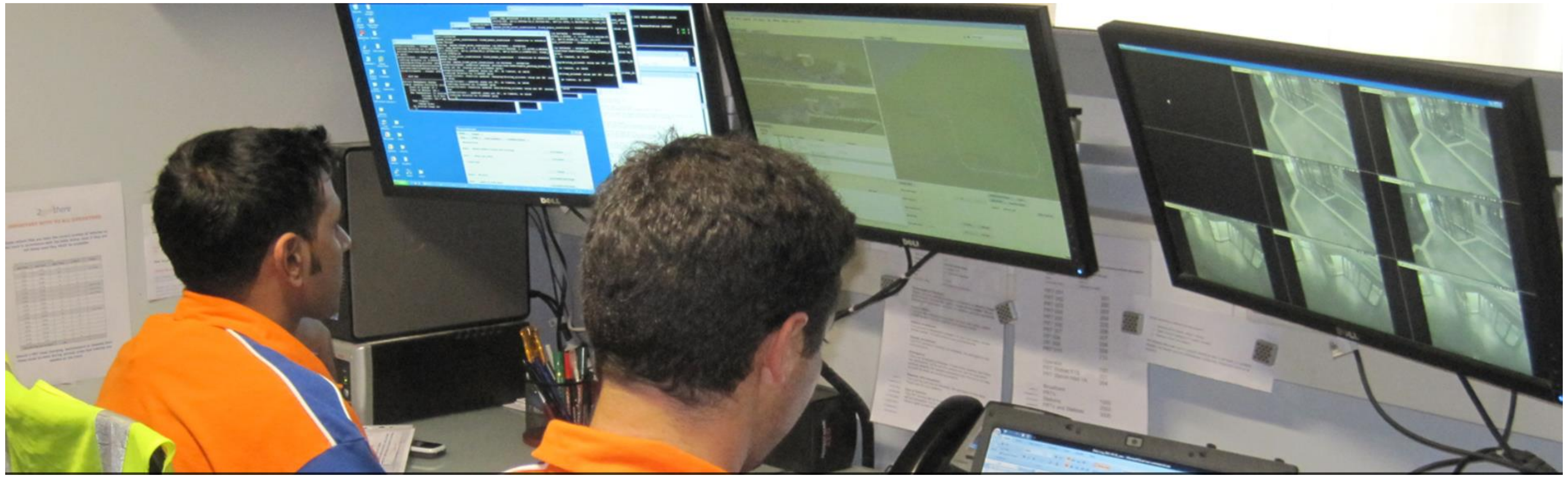
Bus

Tram-train

Metro

2getthere

- 전용 가이드웨이에서 운행하는 자율 주행 셔틀(안전성, 유연성 보장)
- 4~6人 승객 PRT 차량(택시), 16~24人 승객 GRT 차량(버스)
- 대중교통 운영 모니터링 및 감독(TOMS), 차량 제어 시스템(VCS), 안내 제어 시스템(GCS)



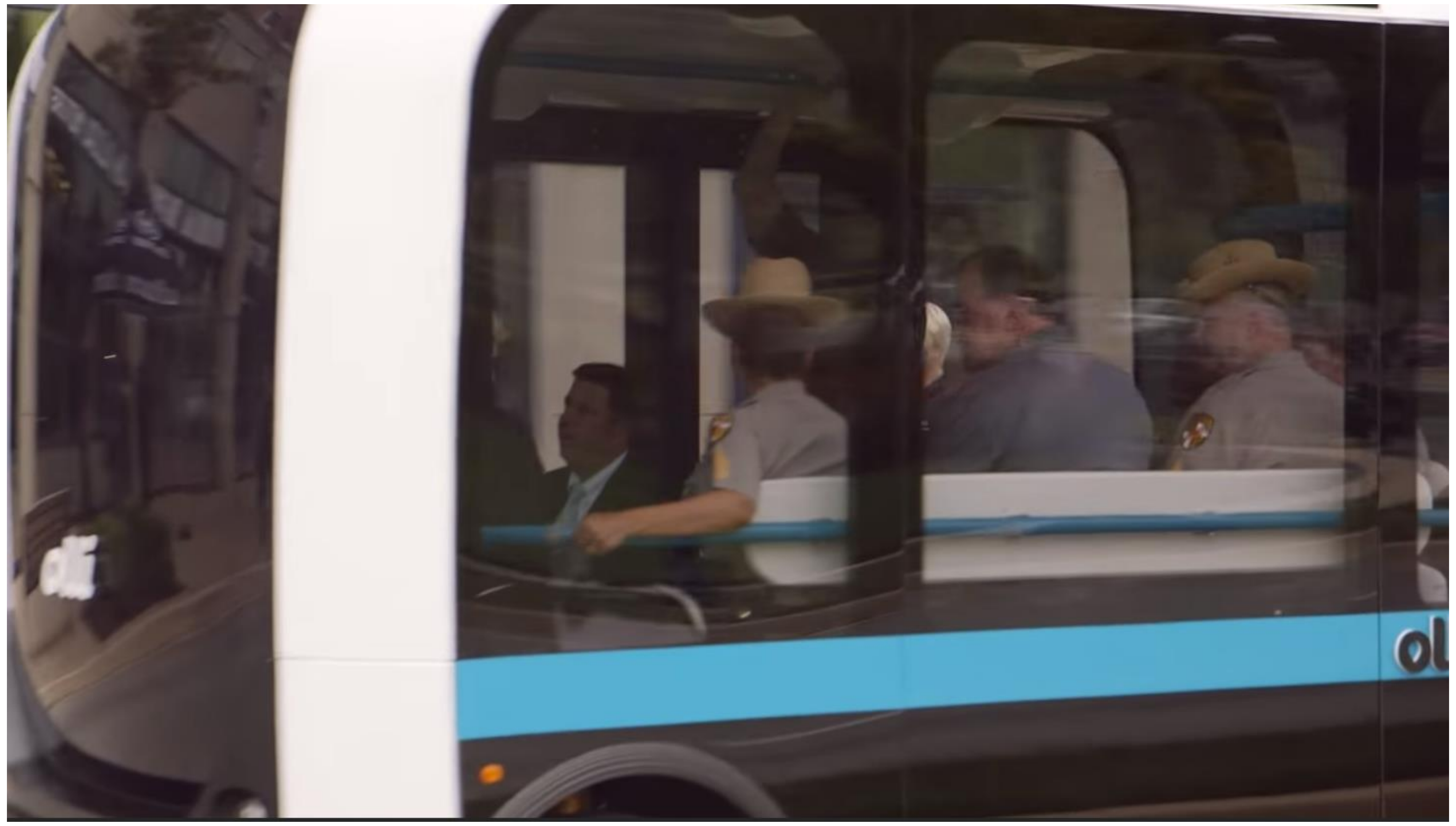
TOMS

GRT

https://www.2getthere.eu/technology/vehicle-types/

- 다수의 마이크로 공장을 이용한 오픈 소스 차량 디자인의 소량 생산
- 자율주행 버스 Olli 개발(IBM Watson의 고급인지능력 통합)
- 360도 센서
- 원격 차량 모니터링, 관리 및 경고 시스템

Autonomous vehicle



Inside the vehicle

AURO

- 자율주행 버스 개발
  (대학 캠퍼스, 비즈니스 파크 및 주거 지역 내 여행용 - 2인승, 6인승, 12인승 셔틀 버스)
- 기후 제어, Lidar 및 Camera로 360도 감지, 장애물 탐색 및 회피 등

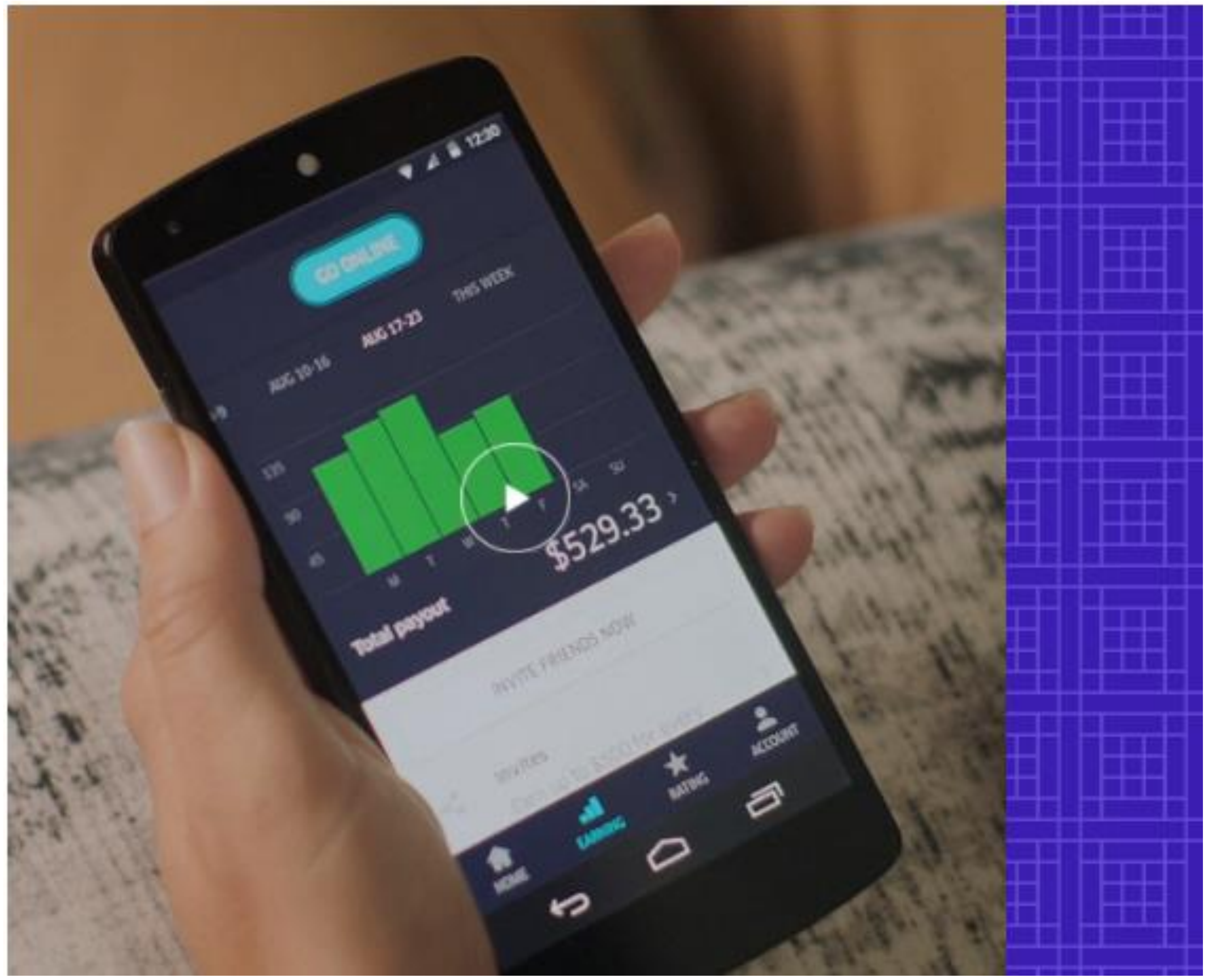
Autonomous vehicle

Running the car

UBER

- 물류 서비스
- 모바일 어플리케이션을 통해 승용차와 운전자를 연결
- 원하는 차량 서비스 선택 가능(이코노미, 프리미엄, 카풀 등)

Utilization the Uber taxi

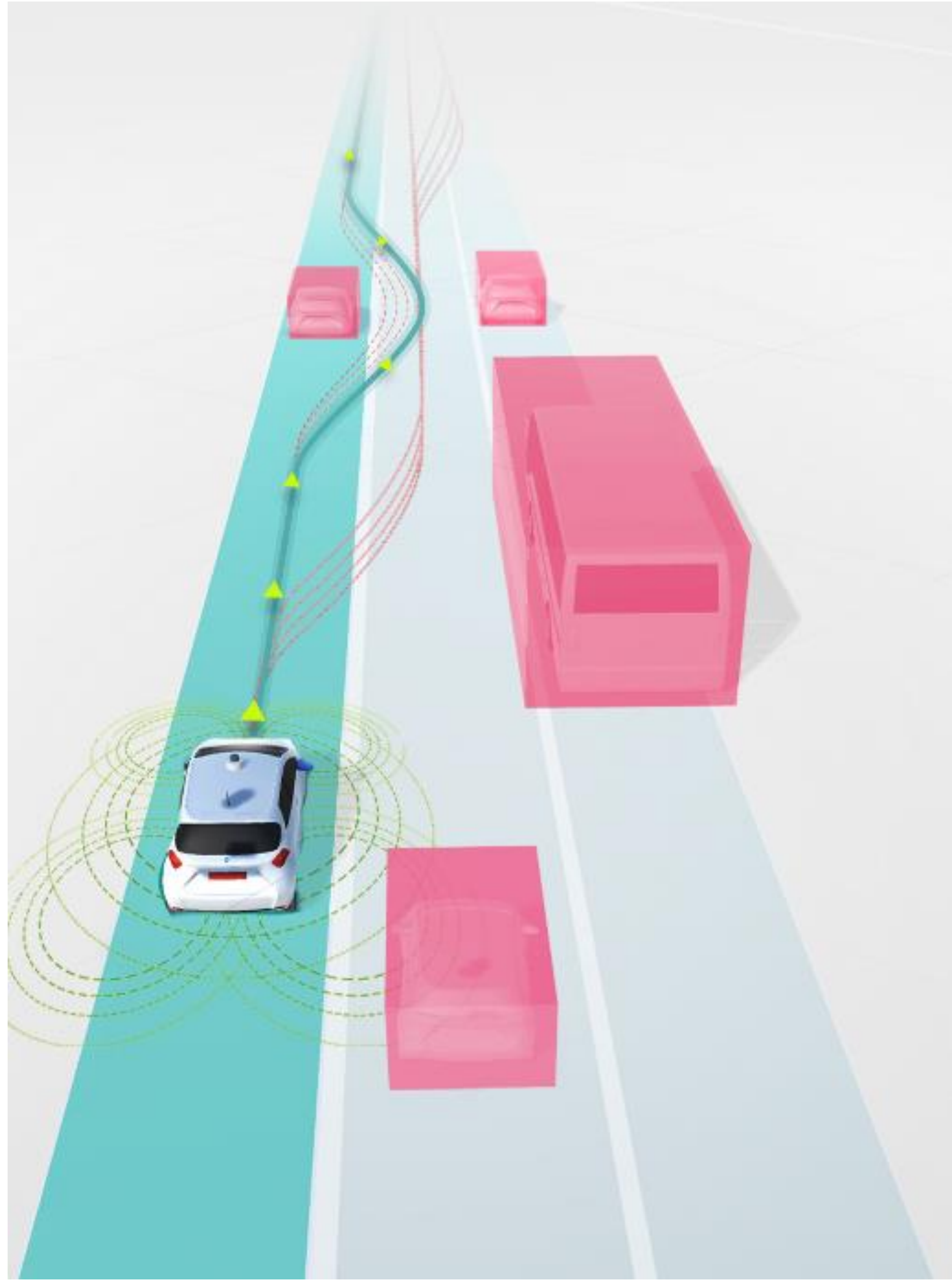


Apps for driver partners

https://www.uber.com/ko-KR/

• 자율 주행을 위한 확장 가능 full-stack 소프트웨어 솔루션
• 대규모 자치 차량을 통해 point-to-point 이동성을 제공하는 솔루션 개발 中.
(도시에서의 자율 주행, 스마트폰 기반 승차, 차량 경로 및 관리, 원격 조작을 통해 차량을 제어하는 소프트웨어)
• 미국, 싱가포르 및 유럽에서 테스트 完.

Autonomous vehicle



NuCore™ software

Thank you very much!

Paul Kang
paulkang@prosensetek.com

PROSENSE

Drive for Excellence!