Real-Time Object Detection and Semantic Segmentation

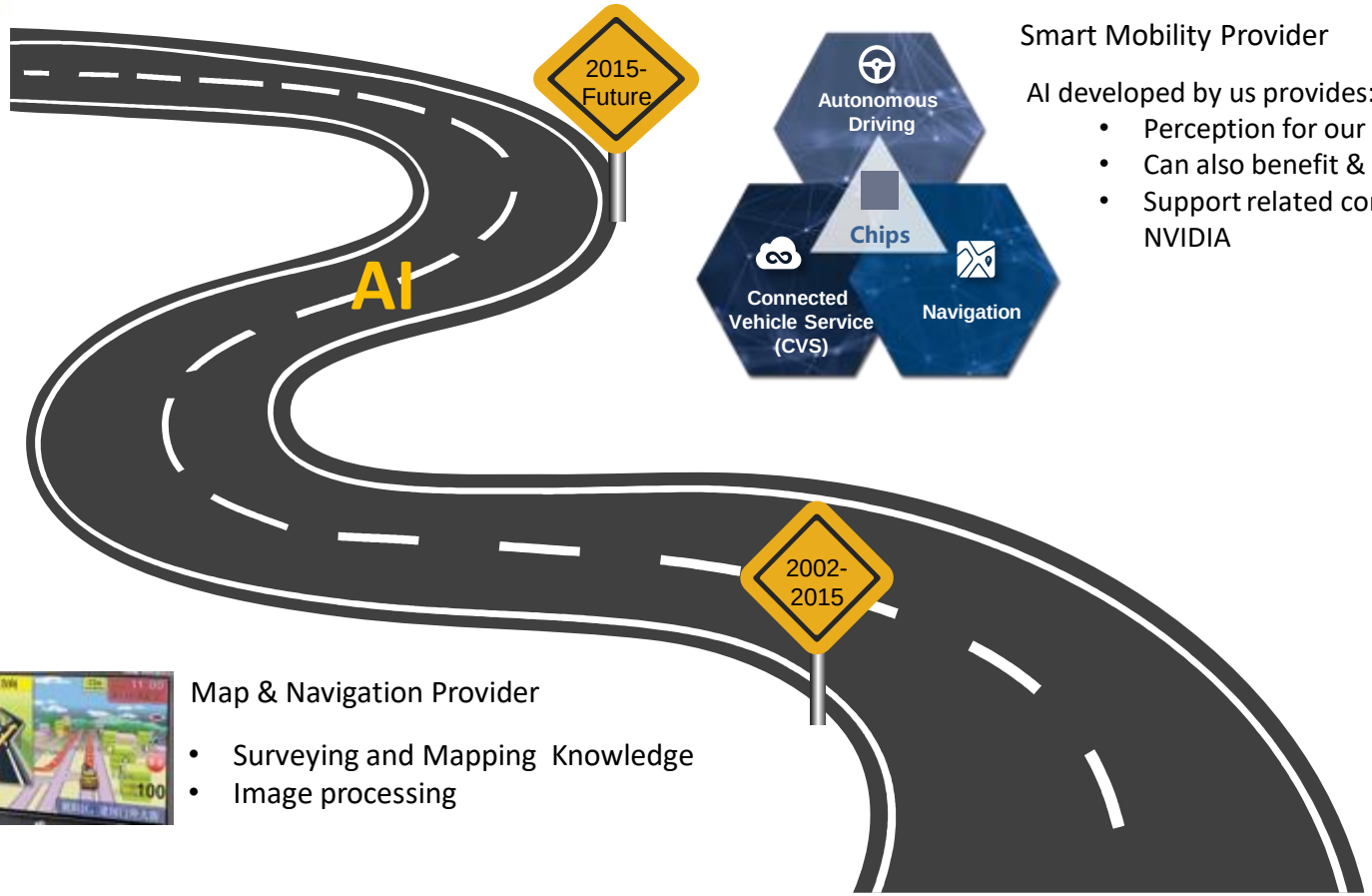
Wei Li
Corporate Development manager

Geetank Raipuria
Computer Vision Engineer

NavInfo Europe

Presented at GTC 2019
Session S9351

NavInfo - our growth benefits from AI



Smart Mobility Provider

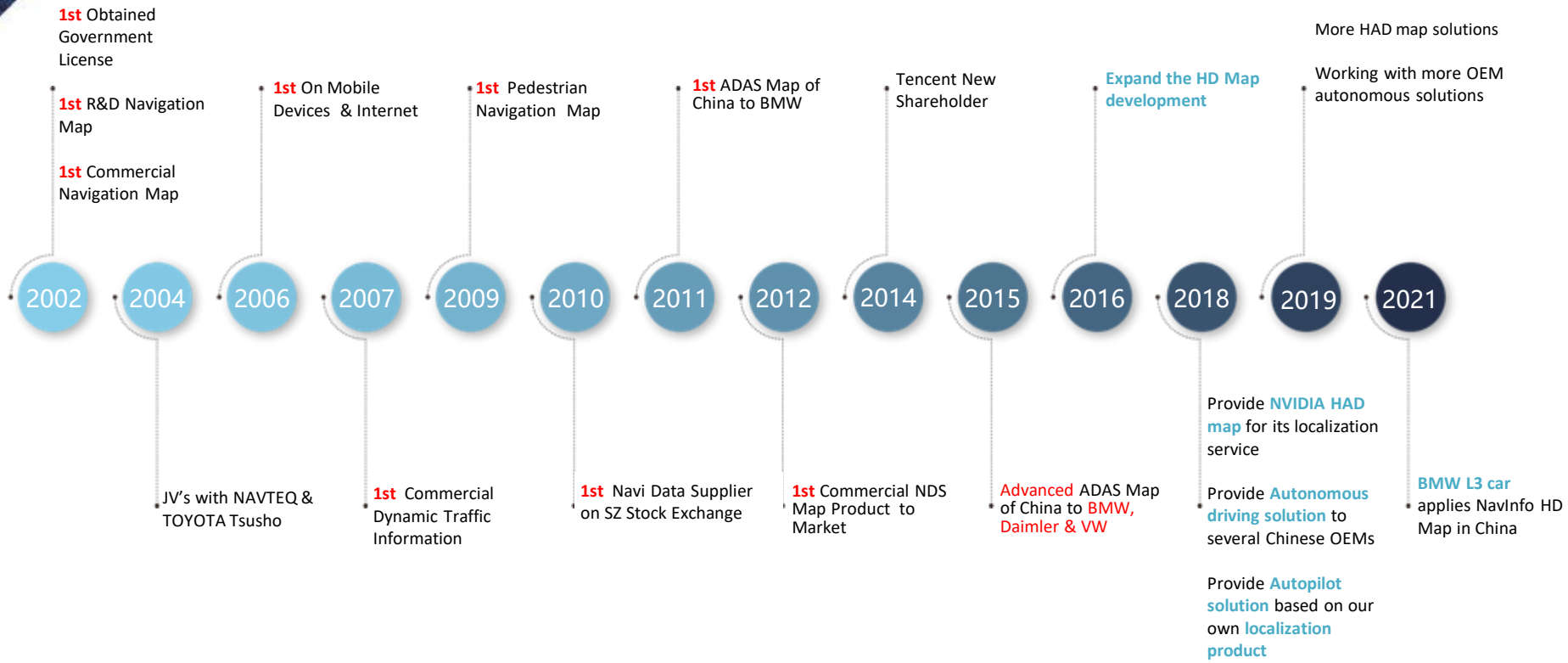
AI developed by us provides:

- Perception for our autonomous solutions
- Can also benefit & apply to other industries
- Support related companies/partners, like NVIDIA

Map & Navigation Provider

- Surveying and Mapping Knowledge
- Image processing

NavInfo's business growth path



NavInfo's Footprint

America

- *International Business Expansion*

Netherlands

1) NavInfo EU

- *Advanced AI Research Lab*
- *EU Business Expansion*

2) AIIM

- *Autonomous driving & Robotics solution provider*

3) Mapscape

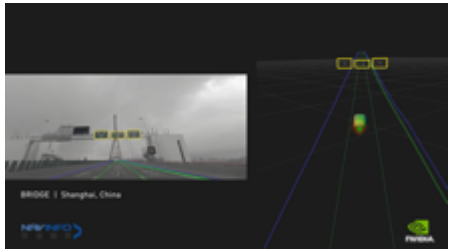
China

- *21 Localization Bases for Data Collection and Technology Services*
- *6 R&D Centers (Shanghai, Xi'an, Shenyang, Wuhan, Hefei, Shenzhen)*
- *Beijing Headquarters*

Singapore

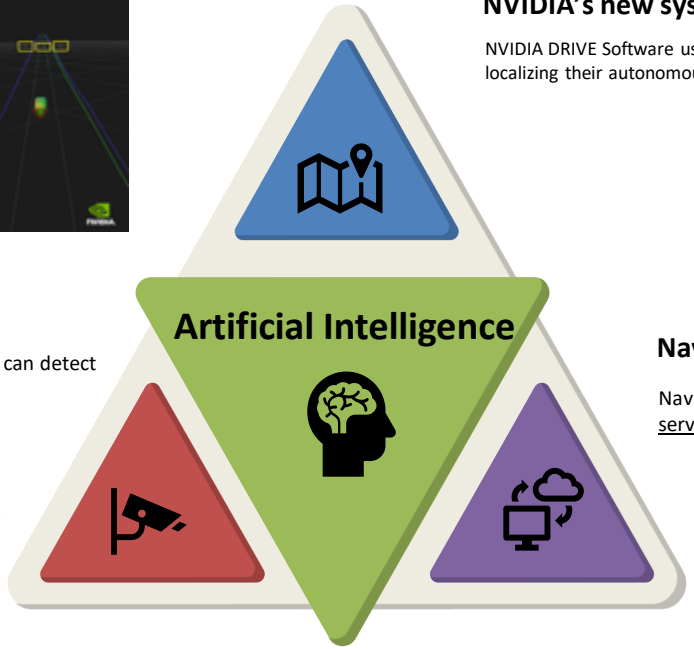
- *Southeast Asia Business Expansion*

Cooperation between NavInfo & NVIDIA



NVIDIA's new system - DRIVE Localization

NVIDIA DRIVE Software use NavInfo's HD Map to give customer a way of localizing their autonomous driving car.

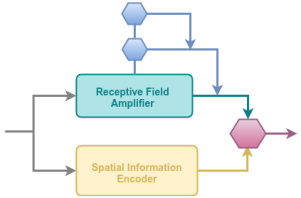


NavInfo's Perception Technology

NavInfo's research lab developed a vision based system that can detect and classify objects in real time on NVIDIA Xavier

NavInfo's Training Model

NavInfo trains and optimizes its models on NVIDIA DGX-1 servers



We support and benefit each other's achievements, driven by AI, to generate better products and services to our customers and end users

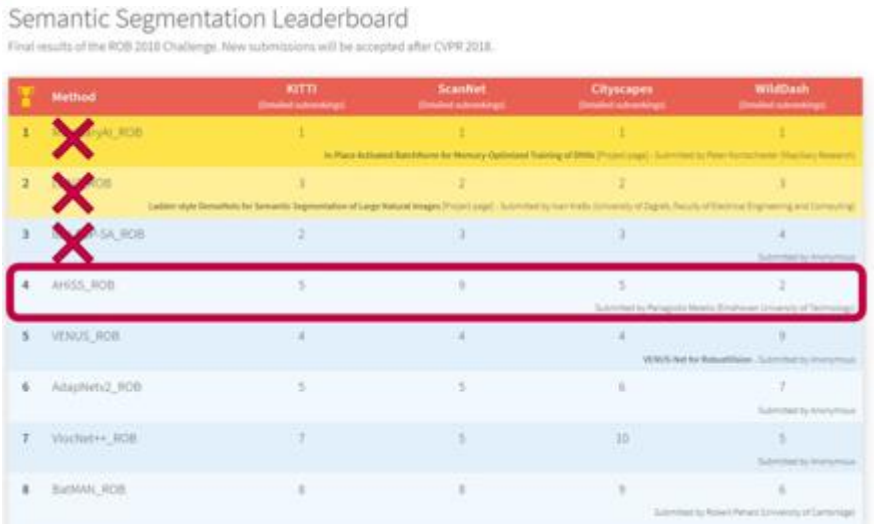
NavInfo Service Offerings in Europe

Autonomous driving and robotic solutions

AI based solutions for different industries

AI based algorithms

Results from TU/e sponsored by NavInfo Europe



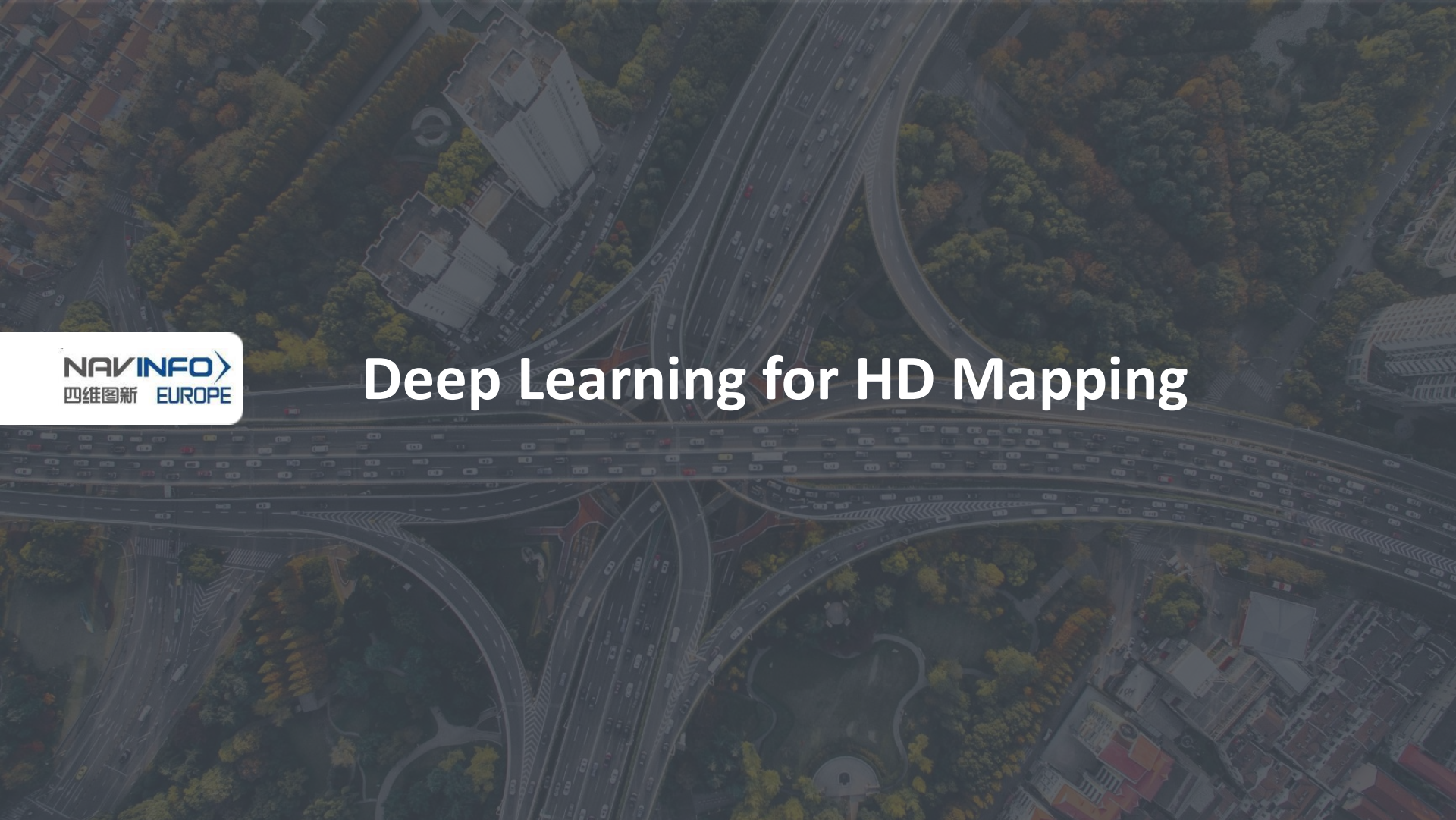
No. 4 in the world in the Semantic Segmentation

Reference: [1] Panagiotis Meletis and Gijs Dubbelman, "Training of Convolutional Networks on Multiple Heterogeneous Datasets for Street Scene Semantic Segmentation", In IEEE IV 2018. June 2018, 0–8.

KITTI Leaderboard

Method	Name	Tra. error	Rot error.	Comp.
V-LOAM	LiDAR SLAM	0.60 %	0.0014 [deg/m]	0.1 s / 2 cores
LOAM	LiDAR SLAM	0.61 %	0.0014 [deg/m]	0.1 s / 2 cores
IMLS-SLAM++	LiDAR SLAM	0.64 %	0.0015 [deg/m]	1.25 s / 1 core
SOFT2	SLAM	0.65 %	0.0014 [deg/m]	0.1 s / 2 cores
IMLS-SLAM	LiDAR SLAM	0.69 %	0.0018 [deg/m]	1.25 s / 1 core
MC2SLAM	LiDAR SLAM	0.69 %	0.0016 [deg/m]	0.1 s / 4 cores
ESO	**Stereo odometry**	**0.80 %**	**0.0026 [deg/m]**	**0.08 s / 4 cores**
sGAN-VO	Stereo odometry	0.81 %	0.0025 [deg/m]	0.1 s / 1 core
LG-SLAM	LiDAR SLAM	0.82 %	0.0020 [deg/m]	0.2 s / 2 cores
RotRocc+	Stereo odometry	0.83 %	0.0026 [deg/m]	0.25 s / 2 cores
GDVO	Stereo odometry	0.86 %	0.0031 [deg/m]	0.09 s / 1 core
CPFG-slam	LiDAR odometry	0.87 %	0.0025 [deg/m]	0.03 s / 4 cores
SOFT	Stereo odometry	0.88 %	0.0022 [deg/m]	0.1 s / 2 cores
RotRocc	Stereo odometry	0.88 %	0.0025 [deg/m]	0.3 s / 2 cores
DVSO	Stereo odometry	0.90 %	0.0021 [deg/m]	0.1 s / GPU
SSO	Stereo odometry	0.93 %	0.0021 [deg/m]	0.1 s / 1 core
LIMO	LiDAR SLAM	0.93 %	0.0026 [deg/m]	0.2 s / 2 cores
Stereo DSO	Stereo odometry	0.93 %	0.0020 [deg/m]	0.1 s / 1 core
Elbrus	Stereo odometry	0.98 %	0.0023 [deg/m]	0.1 s / 1 core
ROCC	Stereo odometry	0.98 %	0.0028 [deg/m]	0.3 s / 2 cores

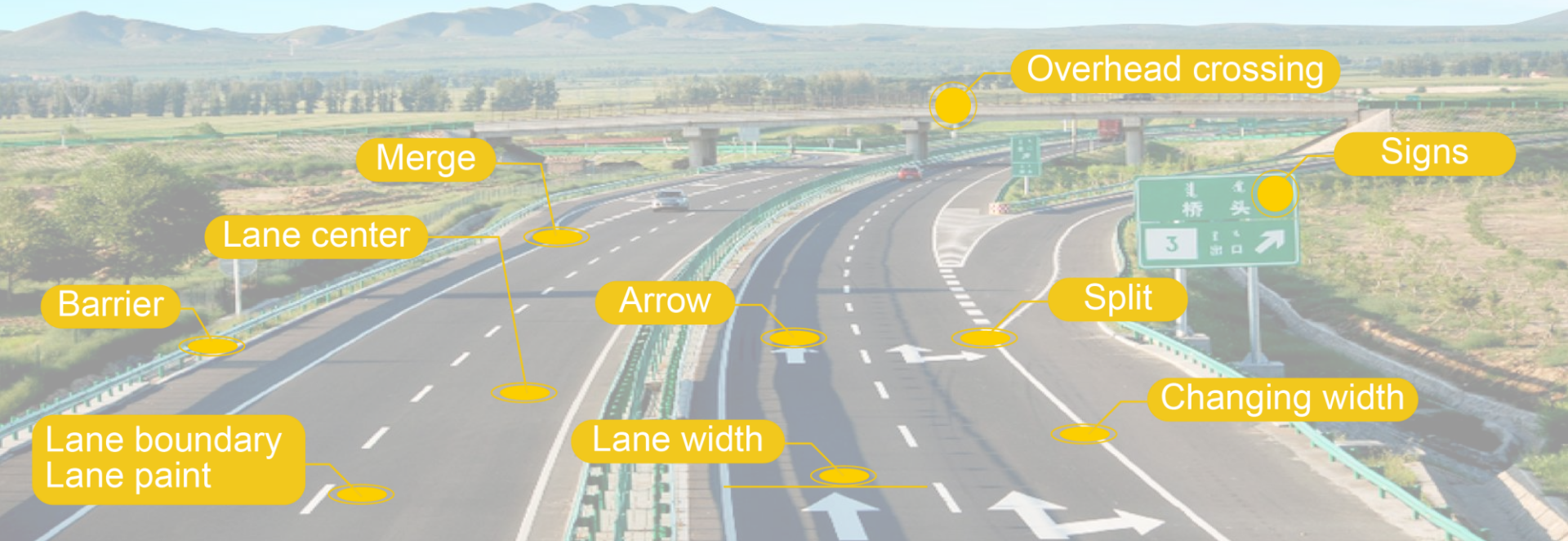
Results with the highest speed and lowest error in Stereo Odometry (not Lidar, not SLAM)



Deep Learning for HD Mapping

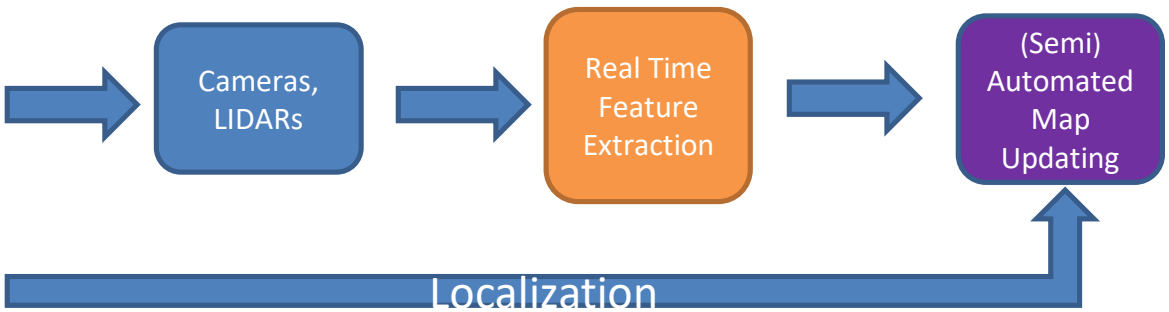
HD Mapping

Includes highly accurate lane and road features.



Deep Learning for HD Mapping Feature Extraction

- Deep learning provides automating feature extraction from video feed of collection vehicles



Real Time Object Detection

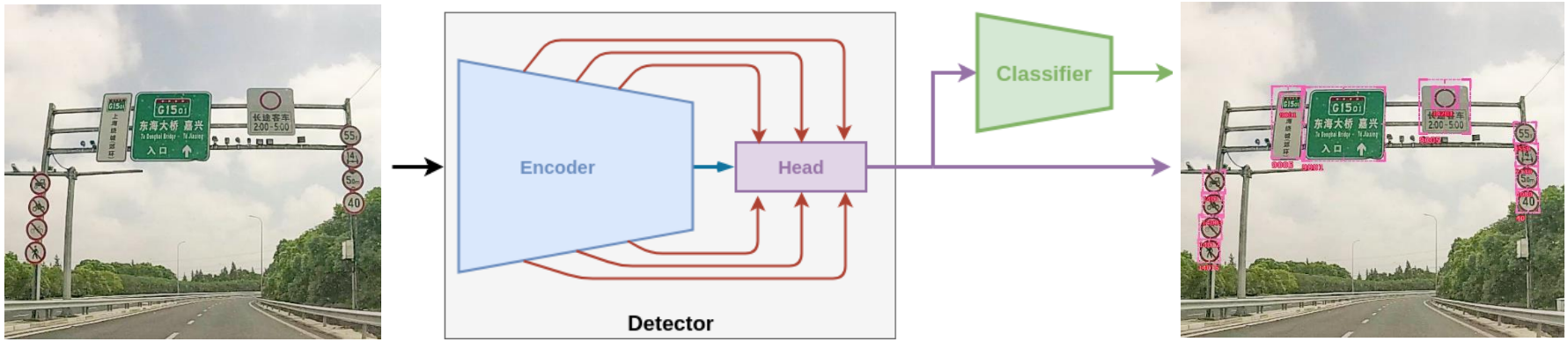
NAVINFO 四维图新 EUROPE

Geetank Raipuria
Computer Vision Engineer
Advanced Research Lab
NavInfo Europe

Andrei Pata
Software Engineer
Advanced Research Lab
NavInfo Europe

Real-Time Object Detection System

Object Detector based on Deep Convolutional Neural Network architecture, to localize and classify road signs and traffic lights from a real-time camera feed



Two Stage System: Best of both worlds

High Accuracy

Low Inference time

Real-Time Object Detection System

Features Supported

350+ supported classes including

Traffic Signs

- Regulatory Signs
- Warning Signs
- Guide Signs
- Information Signs
- Road Work Signs

Signboards

Traffic Lights

Digital Traffic Signs

Real Time Performance

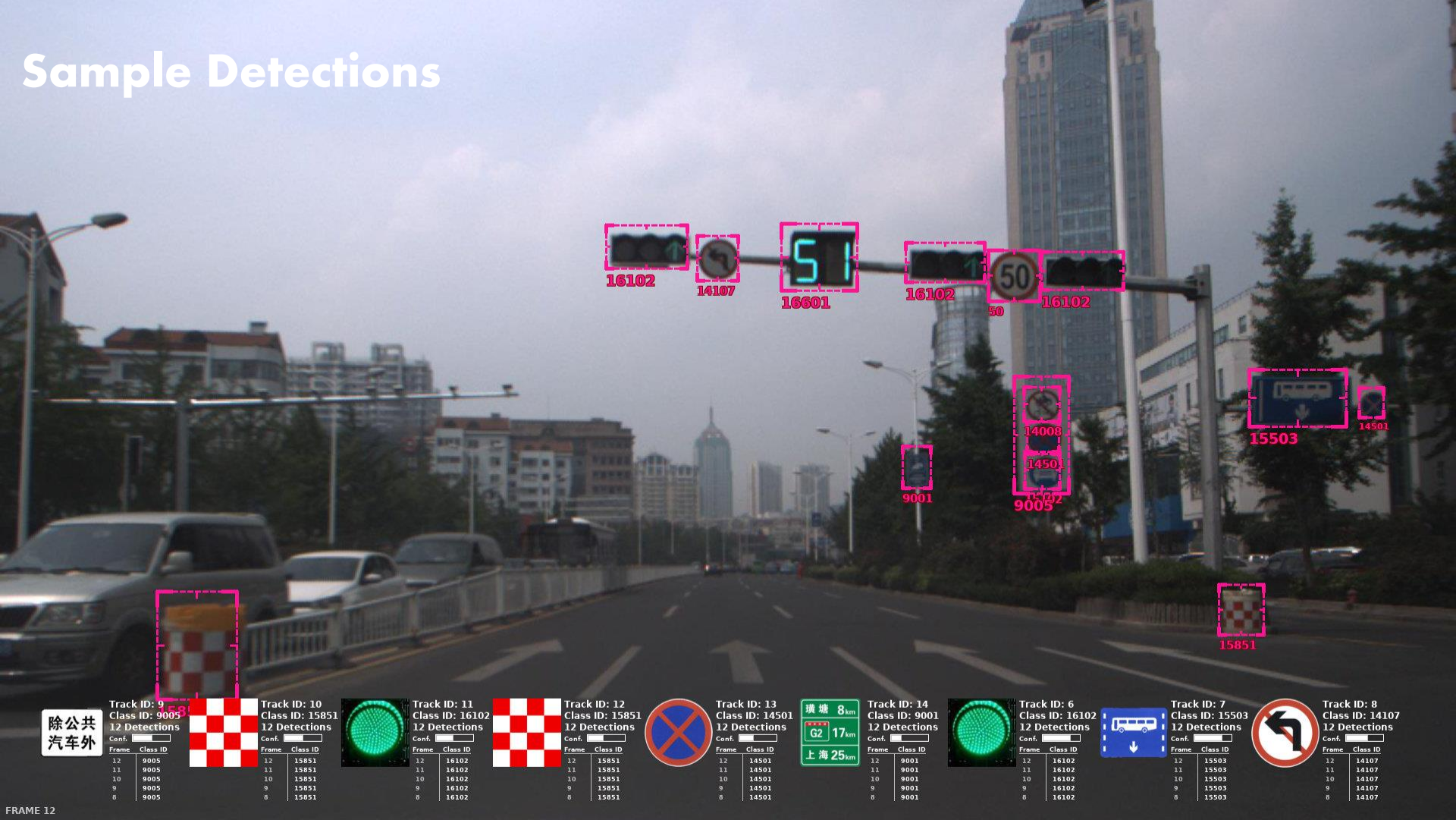
2-3x Speedup using Tensor RT

About : **35 fps** on INT8 on NVIDIA Xavier SoC
: **110 fps** on Titan XP
Inference at Full HD resolution (1920x1080)

Able to extract and classify object as small as 25x25 pixels

Robust to extreme lighting conditions

Sample Detections



Track ID: 9
Class ID: 9005
12 Detections

Frame	Class ID
12	9005
11	9005
10	9005
9	9005
8	9005

Track ID: 10
Class ID: 15851
12 Detections

Frame	Class ID
12	15851
11	15851
10	15851
9	15851
8	15851

Track ID: 11
Class ID: 16102
12 Detections

Frame	Class ID
12	16102
11	16102
10	16102
9	16102
8	16102

Track ID: 12
Class ID: 15851
12 Detections

Frame	Class ID
12	15851
11	15851
10	15851
9	15851
8	15851

Track ID: 13
Class ID: 14501
12 Detections

Frame	Class ID
12	14501
11	14501
10	14501
9	14501
8	14501

Track ID: 14
Class ID: 9001
12 Detections

Frame	Class ID
12	9001
11	9001
10	9001
9	9001
8	9001

Track ID: 6
Class ID: 16102
12 Detections

Frame	Class ID
12	16102
11	16102
10	16102
9	16102
8	16102

Track ID: 7
Class ID: 15503
12 Detections

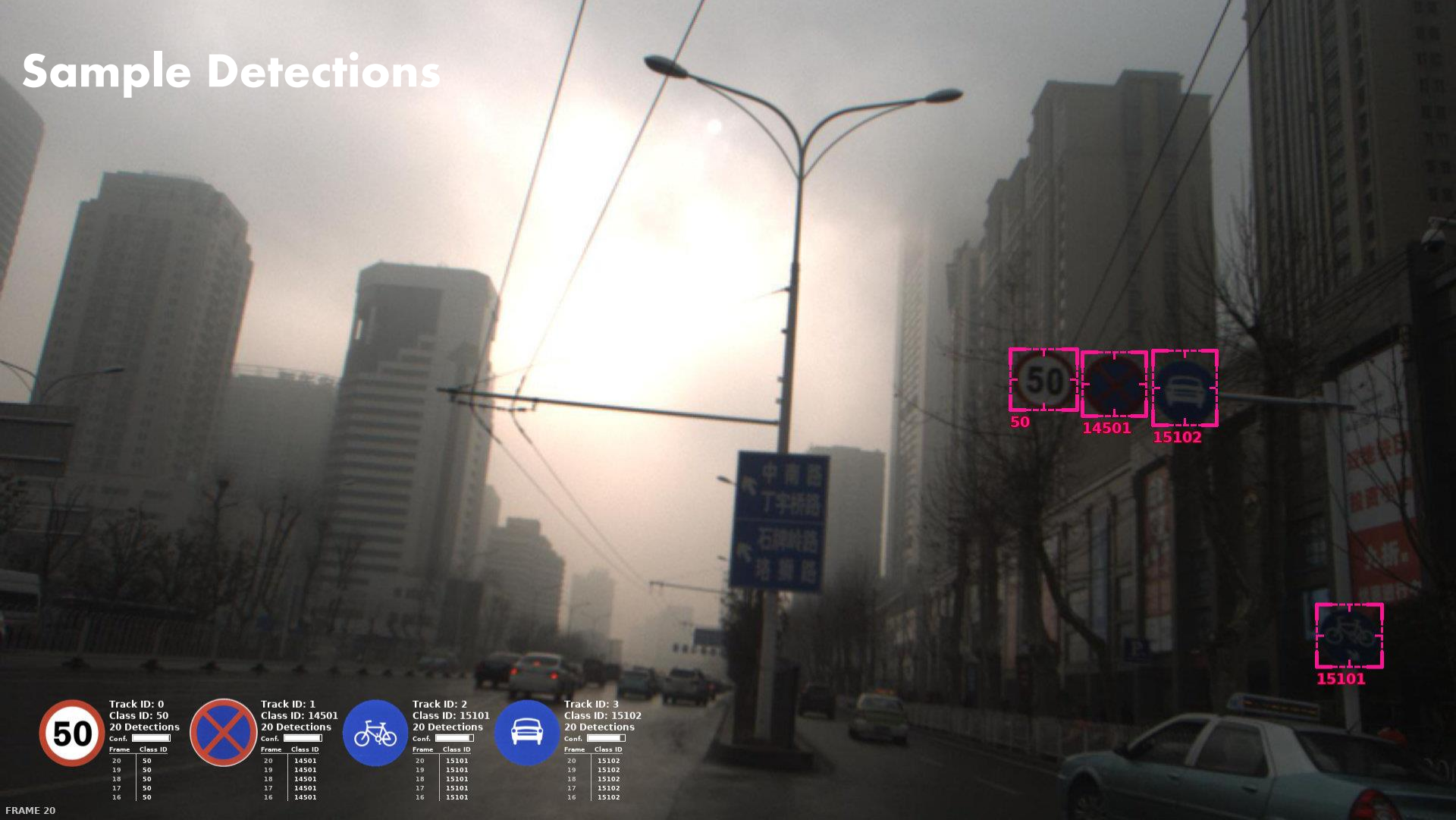
Frame	Class ID
12	15503
11	15503
10	15503
9	15503
8	15503

Track ID: 8
Class ID: 14107
12 Detections

Frame	Class ID
12	14107
11	14107
10	14107
9	14107
8	14107

FRAME 12

Sample Detections



Sample Detections

Track ID: 0
Class ID: 50
20 Detections
Conf.

Frame	Class ID
20	50
19	50
18	50
17	50
16	50

Track ID: 1
Class ID: 14301
20 Detections
Conf.

Frame	Class ID
20	14301
19	14301
18	14301
17	14301
16	14301

Track ID: 2
Class ID: 15102
20 Detections
Conf.

Frame	Class ID
20	15102
19	15102
18	15102
17	15102
16	15102

Track ID: 3
Class ID: 14501
20 Detections
Conf.

Frame	Class ID
20	14501
19	14501
18	14501
17	14501
16	14501

FRAME 20

Online video available at
https://youtu.be/-QtYF0XUZh0

Demo Realtime Object Detection

Advanced Research Lab/NIEU focusses on the development of the AI algorithm, the data processing is done completely in China.
We comply with all Chinese regulations regarding the processing of China data.

Real Time Semantic Segmentation

NAVINFO 四维图新 EUROPE

Ahmed Badar
Computer Vision Engineer
Advanced Research Lab
NavInfo Europe

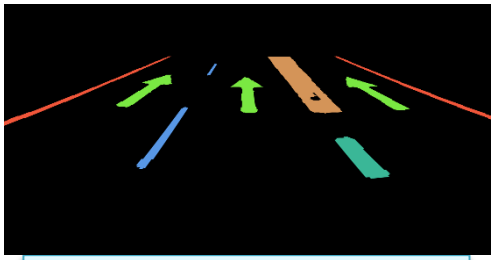
Matti Jukola
Software Engineer
Advanced Research Lab
NavInfo Europe

Real-Time Semantic Segmentation

Deep learning architecture to segment and extract road markings at pixel level

Camera input

Segmentation Network



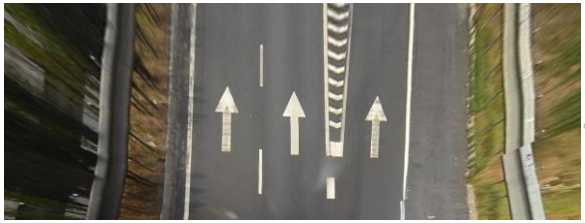
Segmented road markings

Fully Convolutional
Neural Network

Real-Time Semantic Segmentation

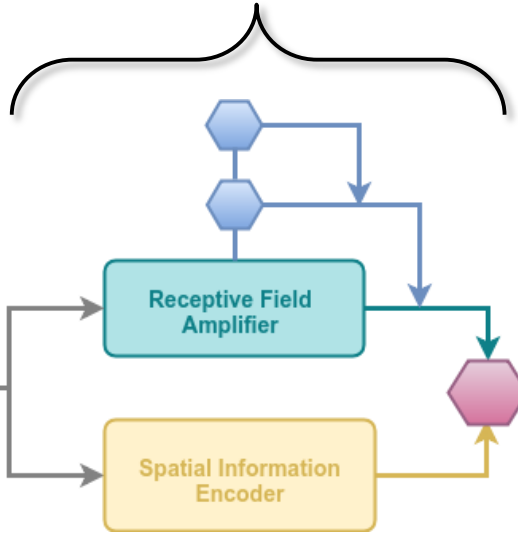
Camera input

Top-view transformation



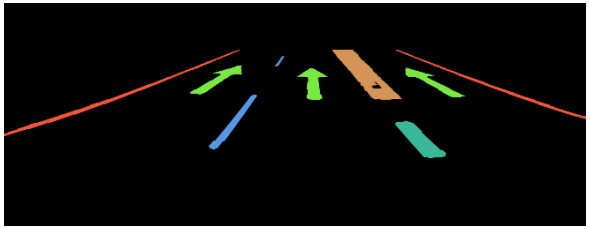
Top-down view

Our model is based on a multi-branch convolutional neural network architecture



Multi-branch segmentation network

Front-view predictions



Front-view transformation

Top-down predictions

Real-Time Semantic Segmentation

Currently supports 40 **Road Marking Classes,** including**:**

- Lane lines
- Arrows
- Text

Real Time Performance

3x Speedup using Tensor RT

About : **90 fps** on NVIDIA Xavier SoC

: **300 fps** on Titan XP

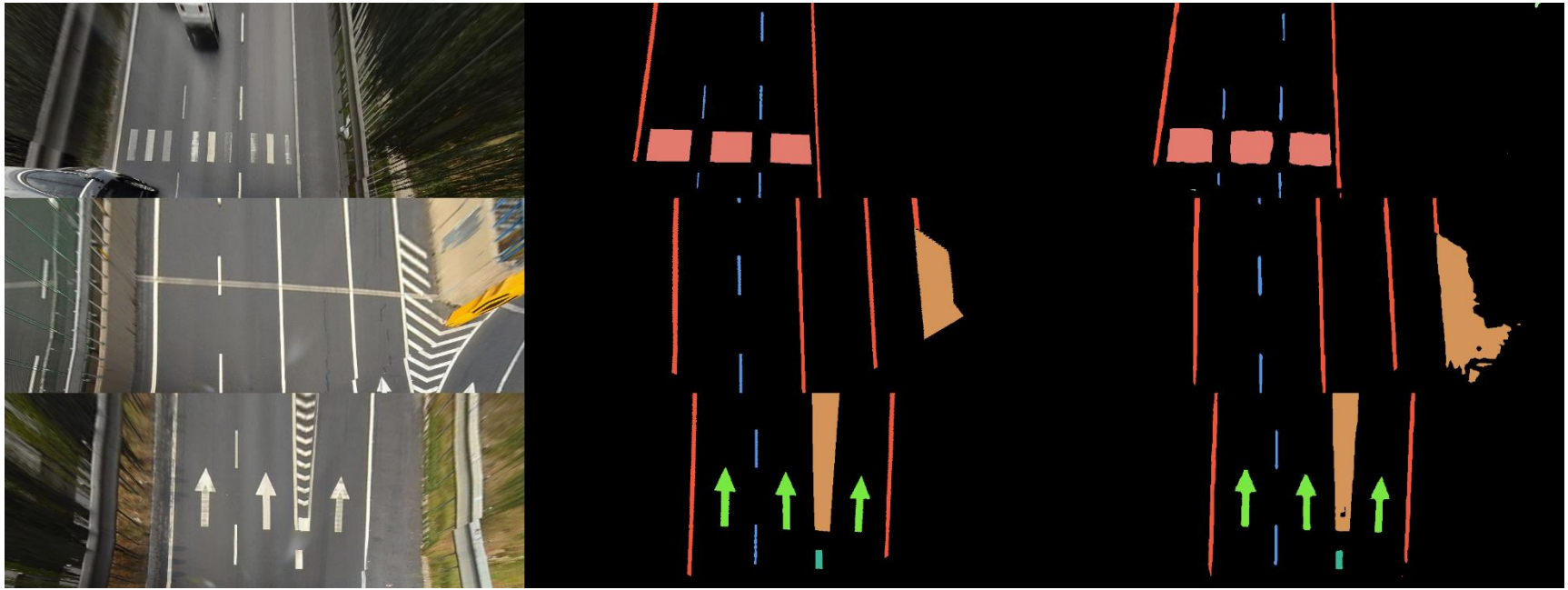
Inference at 1024 × 384 image sizes

This **includes** image transformations (top and front view)

NAVINFO
四维图新

Sample Segmentations

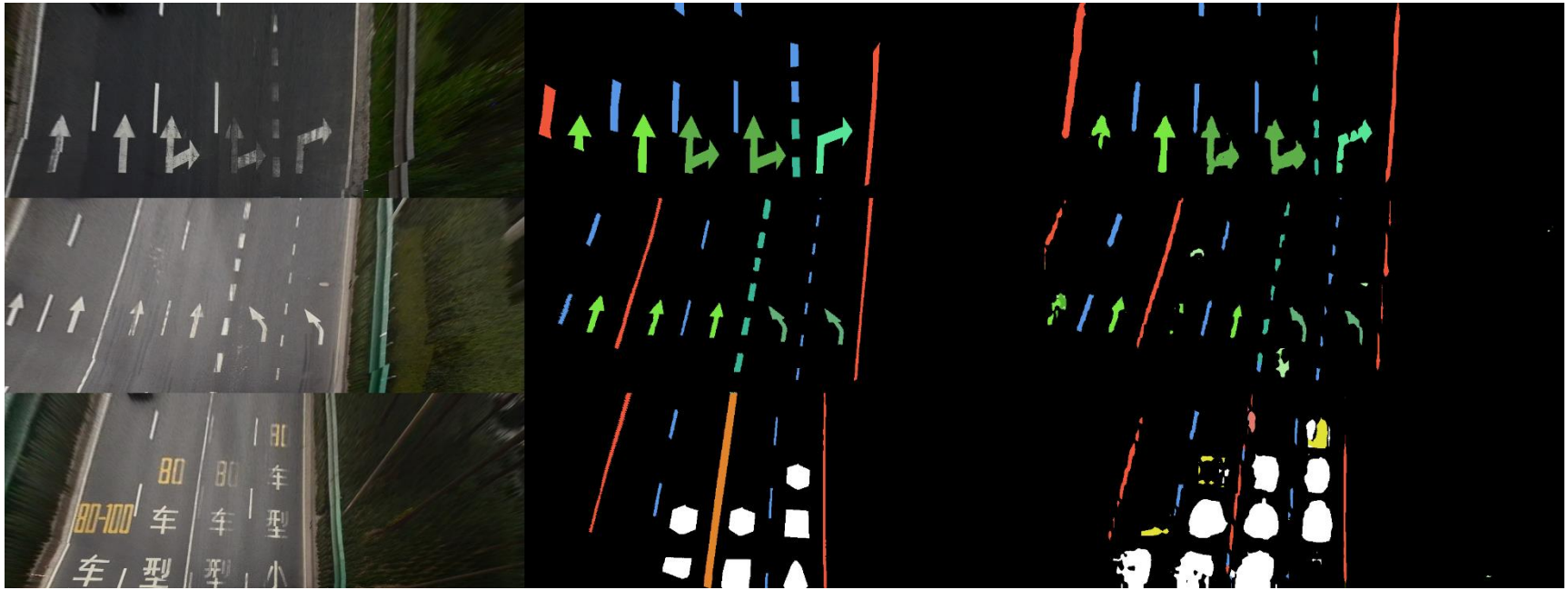
Input Image *Ground Truth* *Prediction*



NAVINFO 四维图新

Sample Segmentations

Input Image *Ground Truth* *Prediction*



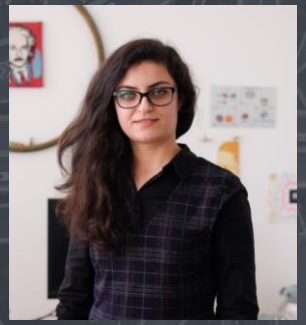
Online video available at
https://youtu.be/E4hU-COkHDo

Demo Realtime Semantic Segmentation

Advanced Research Lab/NIEU focusses on the development of the AI algorithm, the data processing is done completely in China.
We comply with all Chinese regulations regarding the processing of China data.

Real Time Scene Understanding

NAVINFO 四维图新 EUROPE

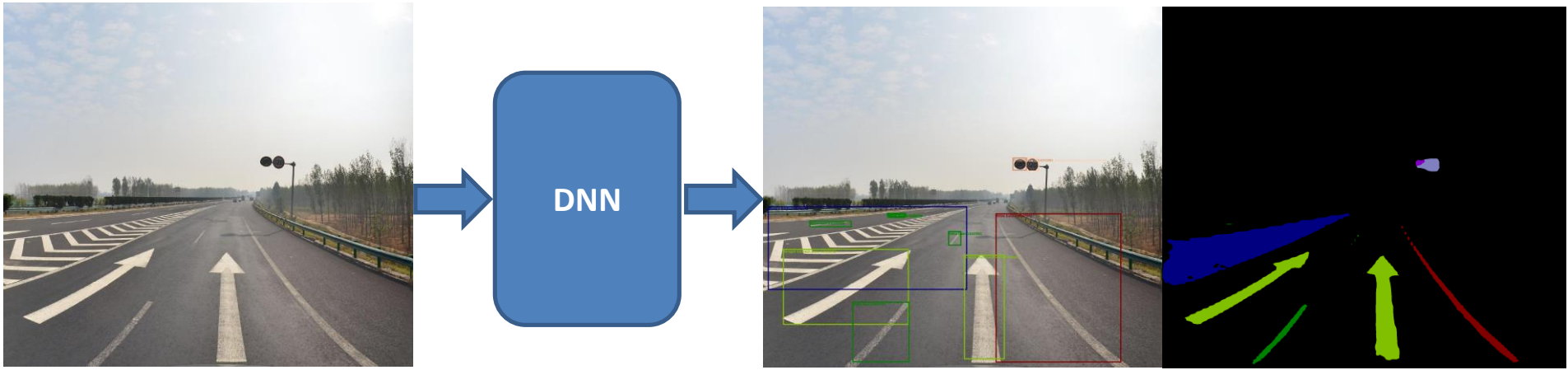


Elahe Arani
Senior AI Researcher
Advanced Research Lab
NavInfo Europe

Mahmoud Gamal
Computer Vision Engineer
Advanced Research Lab
NavInfo Europe

Real-Time Scene Understanding

A real time ***unified*** object detection and semantic segmentation for autonomous driving cars/HD mapping.



Joint Object Detection and Segmentation

Currently supports **40** Road Marking Classes and **350+** Road Sign classes including:

- Traffic Signs
- Gantry Signboards
- Traffic Lights
- Digital Traffic Signs
- Lane Markings
- Text
- Arrows

Performance

Inference at 512x512 image sizes

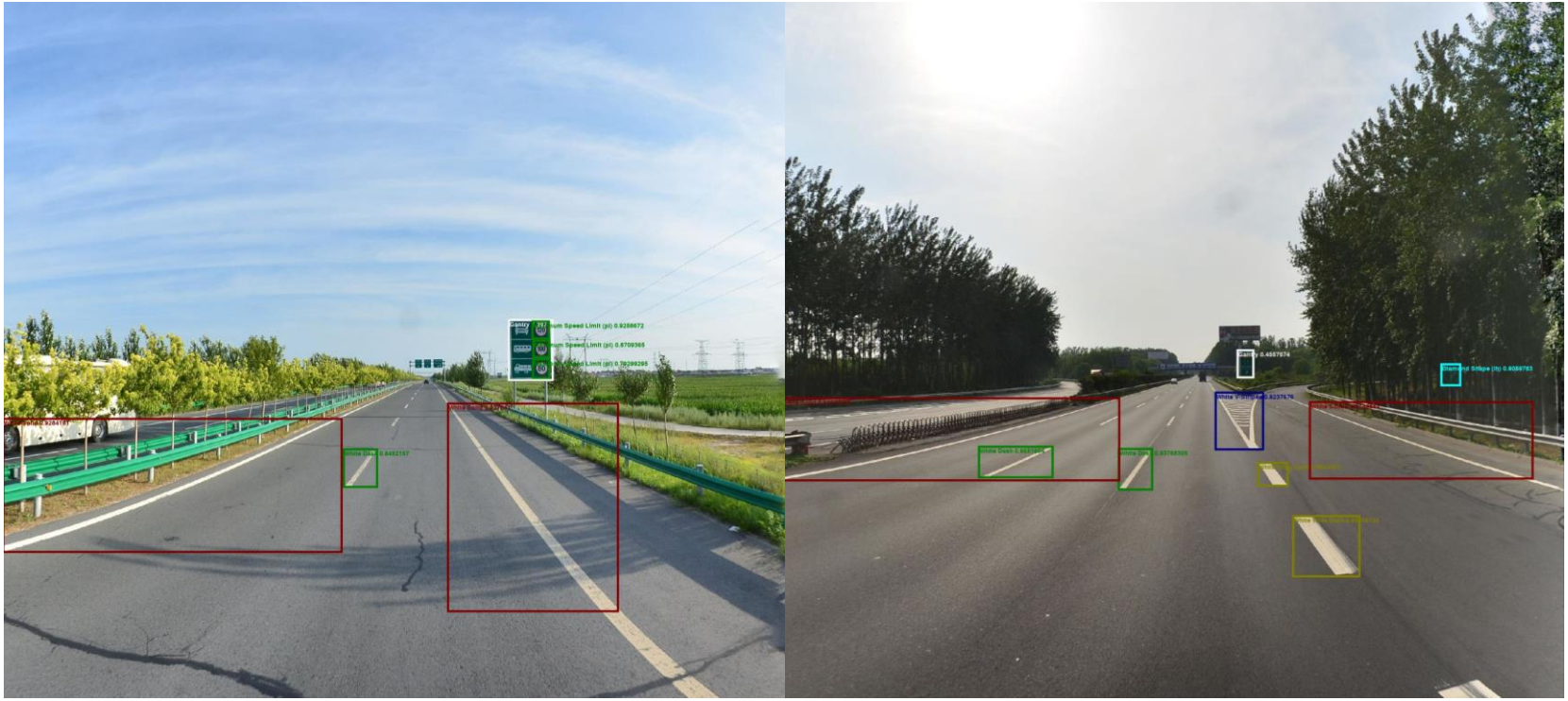
About **45 FPS** on Titan XP

Other Features Supported

- Guard Rails
- Curbs
- Speed Limits on Road

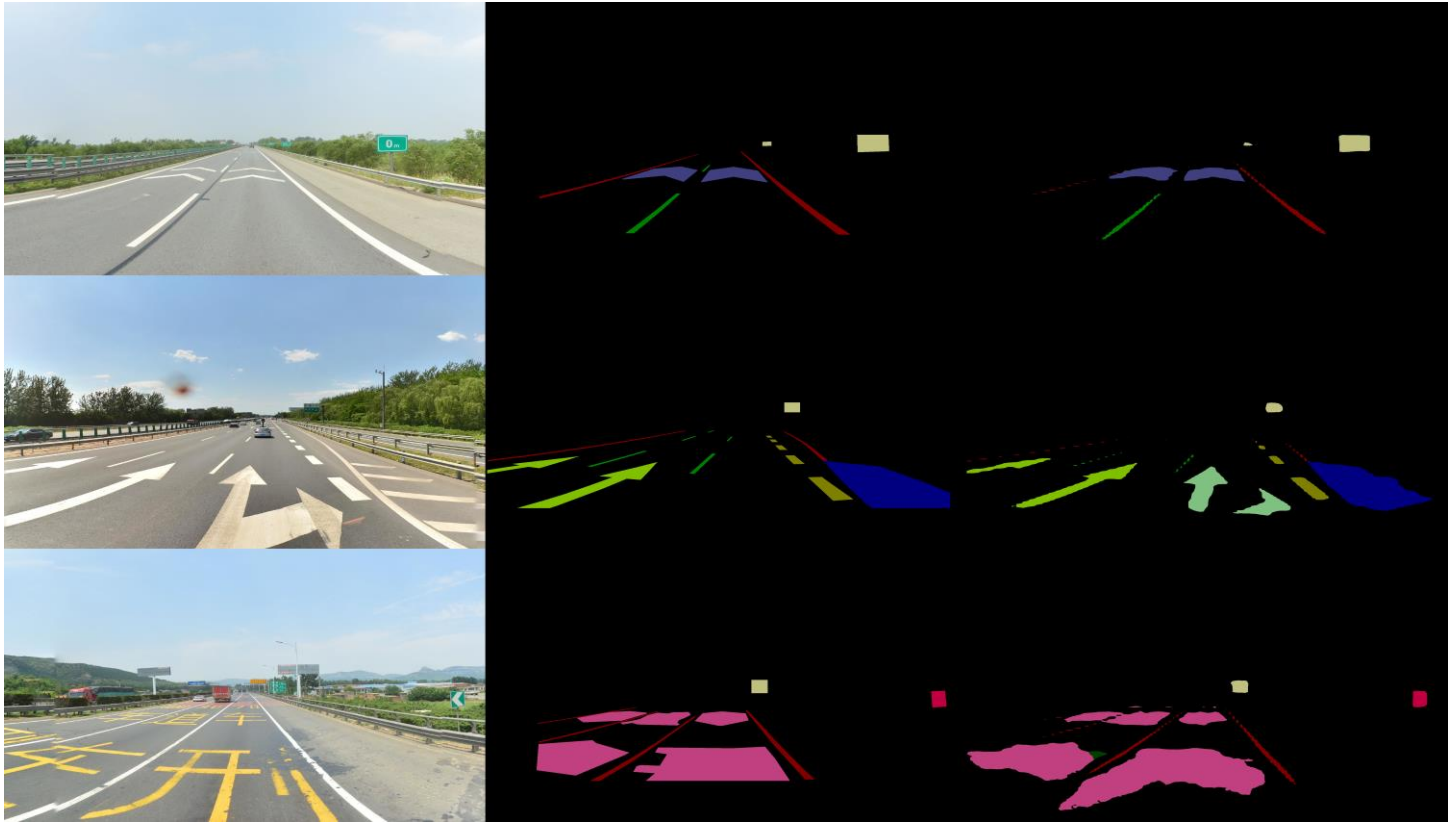
NAVINFO
四维图新

Sample Detections



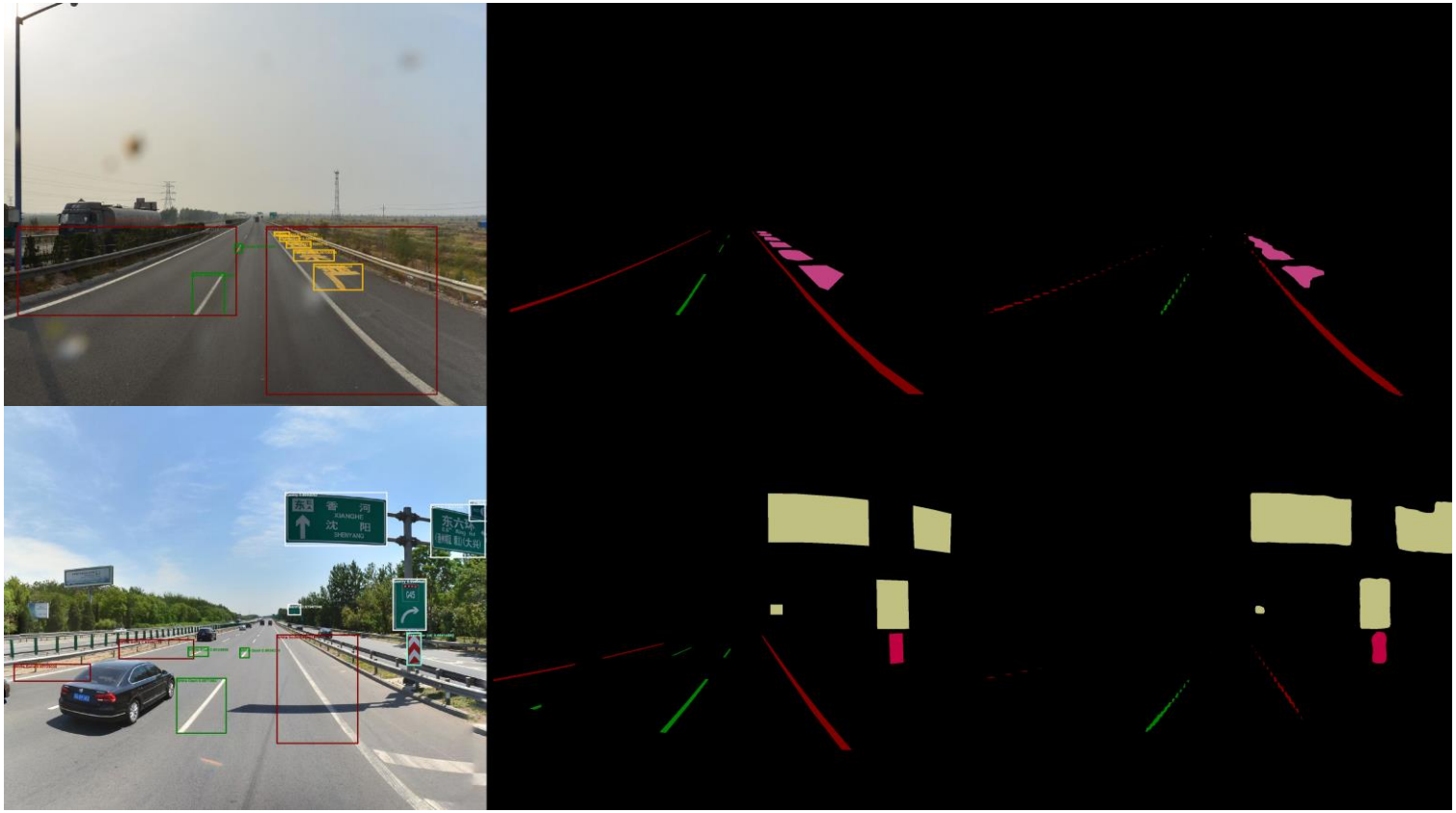
Sample Segmentation

Input Image *Ground Truth* *Prediction*



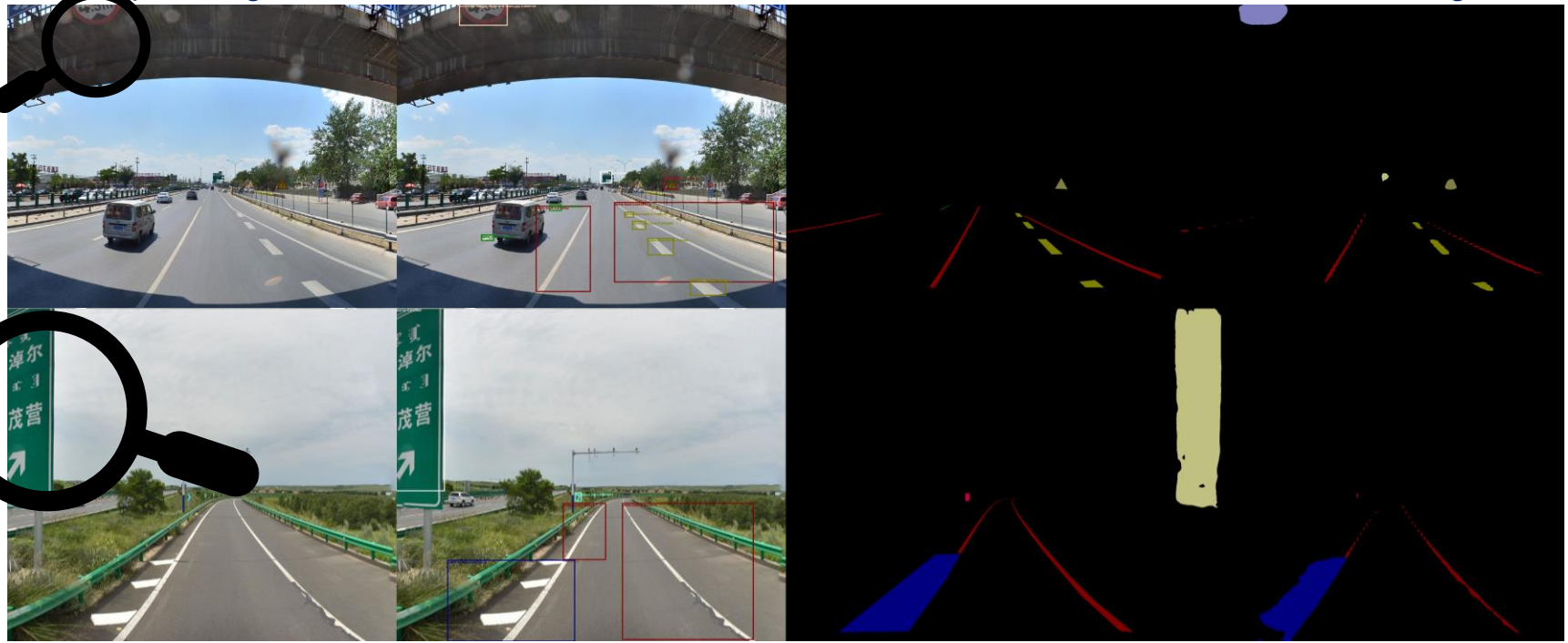
Joint Detections and Segmentation

Prediction (Dec) *Ground Truth* *Prediction (Seg)*



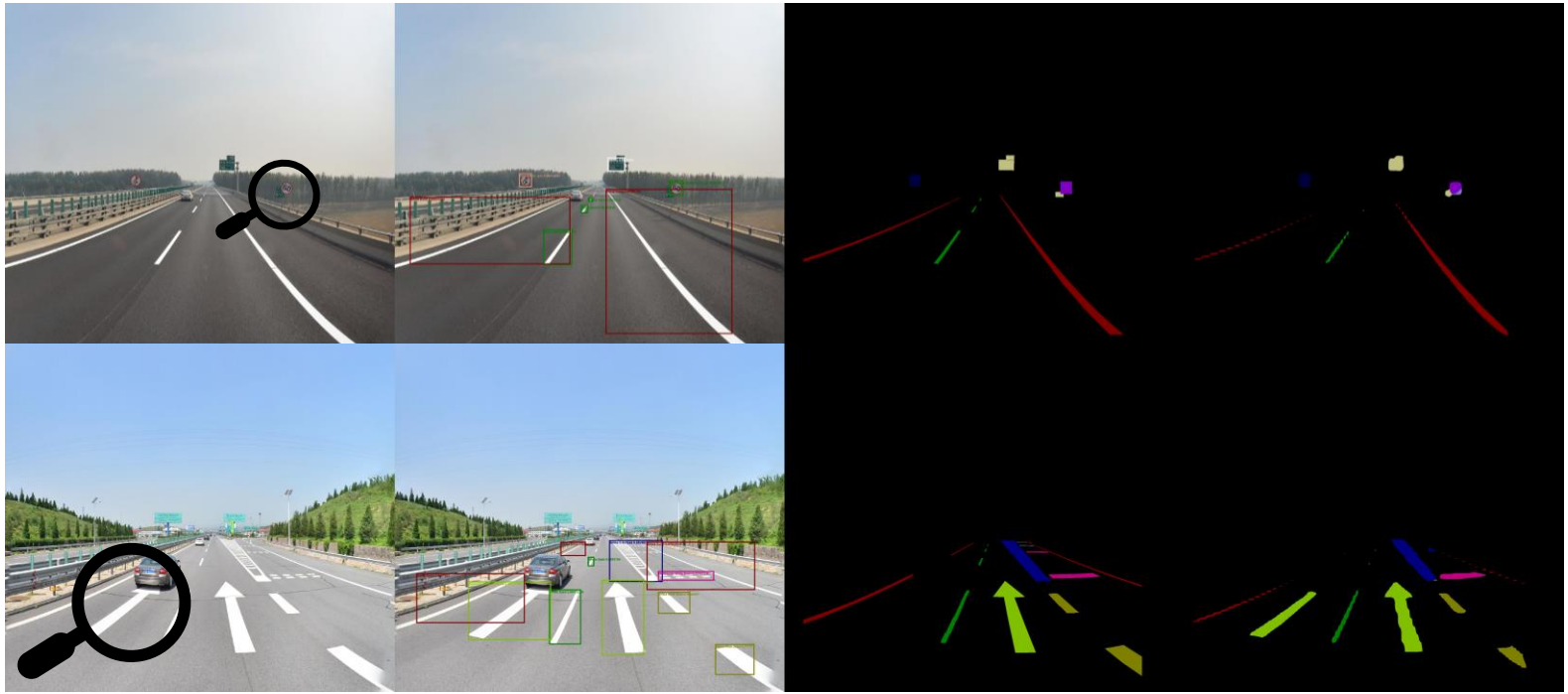
Occlusion Handling

Input Image *Prediction (Dec)* *Ground Truth* *Prediction (Seg)*



Occlusion Handling

Input Image *Prediction (Dec)* *Ground Truth* *Prediction (Seg)*



Online video available at
https://youtu.be/NJVNFfueKb4

Demo Realtime Scene Understanding

Advanced Research Lab/NIEU focusses on the development of the AI algorithm, the data processing is done completely in China.
We comply with all Chinese regulations regarding the processing of China data.

Real-Time Object Detection and Semantic Segmentation

Wei Li
Corporate development manager

Geetank Raipuria
Computer Vision Engineer

NavInfo Europe

Presented at GTC 2019
Session S9351