Scaling up Precise Localization for Autonomous Robots

State Estimation, Multi-Task Learning, and Beyond

Andrei Bârsan — PhD Candidate @ University of Toronto, Research Scientist @ Waabi
andreibarsan.github.io — @andreib — 2021-06-09

Autonomous Driving

Autonomous Driving

- Driving is a leading cause of death in developed countries

Autonomous Driving

- Driving is a leading cause of death in developed countries
- Enhance or replace human drivers — multiple autonomy **levels**

Autonomous Driving

- Driving is a leading cause of death in developed countries
- Enhance or replace human drivers — multiple autonomy **levels**
- Maps enable advanced autonomy and improve **safety**

Autonomous Driving

- Driving is a leading cause of death in developed countries
- Enhance or replace human drivers — multiple autonomy **levels**
- Maps enable advanced autonomy and improve **safety**
- Leveraging maps requires precise info of **where** the vehicle is located

Agenda

Agenda

1. The Role of Localization in Self-Driving

Agenda

1. The Role of Localization in Self-Driving
2. Scalable Map-Based Localization

Agenda

1. The Role of Localization in Self-Driving
2. Scalable Map-Based Localization
3. How Good Does Localization Need to Be?

Agenda

1. The Role of Localization in Self-Driving
2. Scalable Map-Based Localization
3. How Good Does Localization Need to Be?
4. The Future

Autonomy Levels

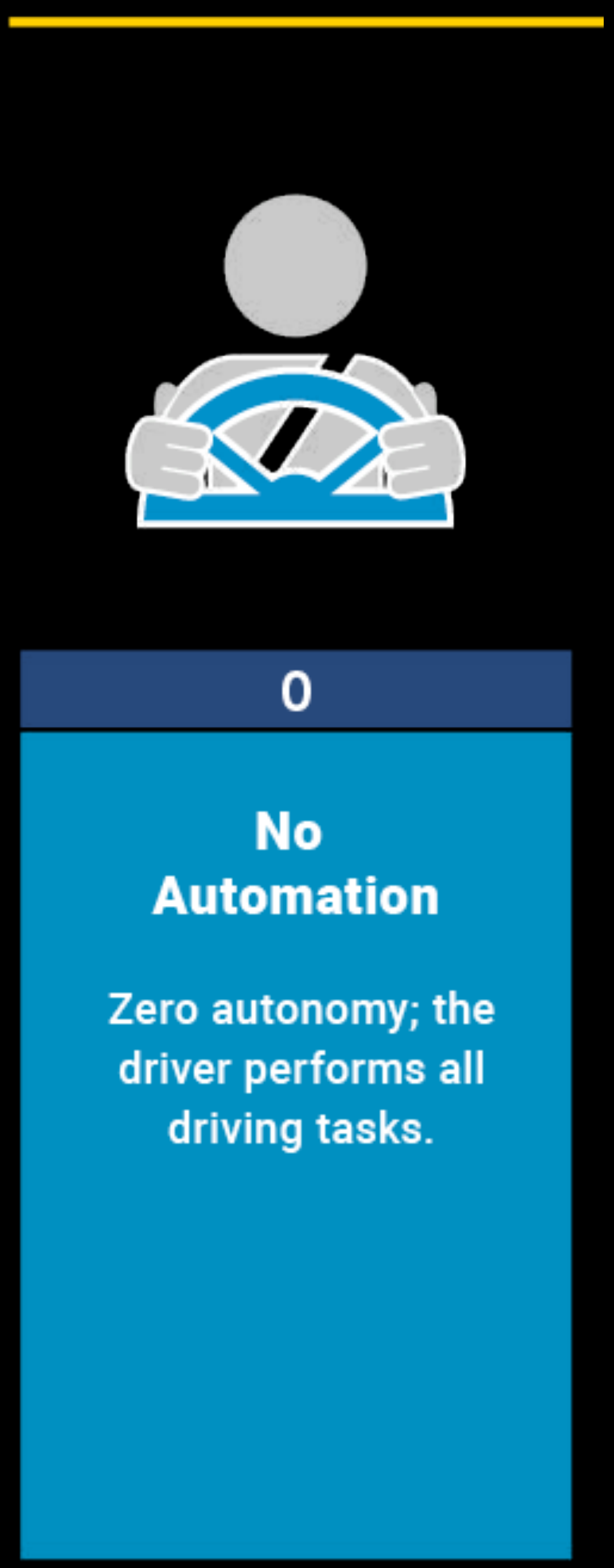


Image source: nhtsa.gov

13069b-082317-v8

Autonomy Levels

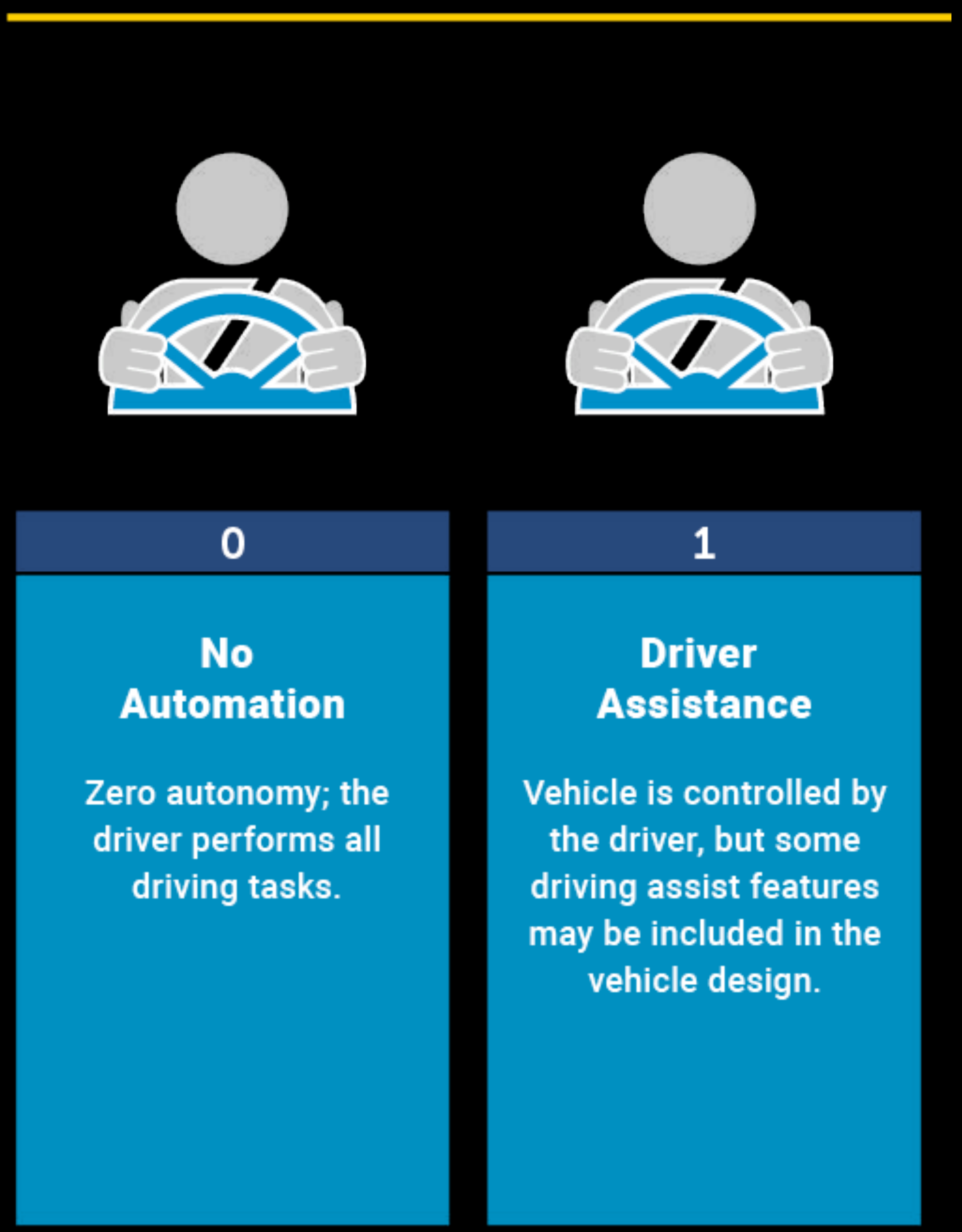


Image source: nhtsa.gov

13069b-082317-v8

Autonomy Levels

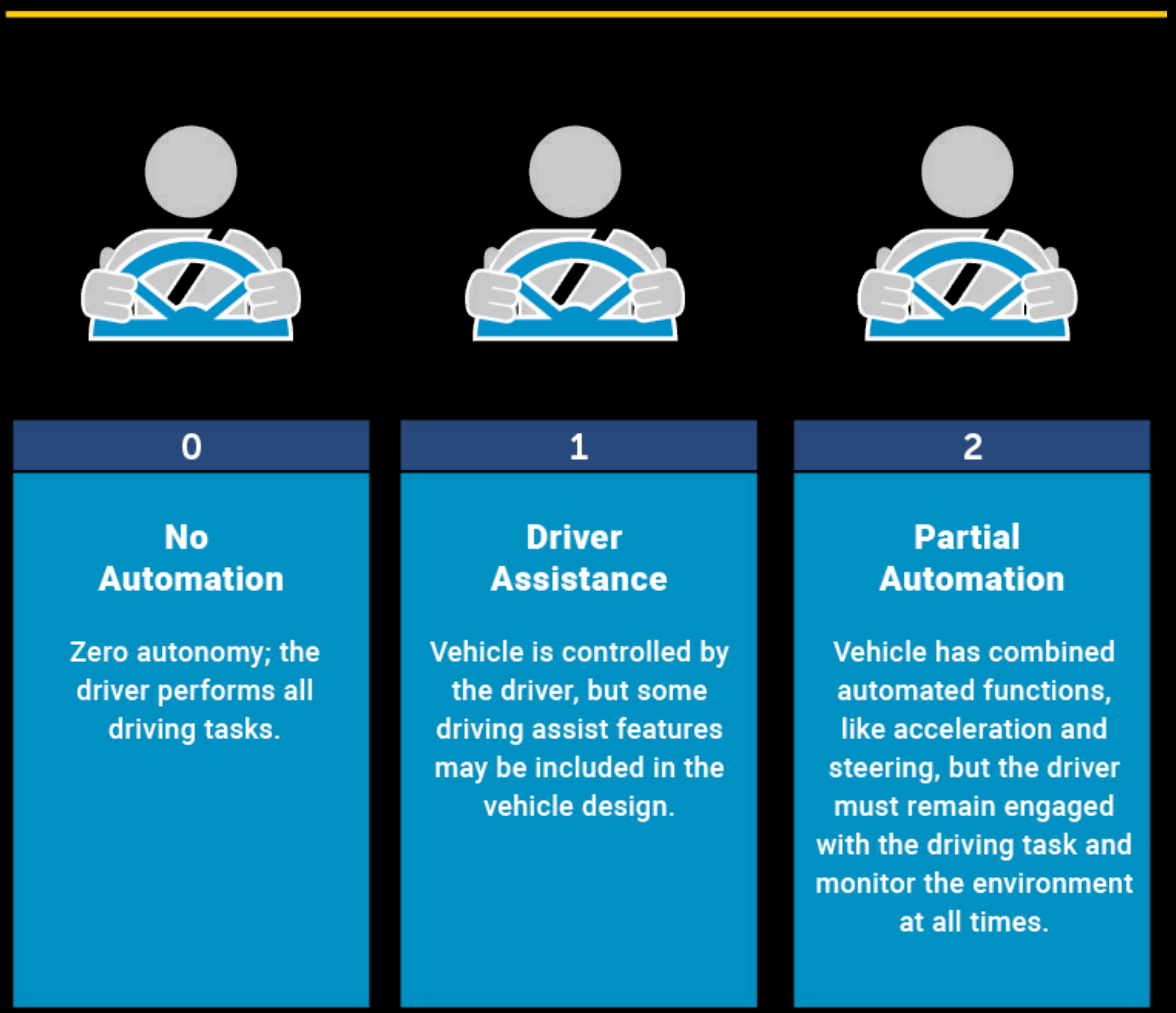


Image source: nhtsa.gov

13069b-082317-v8

Autonomy Levels

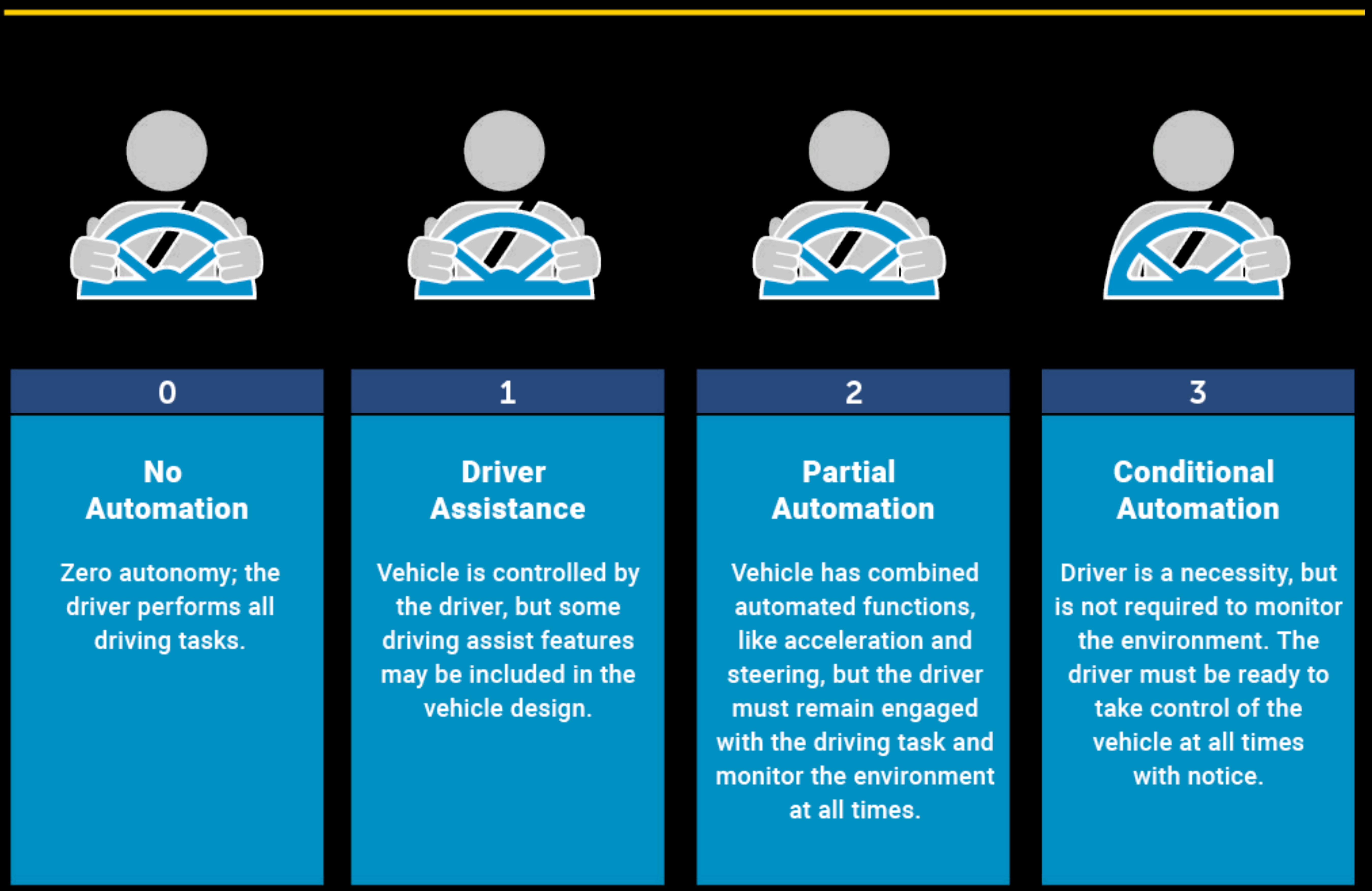


Image source: nhtsa.gov

13069b-082317-v8

Autonomy Levels

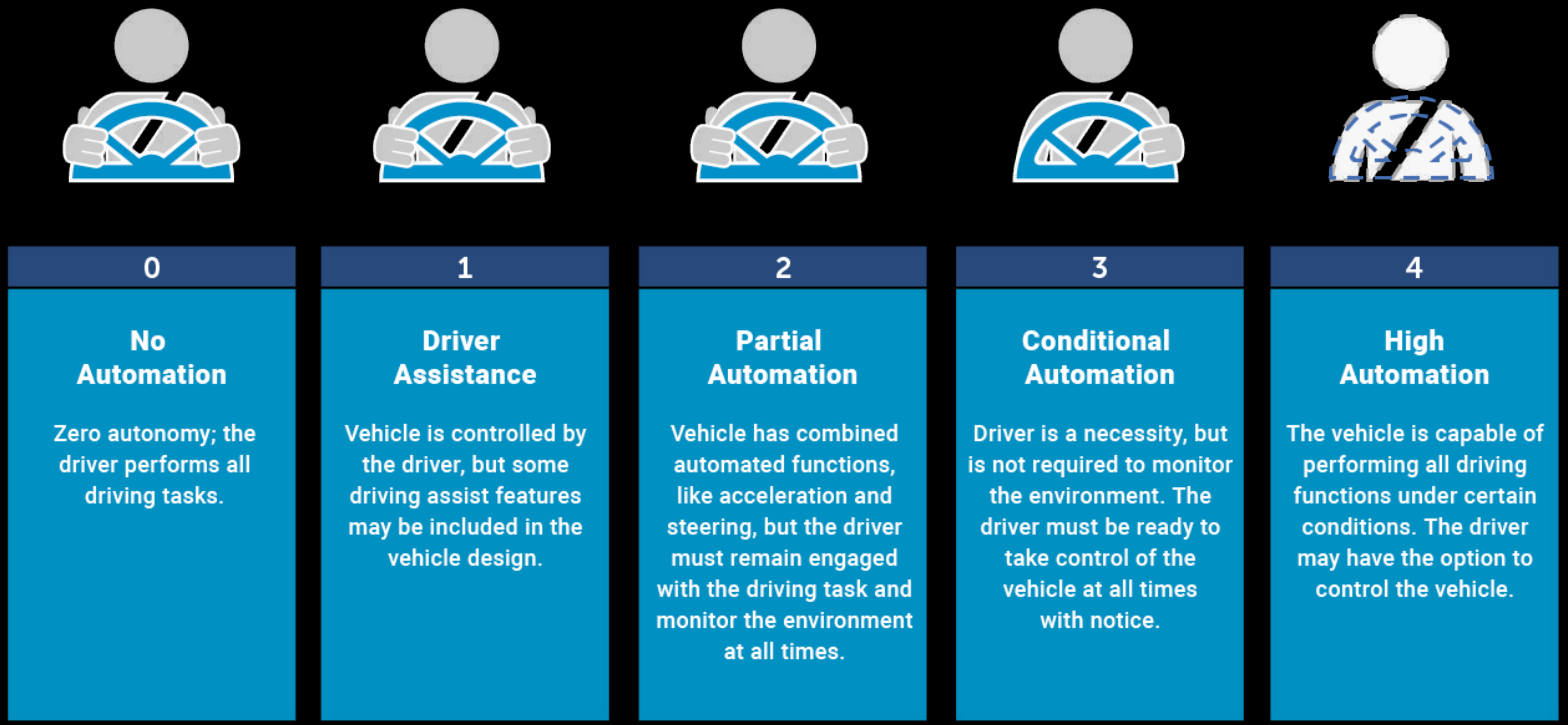


Image source: nhtsa.gov

13069b-082317-v8

Autonomy Levels

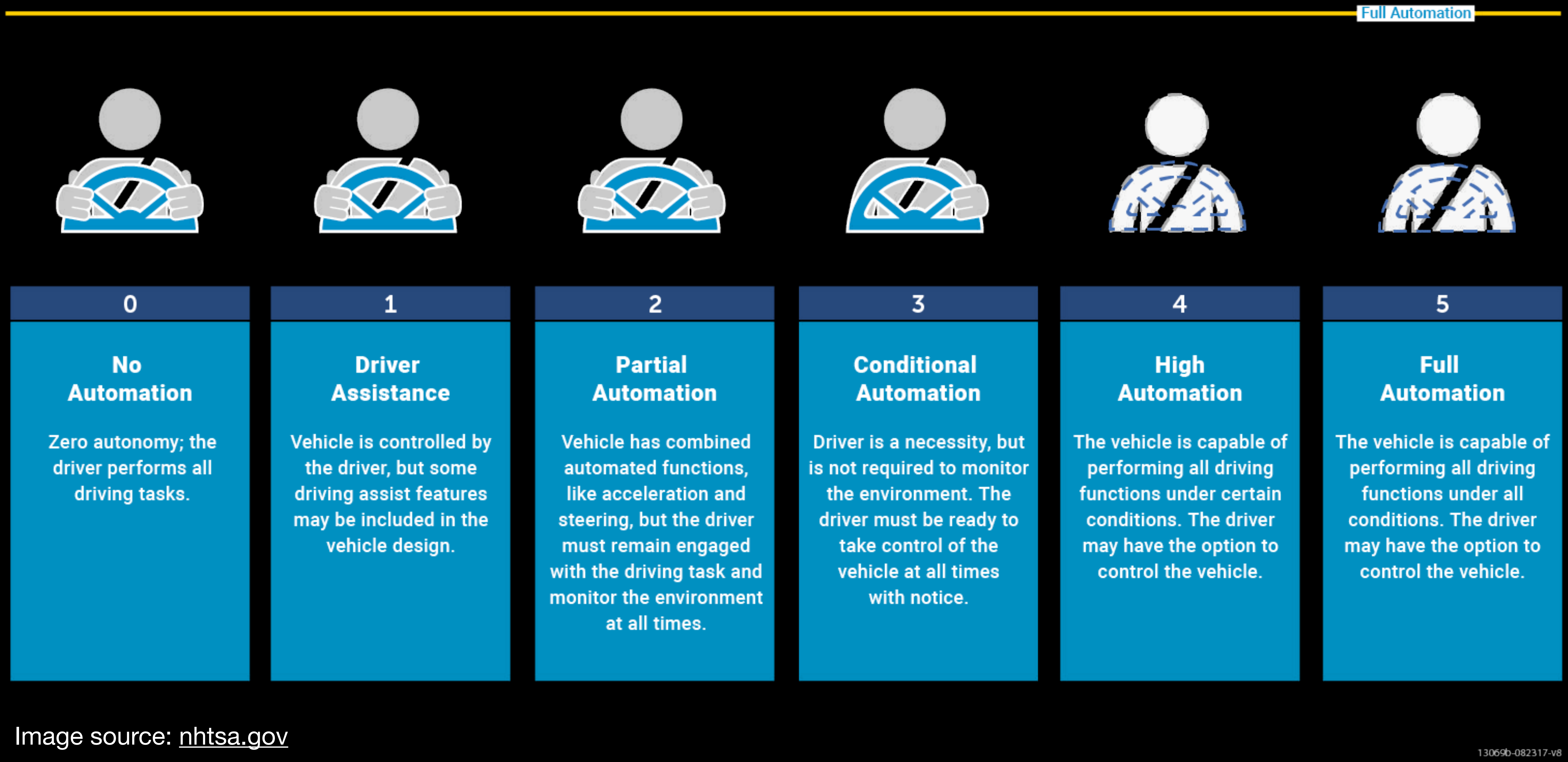


Image source: nhtsa.gov

Autonomy Levels

This talk

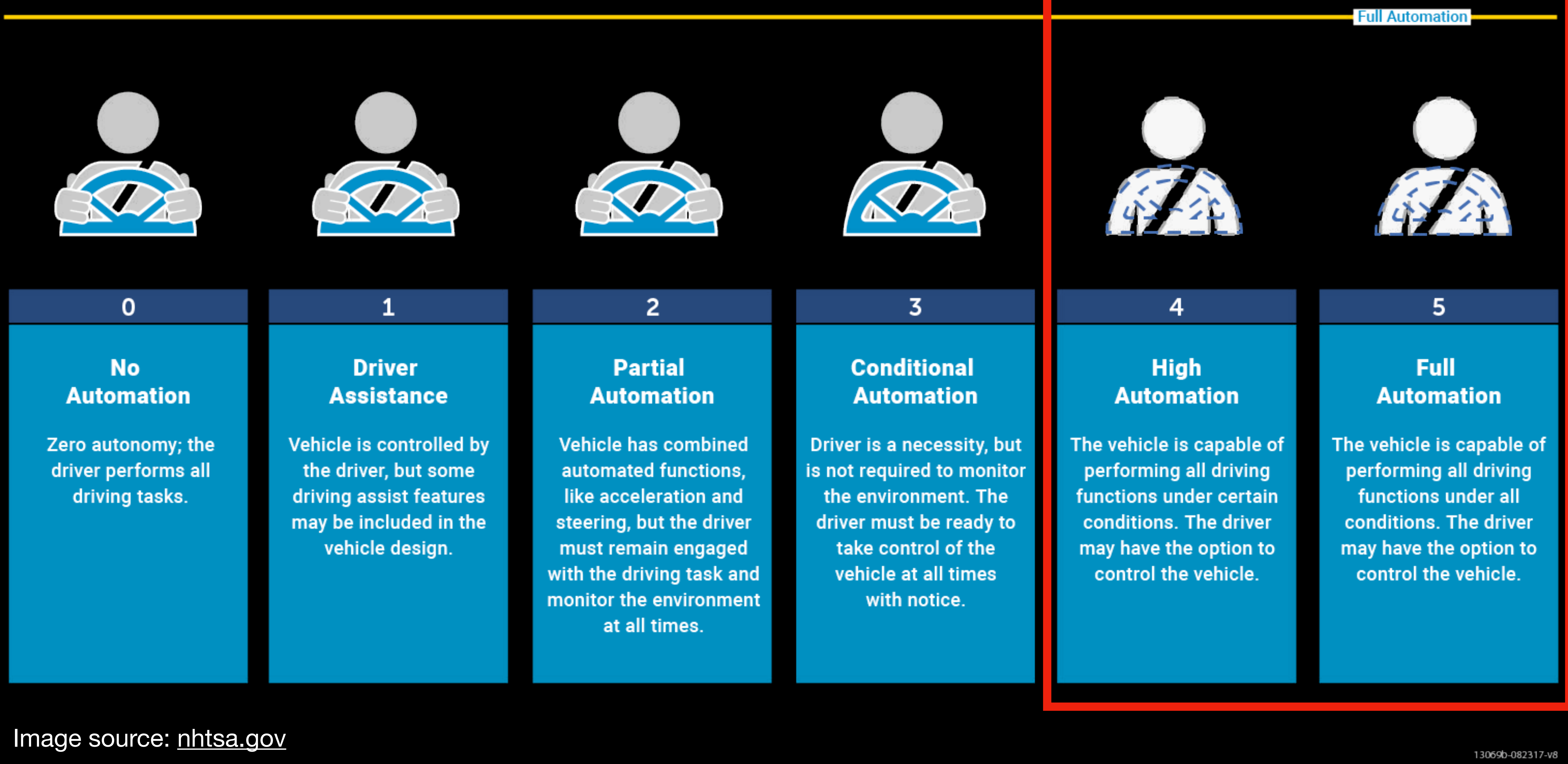
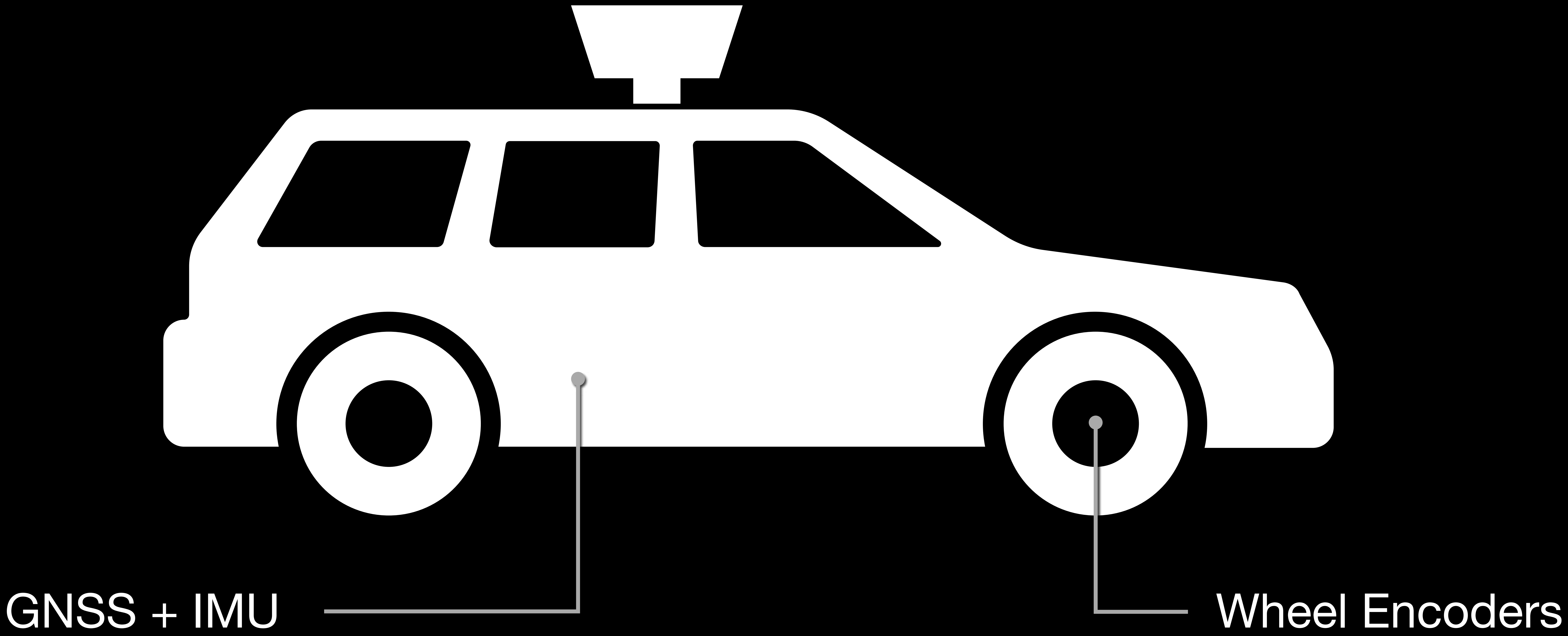


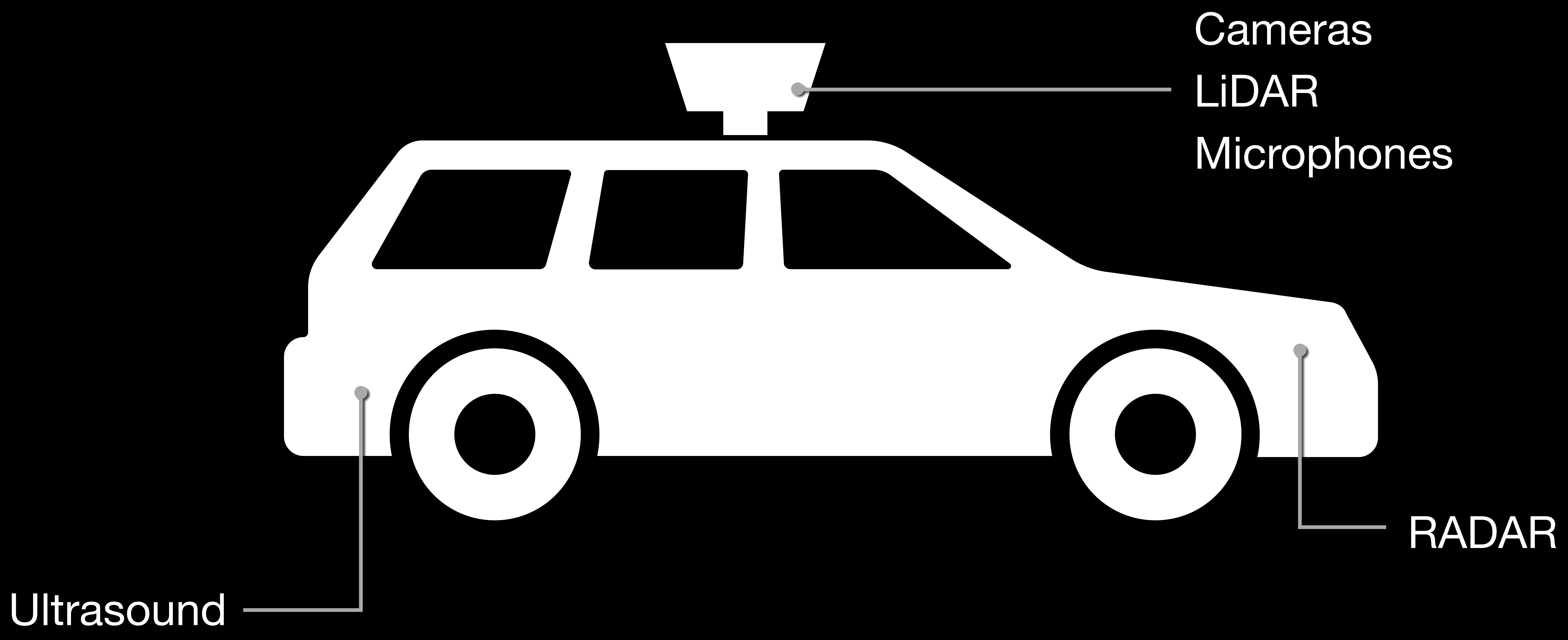
Image source: nhtsa.gov

Autonomy Sensors
Proprioception & GNSS



Autonomy Sensors

Perception



Autonomy Input

Maps → Perception → Prediction → Planning → Control

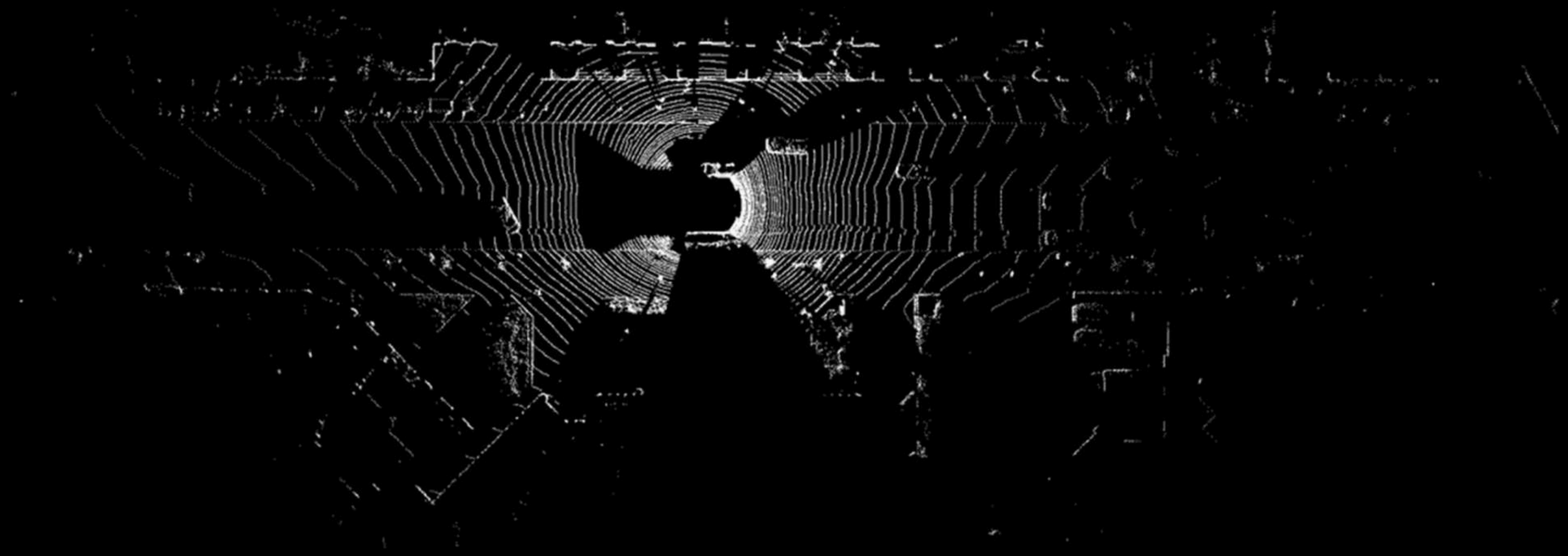
Sensors

Detections Tracks

Long Term Predictions

Motion Trajectory

Steering / Acceleration

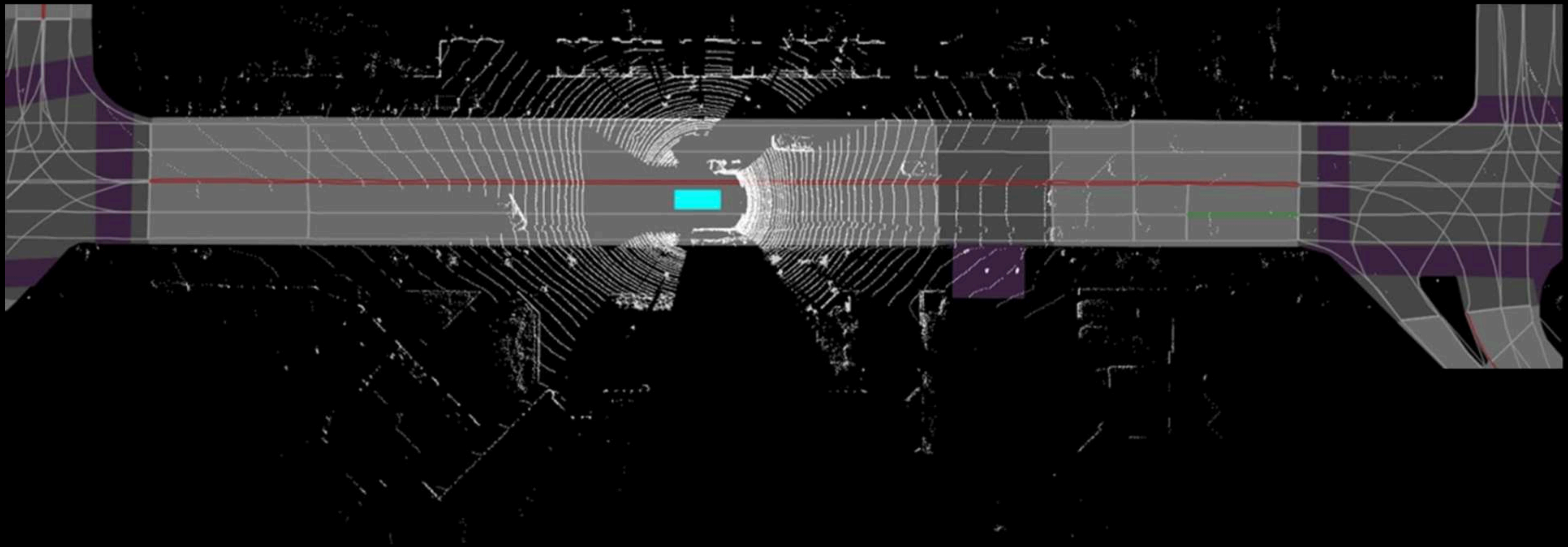


Slide credits: Raquel Urtasun & Sergio Casas

HD Maps

Maps → **Perception** → **Prediction** → **Planning** → **Control**

Sensors | Detections | Tracks | Long Term Predictions | Motion Trajectory | Steering / Acceleration



Slide credits: Raquel Urtasun & Sergio Casas

Perception

Maps

Sensors

→

Perception

Detections Tracks

→

Prediction

Long Term Predictions

→

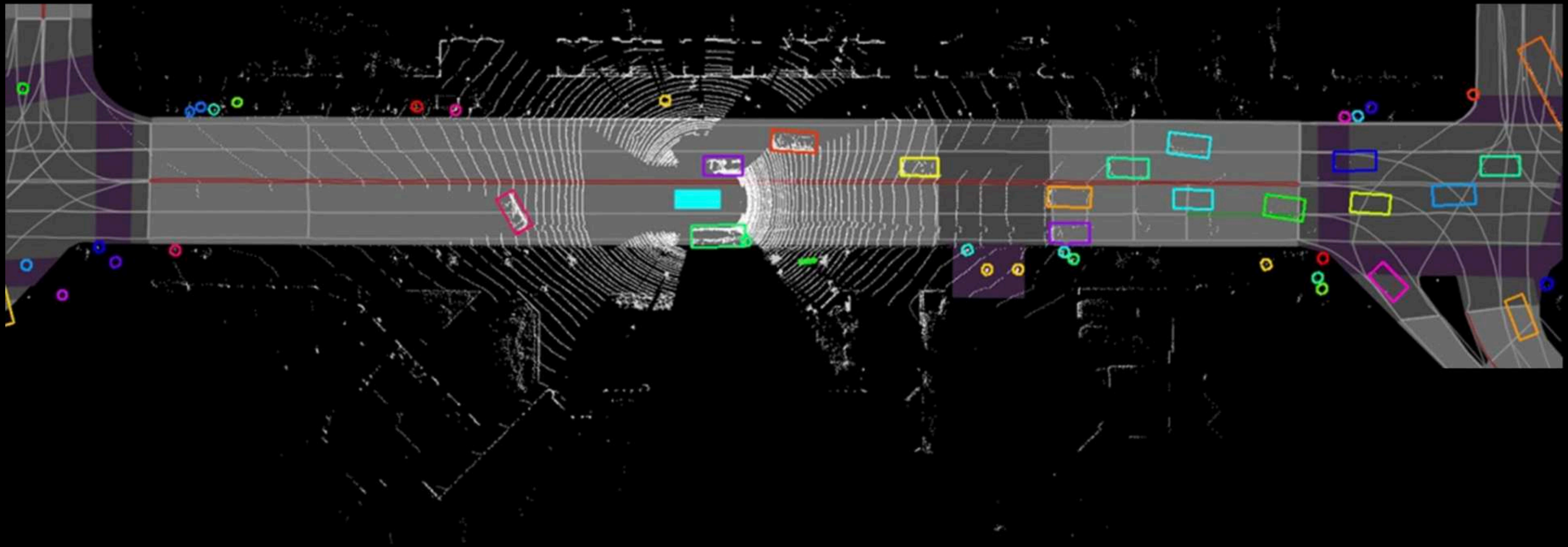
Planning

Motion Trajectory

→

Control

Steering / Acceleration

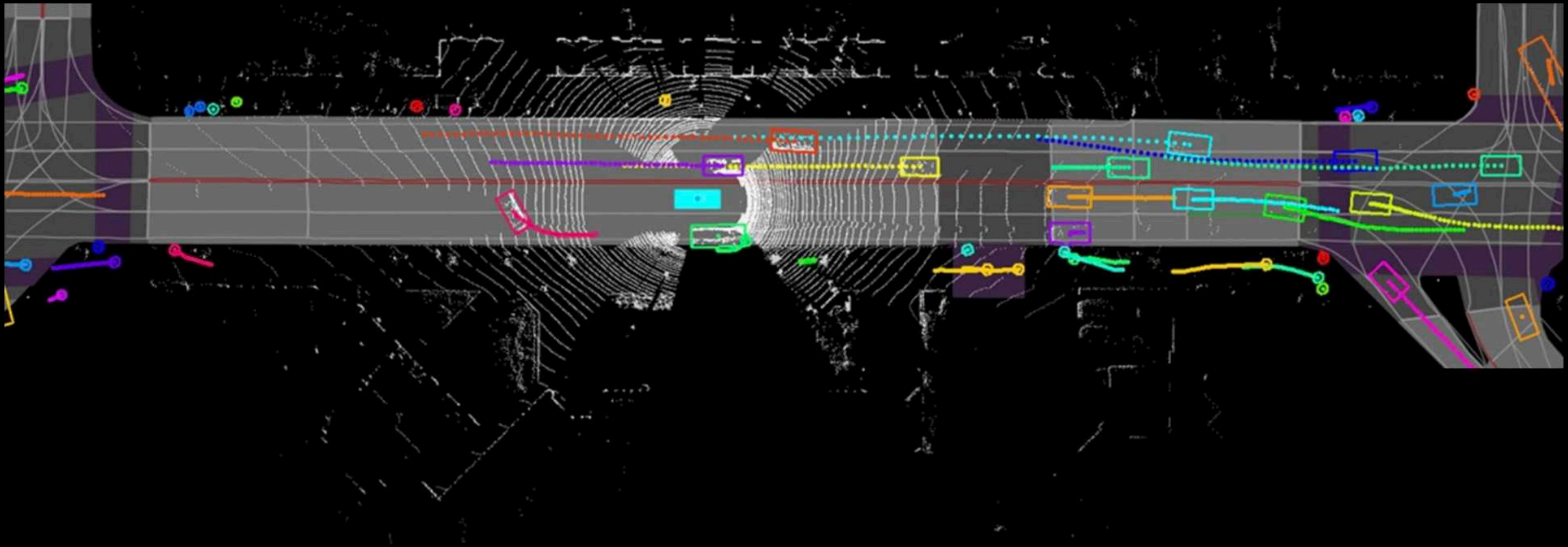


Slide credits: Raquel Urtasun & Sergio Casas

Prediction

Maps → Perception → **Prediction** → Planning → Control

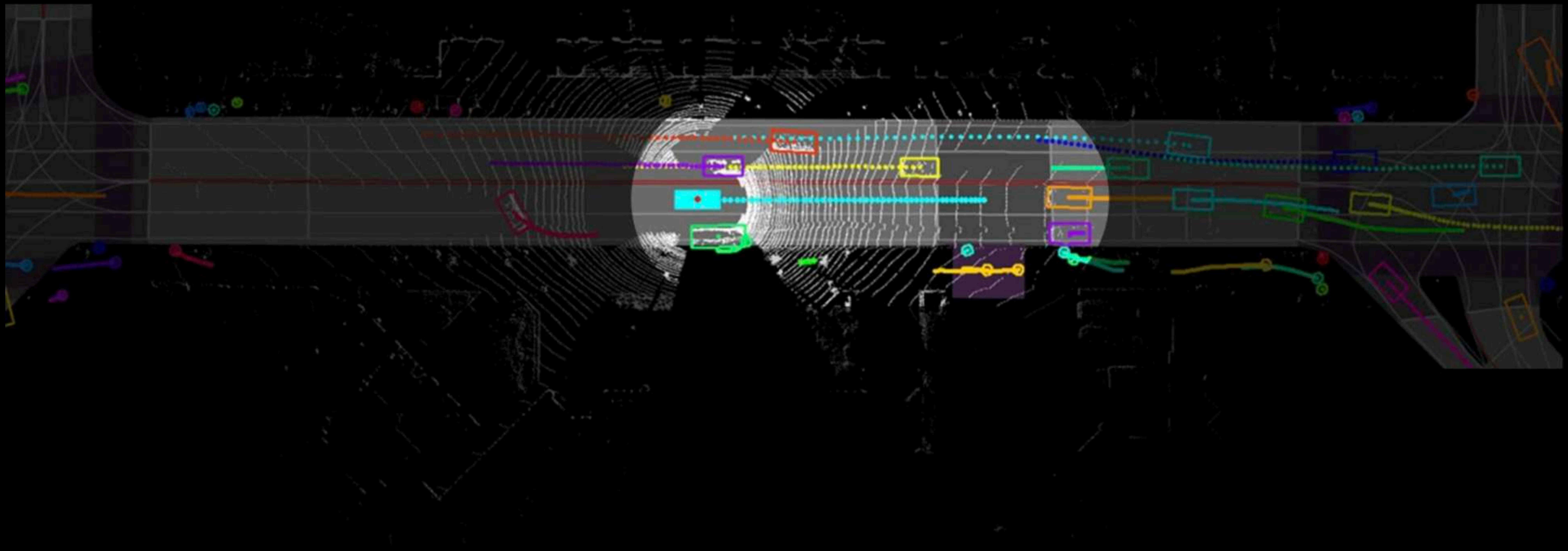
Maps	Perception		Prediction	Planning	Control
Sensors	Detections	Tracks	Long Term Predictions	Motion Trajectory	Steering / Acceleration



Slide credits: Raquel Urtasun & Sergio Casas

Motion Planning

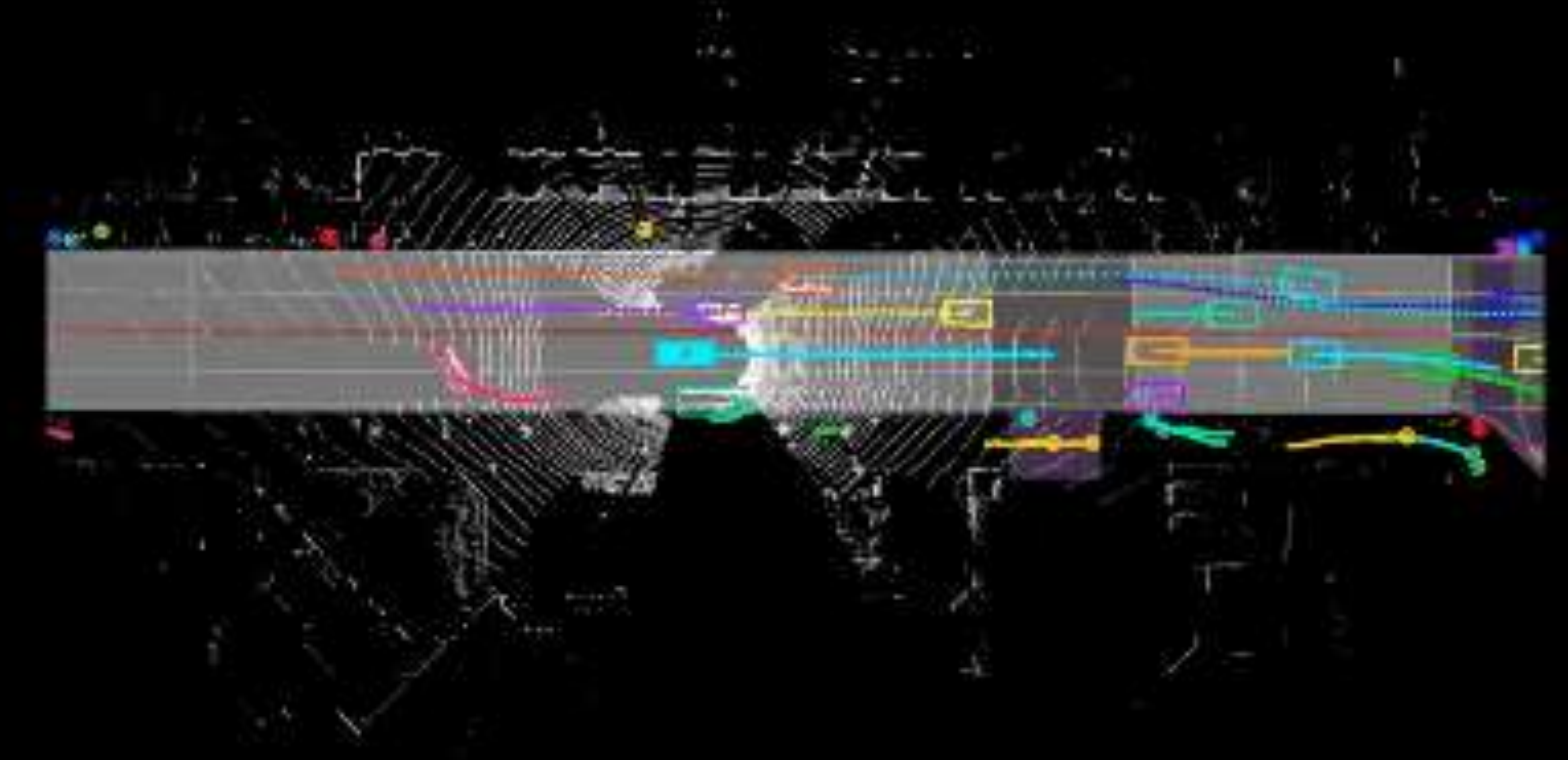
Maps	→	Perception		→	Prediction	→	Planning	→	Control
Sensors		Detections	Tracks		Long Term Predictions		Motion Trajectory		Steering / Acceleration



Slide credits: Raquel Urtasun & Sergio Casas

Vehicle Control

Maps	→	Perception		→	Prediction	→	Planning	→	Control
Sensors		Detections	Tracks		Long Term Predictions		Motion Trajectory		Steering / Acceleration

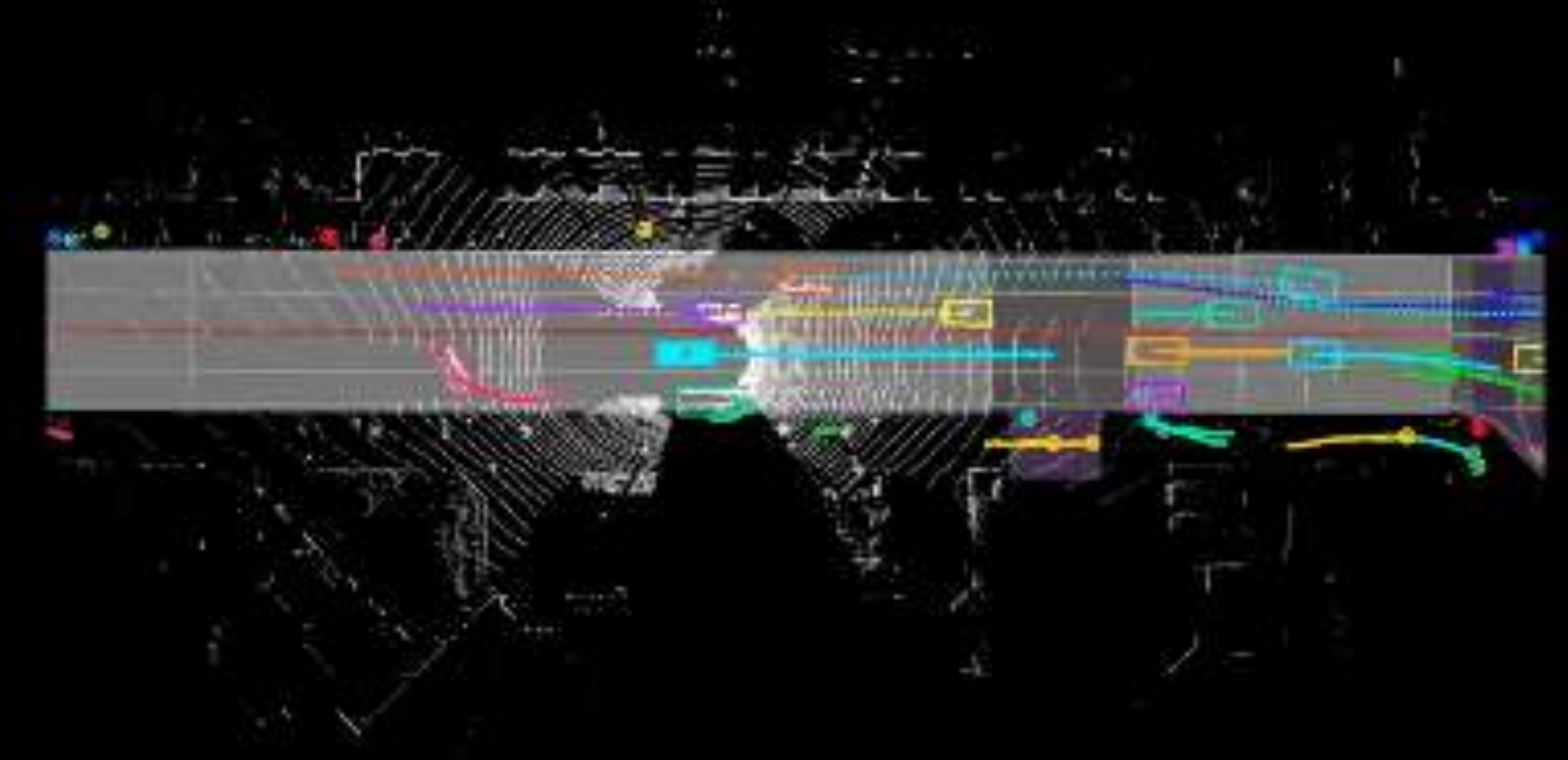


Slide credits: Raquel Urtasun & Sergio Casas

Vehicle Control

Maps → Perception → Prediction → Planning → **Control**

Sensors | Detections, Tracks | Long Term Predictions | Motion Trajectory | Steering / Acceleration



Slide credits: Raquel Urtasun & Sergio Casas

Using High-Definition Maps

- Contents
  - Precise lane boundaries and topology
  - Traffic rules, signs, right of way
  - Crosswalks, intersections, traffic lights
- Applications
  - Improve motion forecasting
  - Robust to occlusions
  - Maps = additional sensor

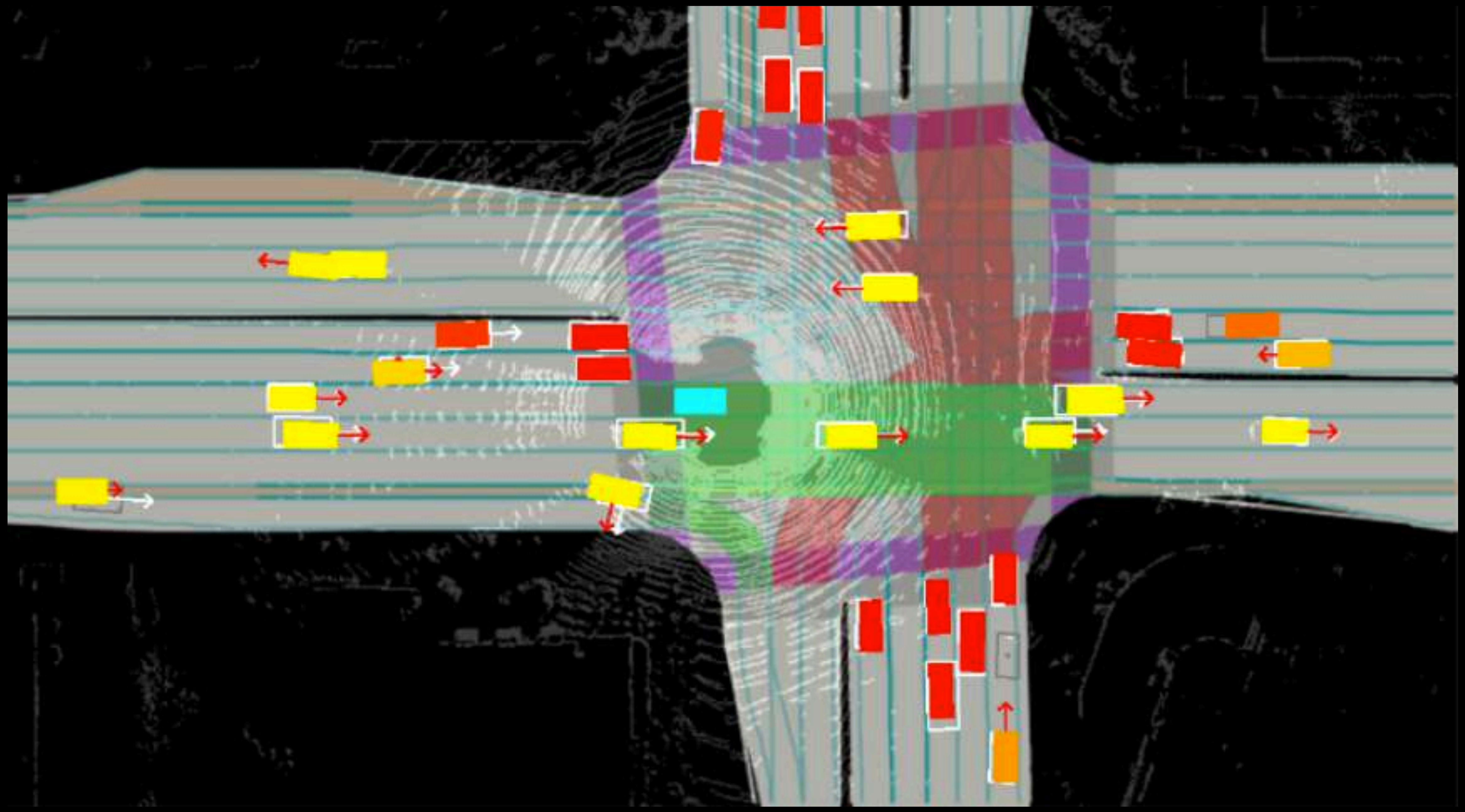
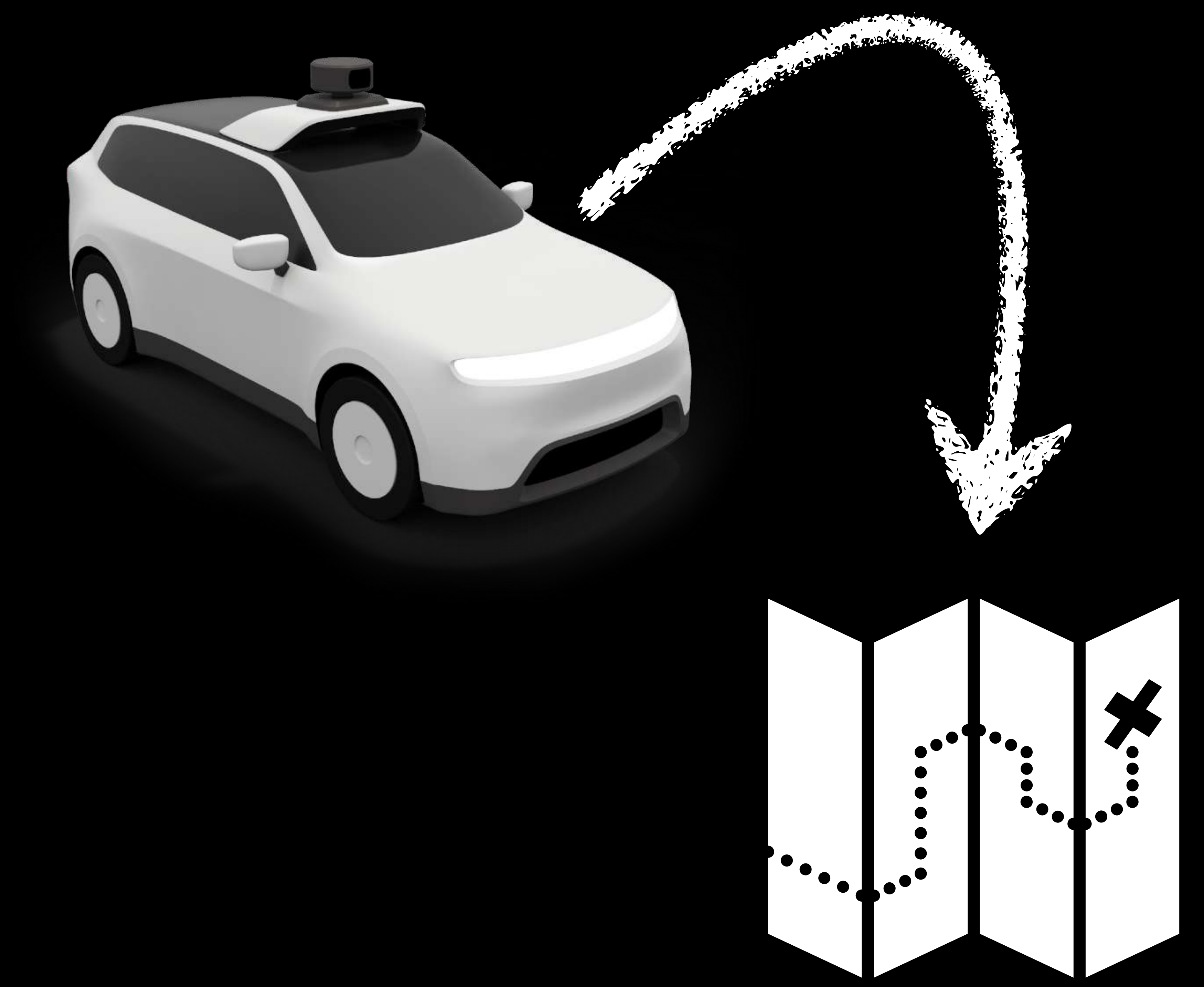


Image credit: IntentNet (Casas et al., 2018)

Why Localize?

- **HD Maps** can improve safety and performance of perception, prediction, and planning.
- Precise ego-localization is required for using maps.

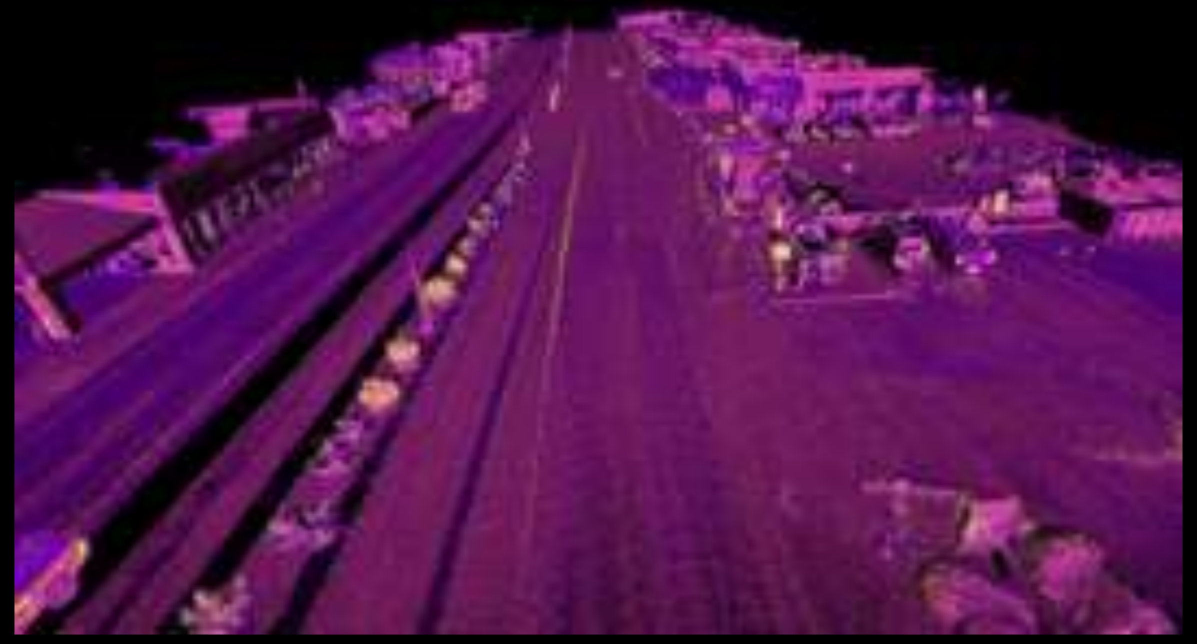


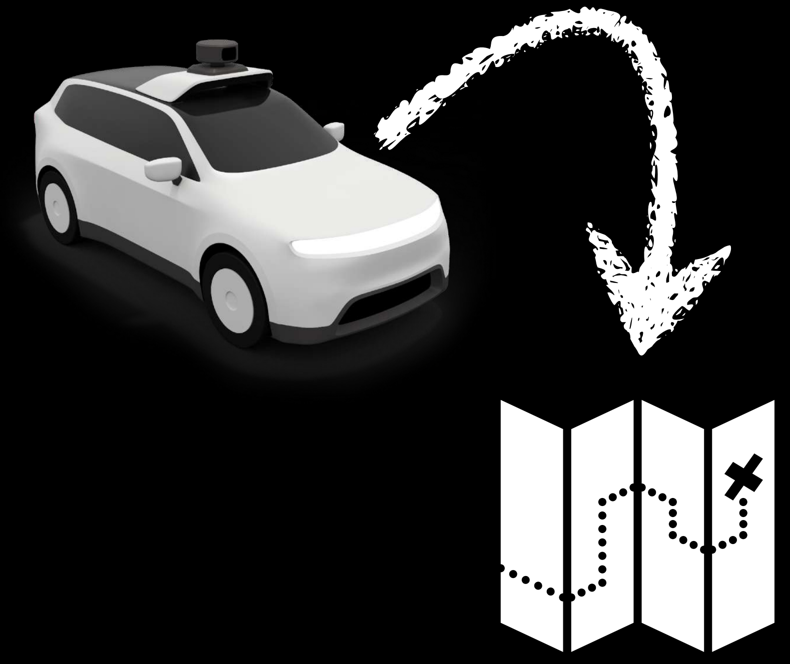
Car graphic source: Uber ATG

Problem Statement

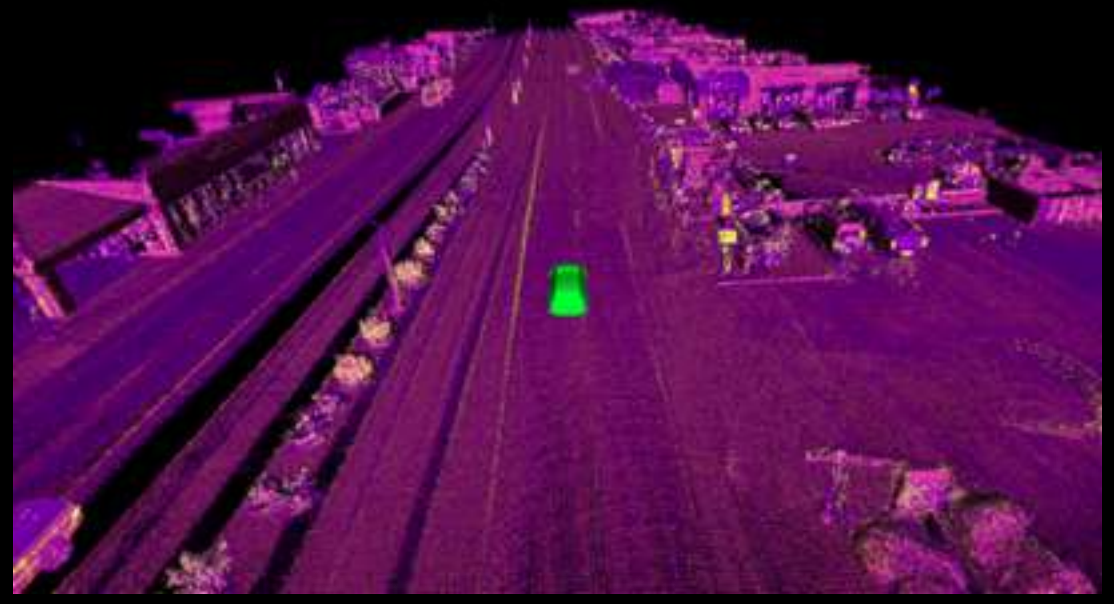
Sensor Observations

Offline Map

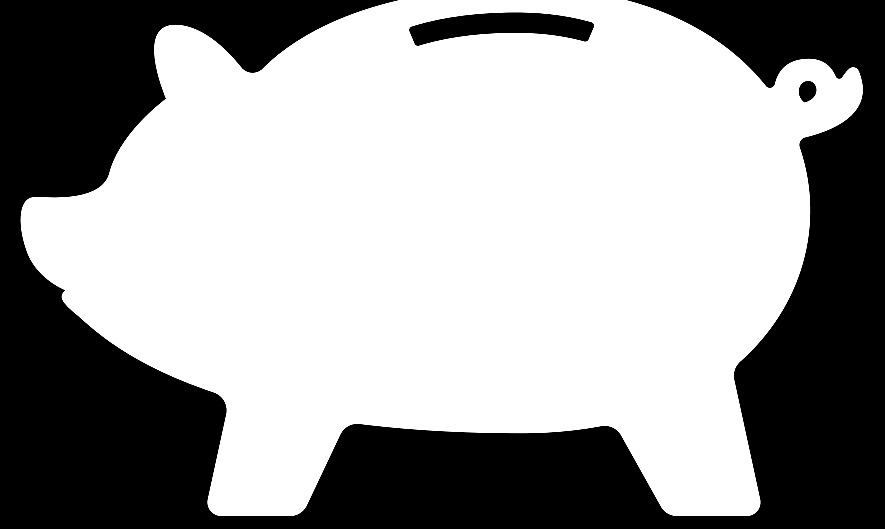




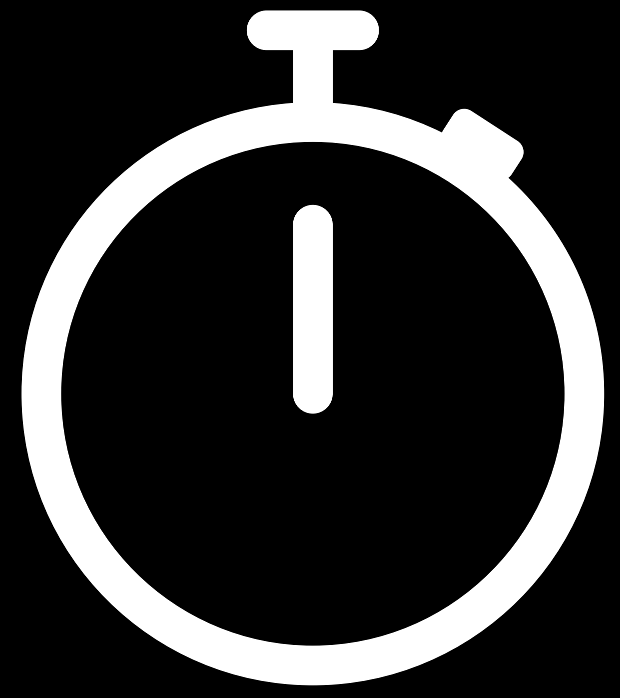
Robot **Location** in Map



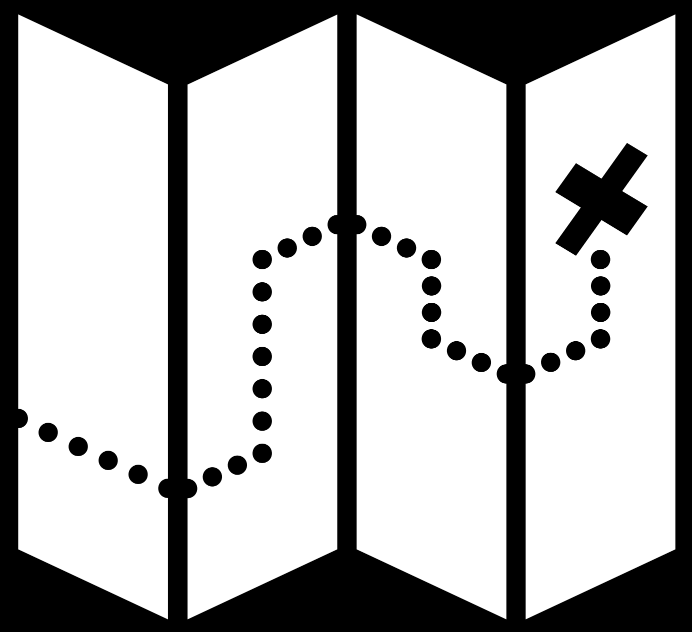
Localizer Desiderata



Low **Cost** for Map
Building & Storage



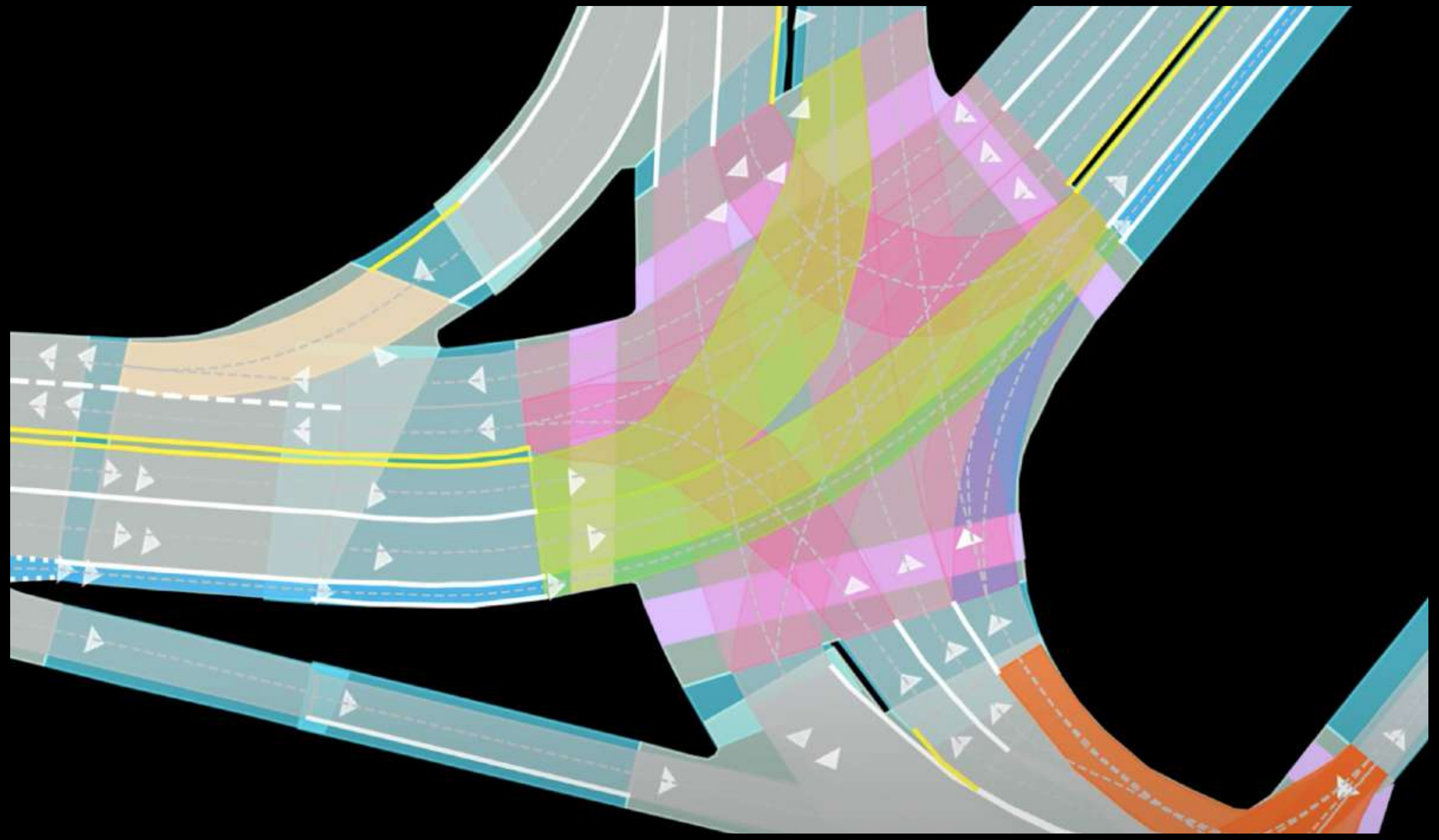
Real-Time
Inference

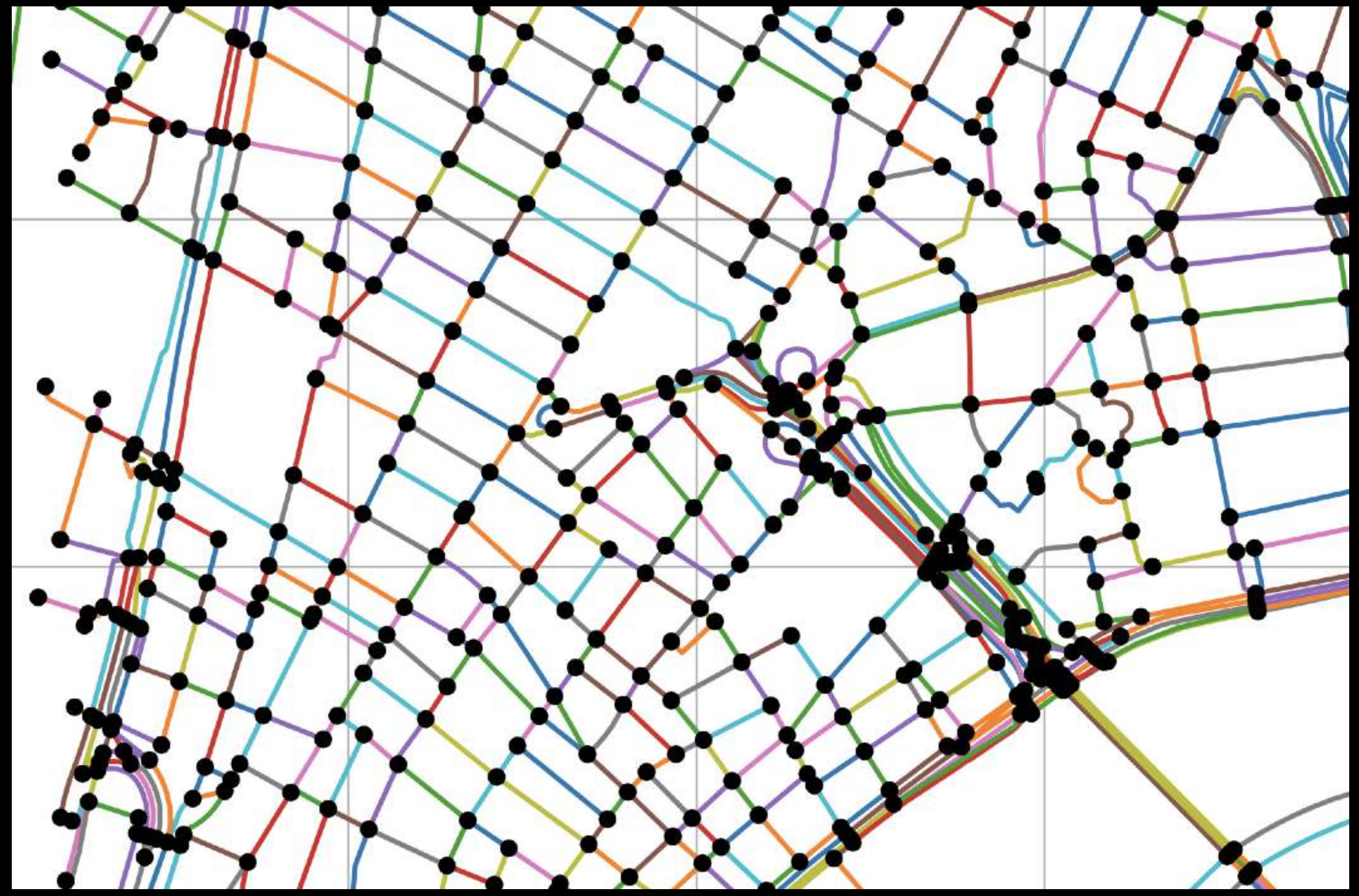


High **Accuracy**
(Centimeter-level)

Types of Localization Information

Semantic Map
(The HD map itself)





Toplogical Map

3d Geometry Map

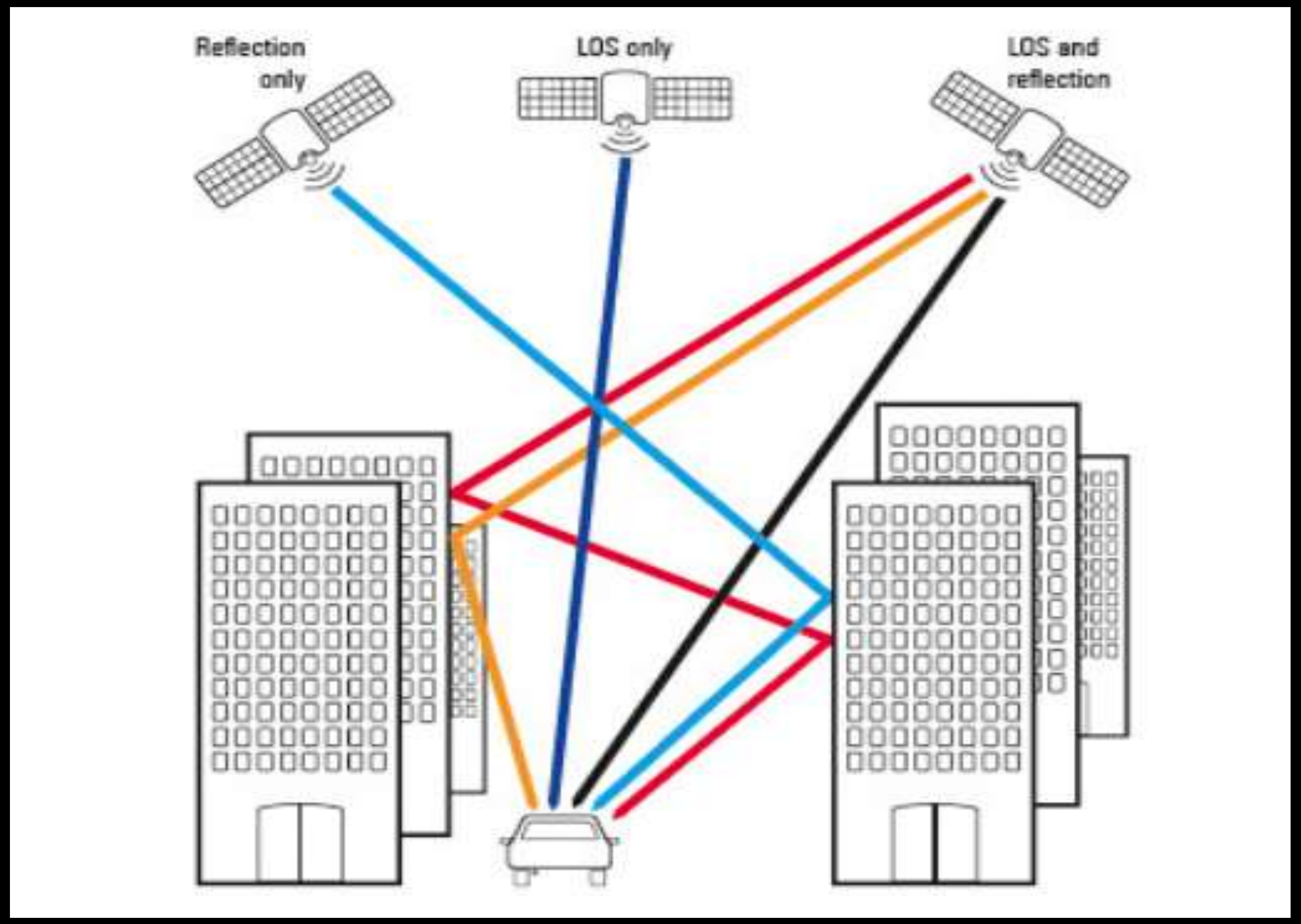
Occupancy Map

Image credits: Uber ATG

Localization Challenges

Dynamic objects

Image credit: Rohde & Schwarz



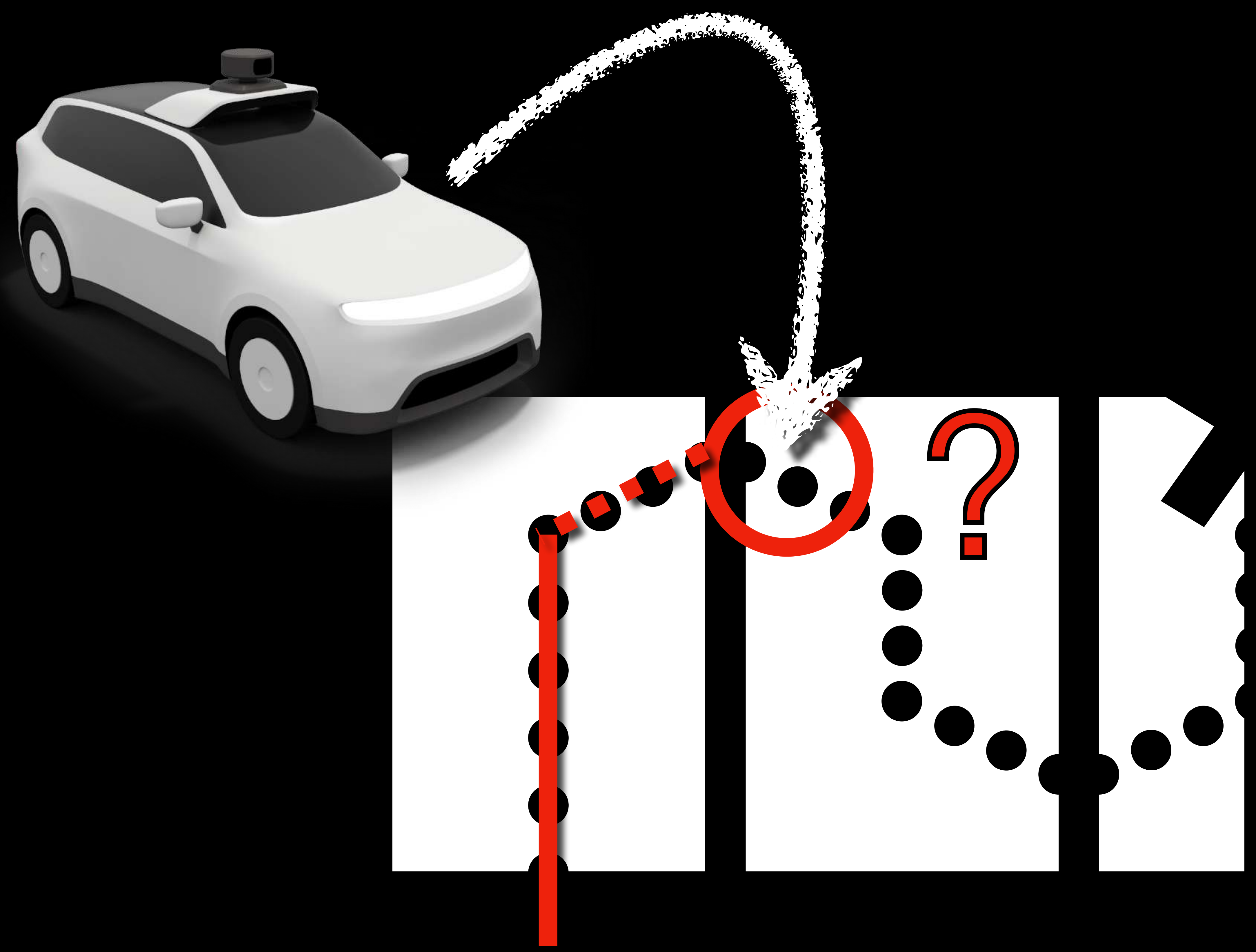
Sensor Noise

Degenerate geometry
(no useful cues)

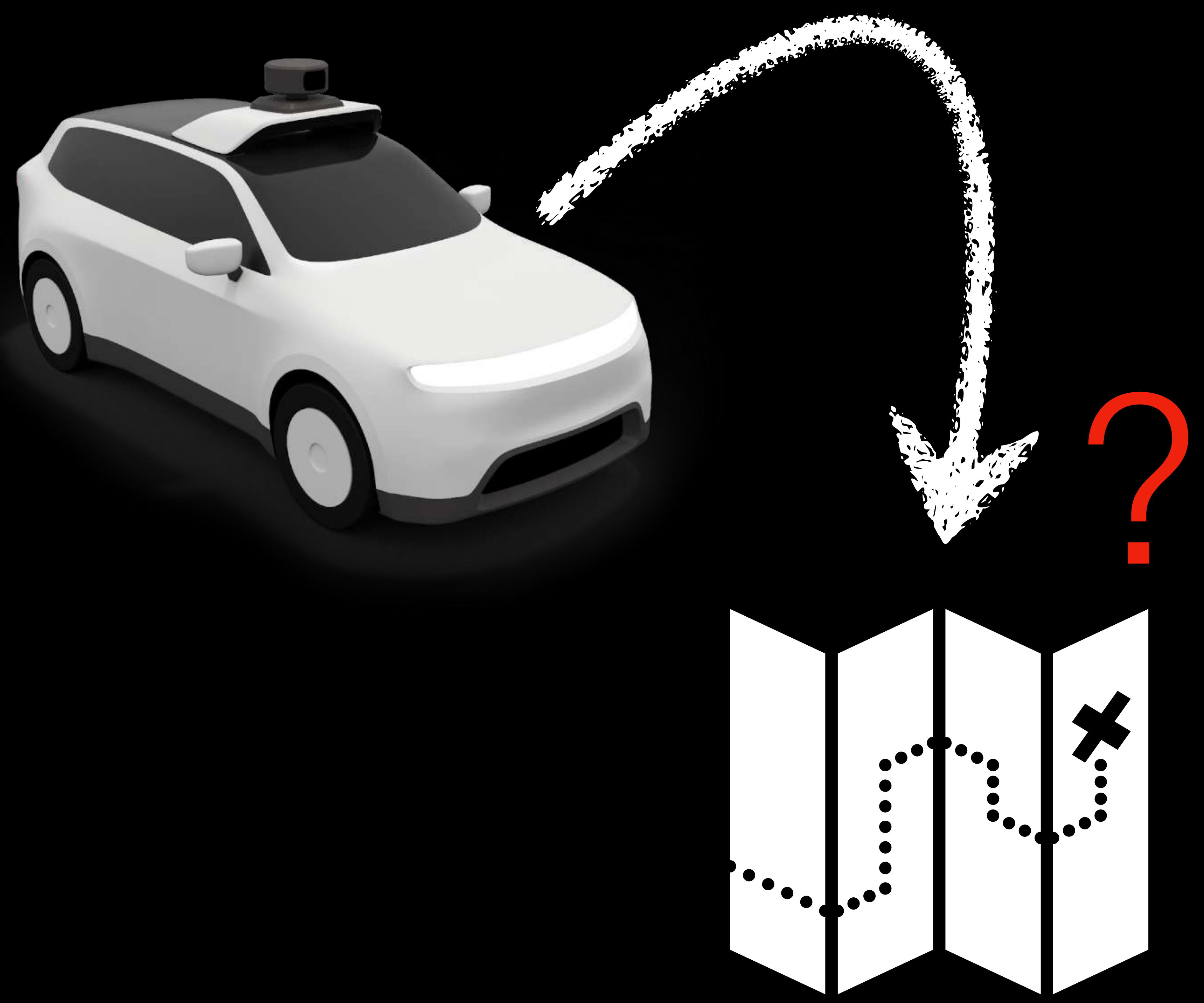
Environment Changes

Types of Localization

Online Localization

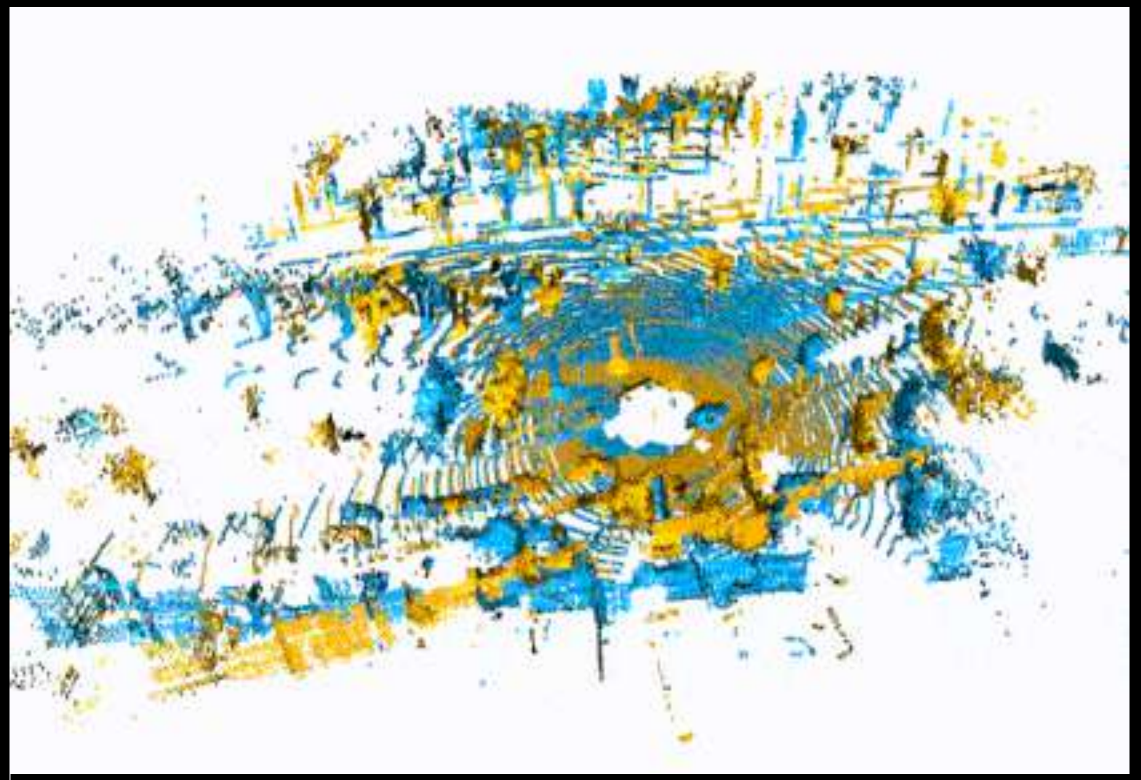


Global Localization

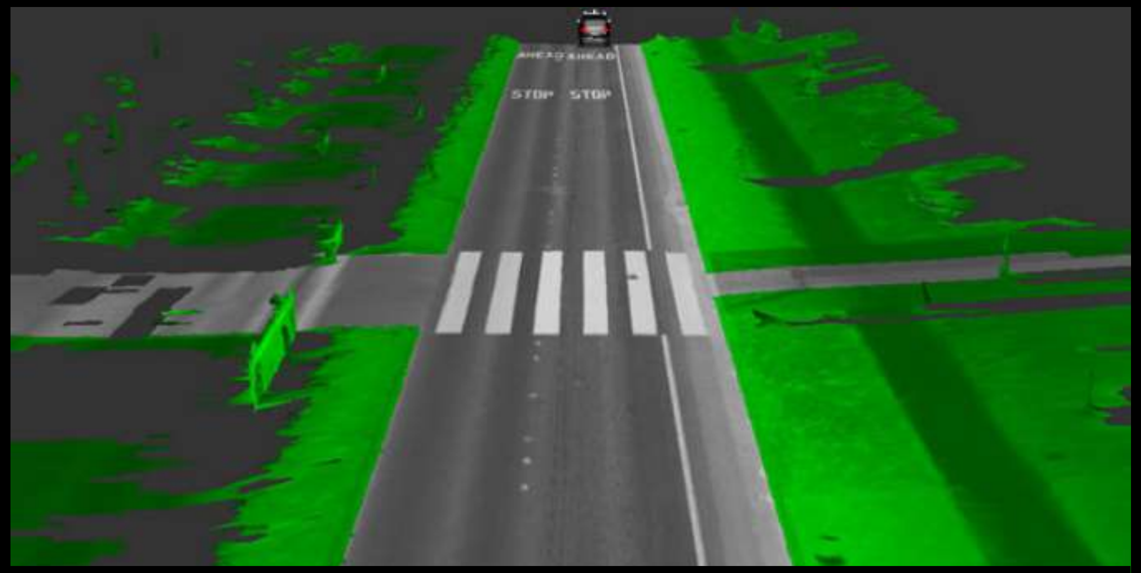


Existing Approaches

Online Localization

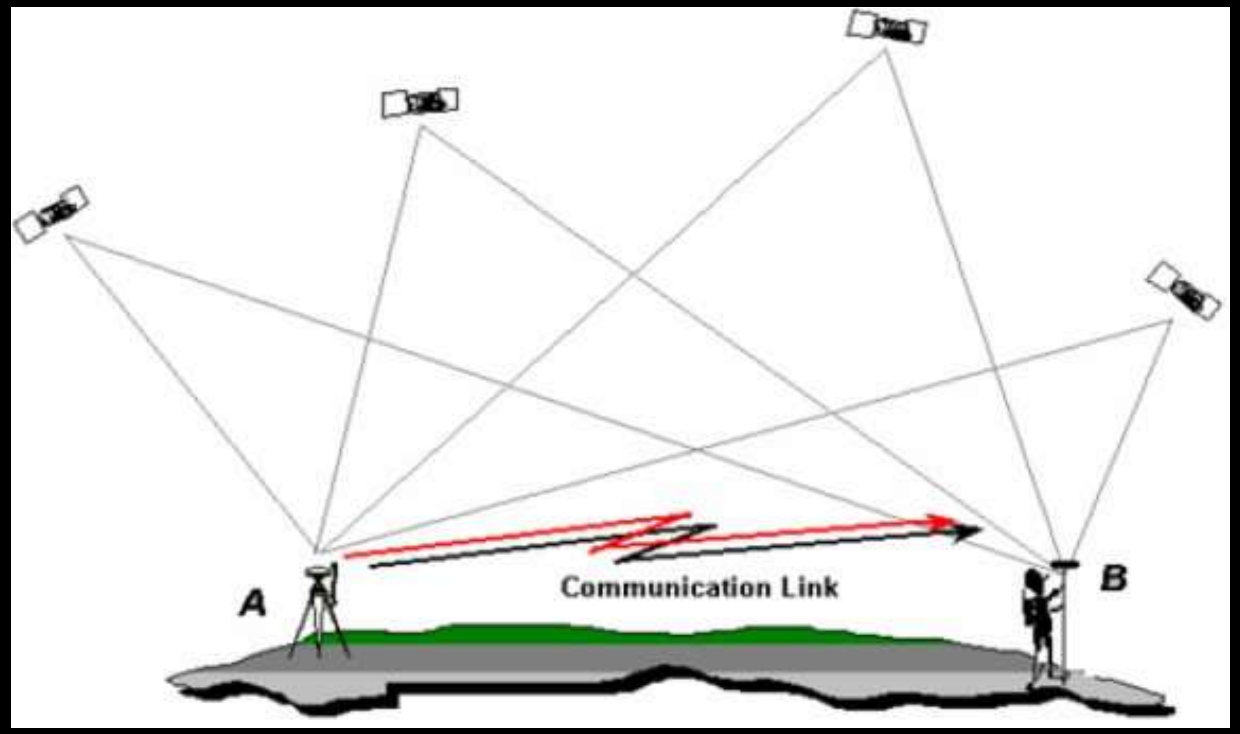


Geometric Alignment

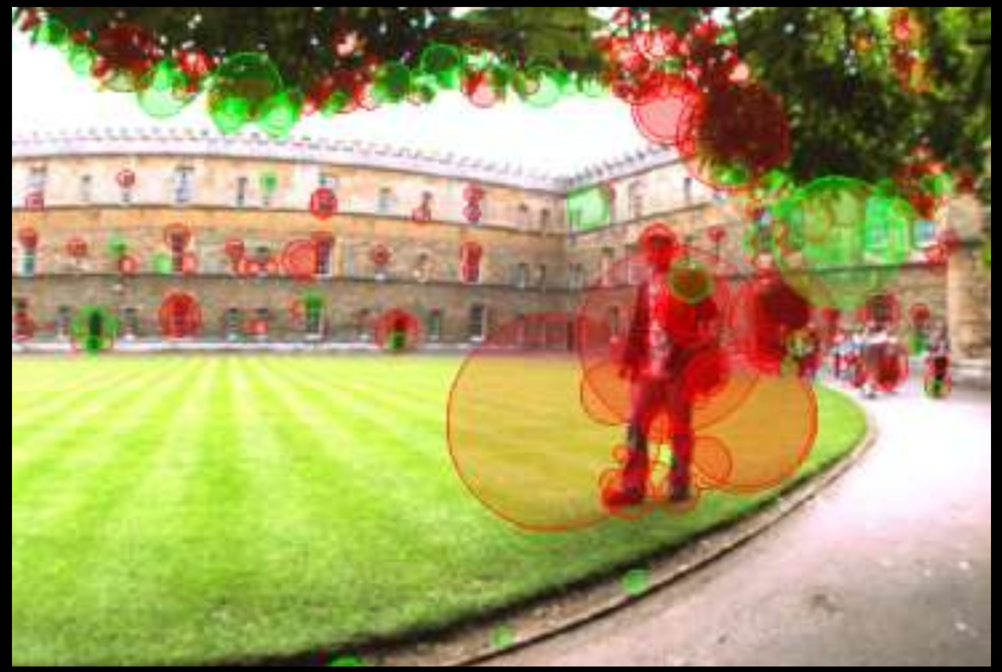


LiDAR Reflectance Matching

Global Localization



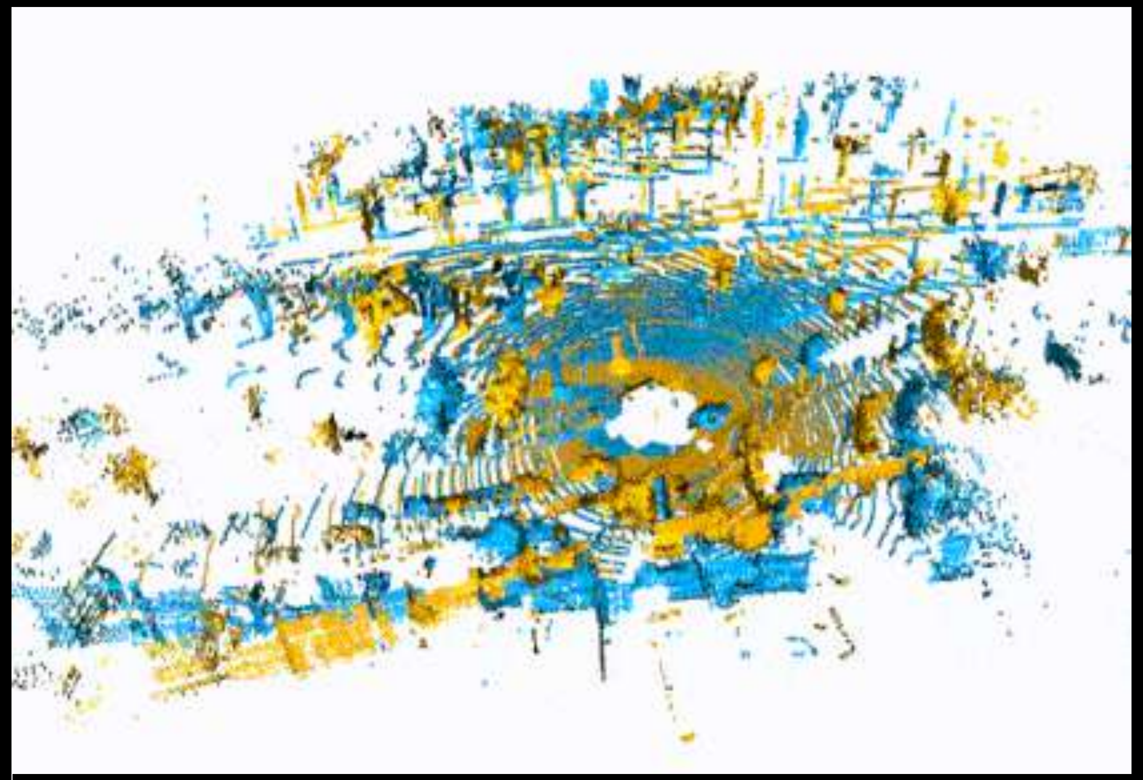
GPS / RTK



Place Recognition

Existing Approaches

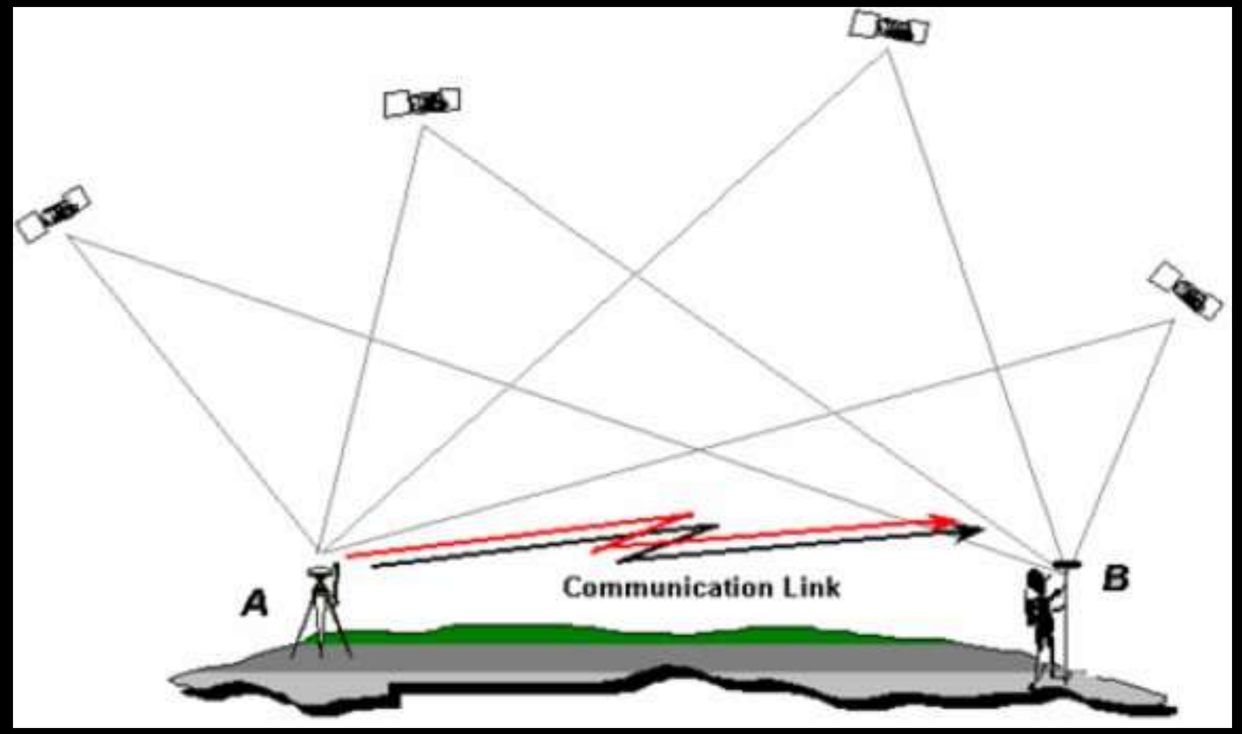
Online Localization



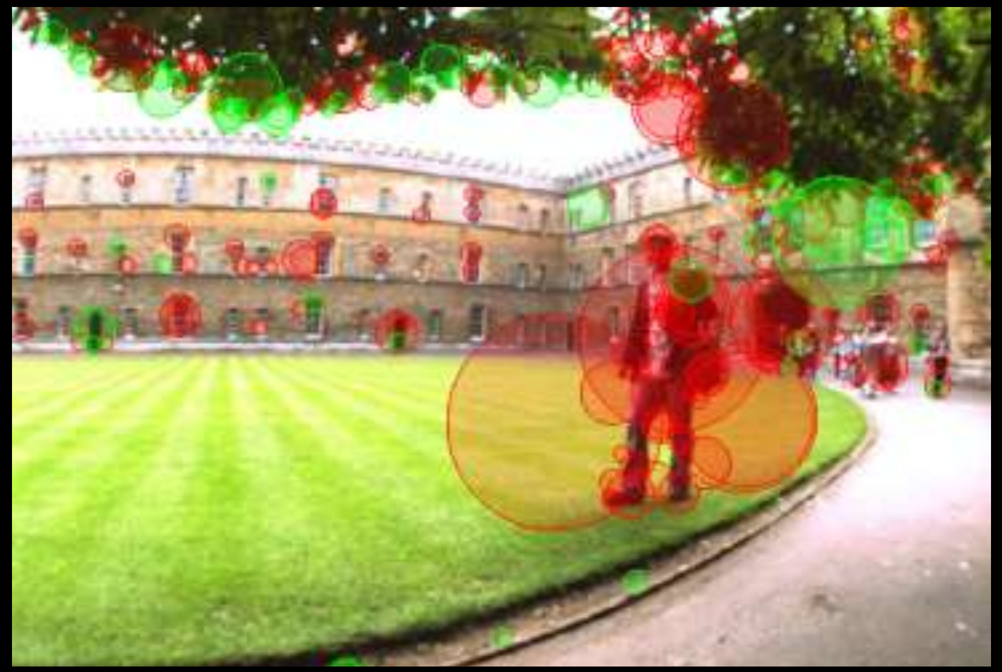
Geometric Alignment

LiDAR Reflectance Matching

Global Localization



GPS / RTK



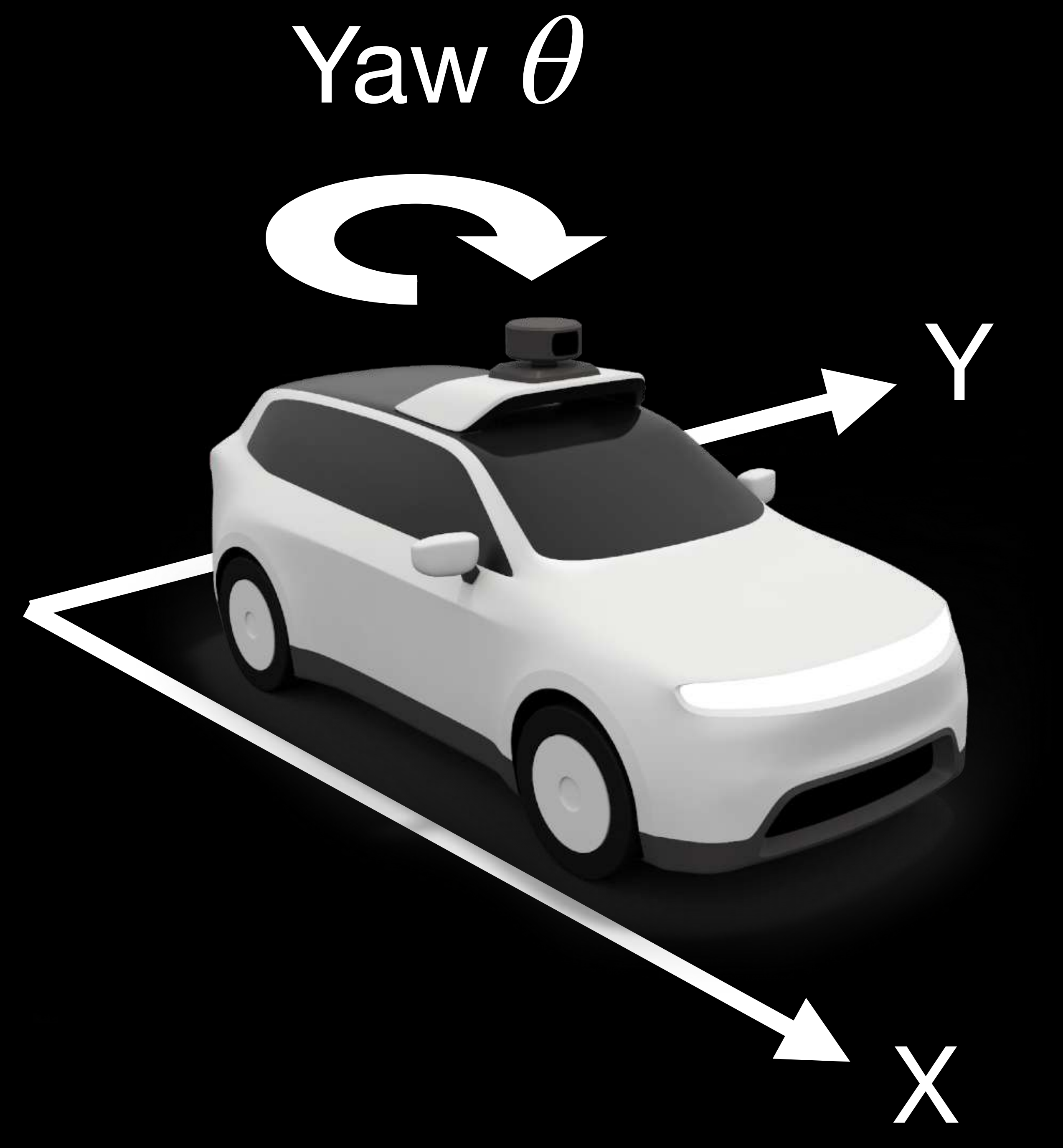
Place Recognition

Scalable Map-Based Localization

Based on joint work with Xinkai Wei, Julieta Martinez, Andrei Pokrovsky, Raquel Urtasun, and Shenlong Wang
See references (CoRL '18, CVPR '19)

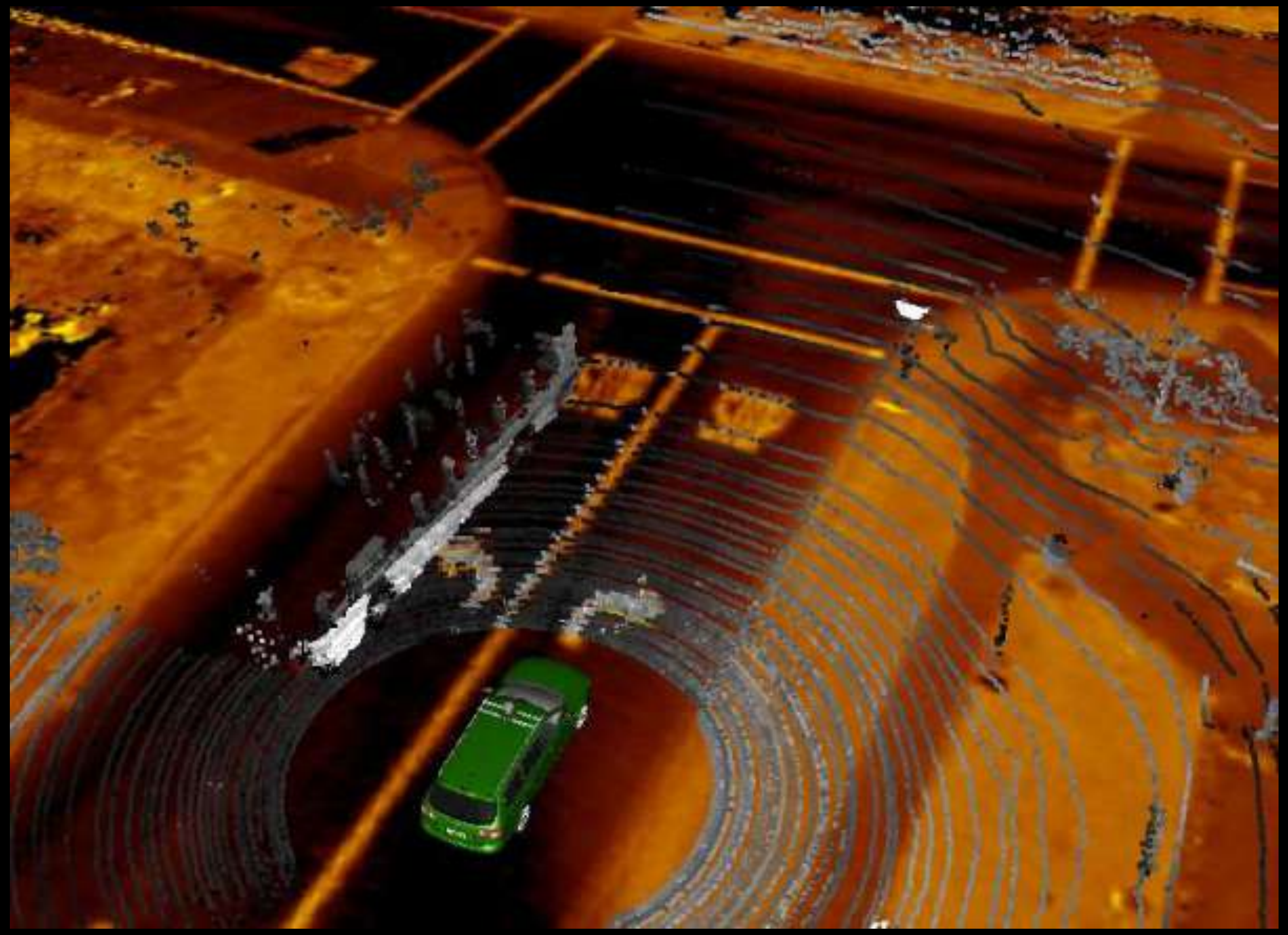
Map-Based LiDAR Localization

- Focus: **Online** localization
- Leverage dense HD maps built in advance
- Use LiDAR
- Vehicle on ground:
  - **Minimal pose:** (X, Y, yaw)
  - **Easy** and **efficient** to model



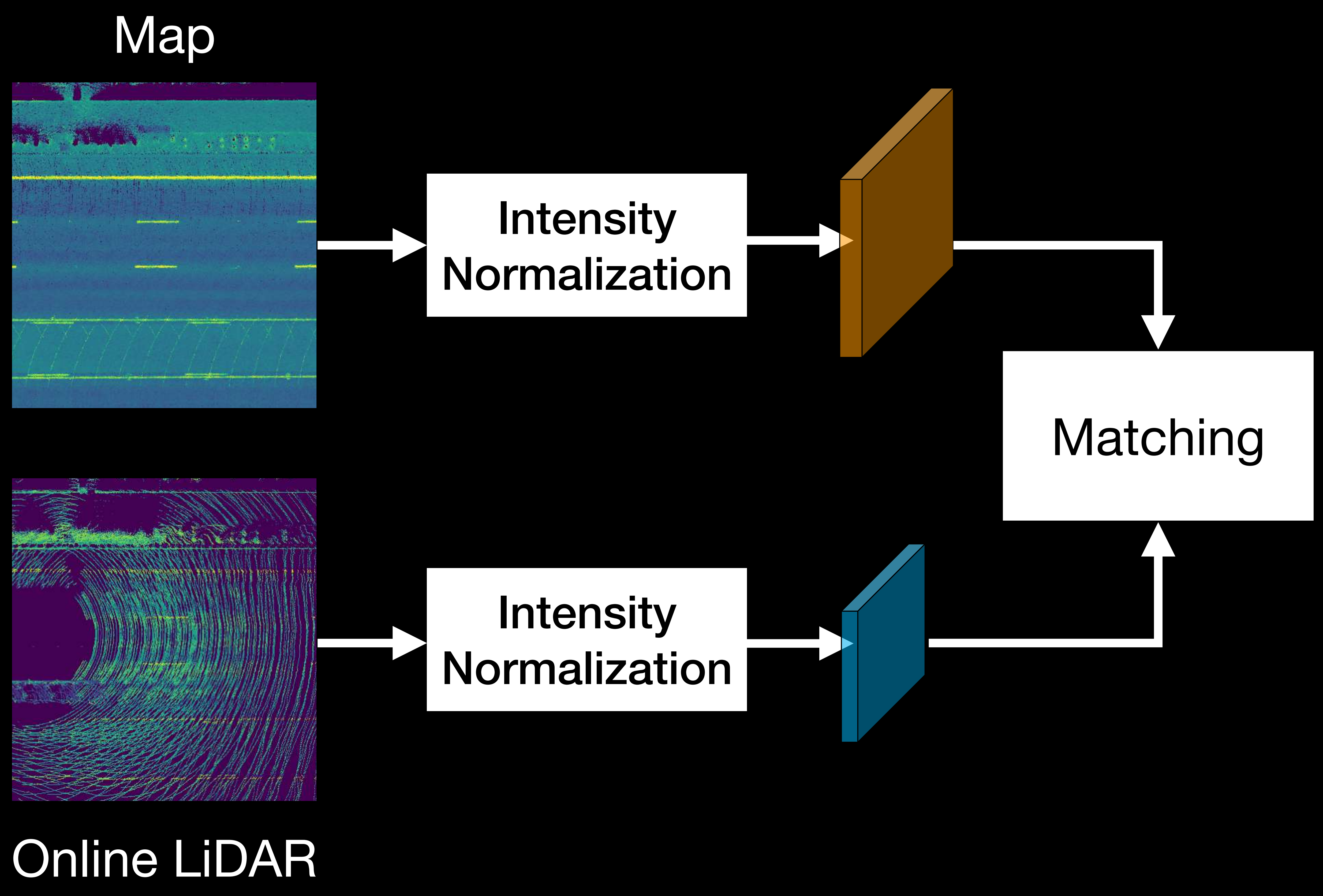
Background: LiDAR Reflectance Matching

- Correlate observations to the map
- Strengths
  - Robust to outliers and nearly featureless environments
  - Can be implemented in a computationally efficient way
- Limitations
  - Requires good initialization (online localization, remember!)
  - Vulnerable to LiDAR mis-calibration and large occlusion
  - High map storage cost

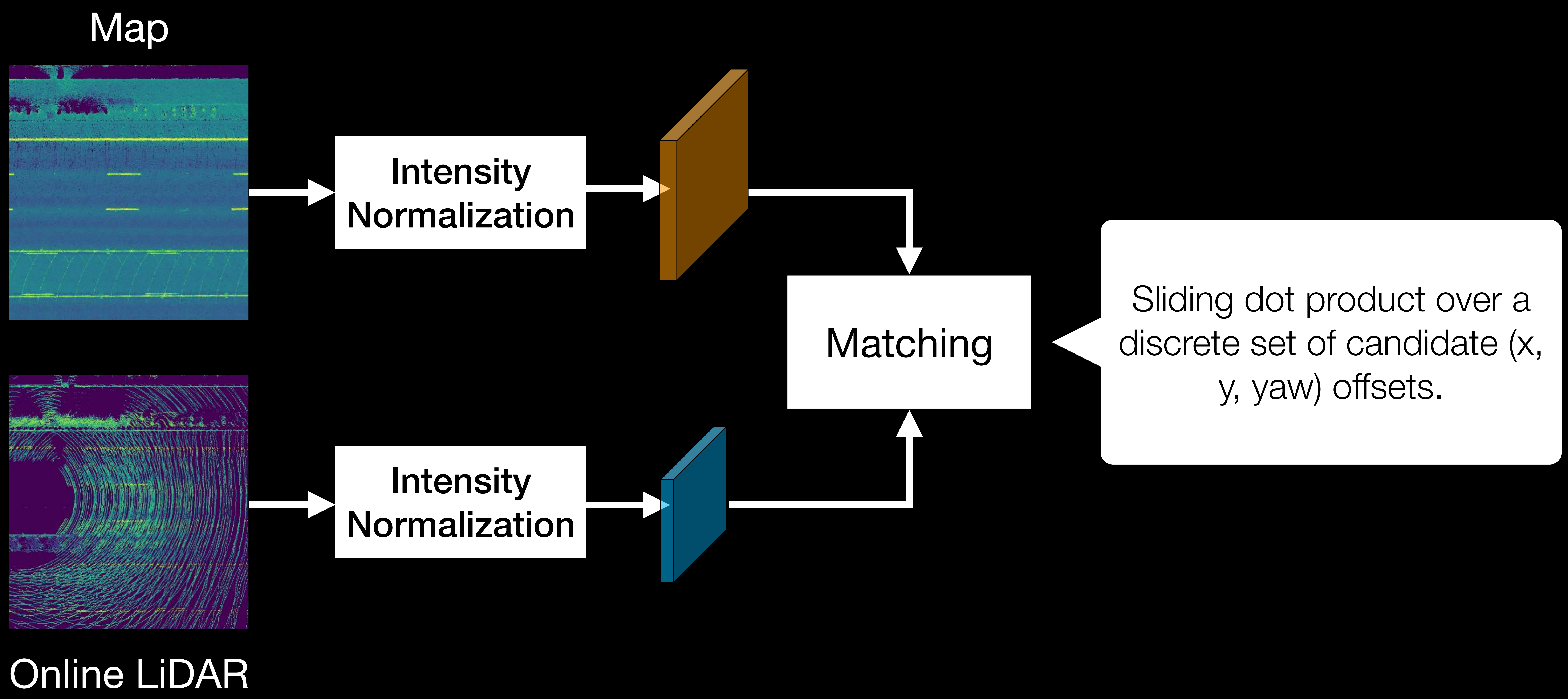


Reference: Map-Based Precision Vehicle Localization in Urban Environments [Levinson, Montemerlo, Thrun, RSS 2013]

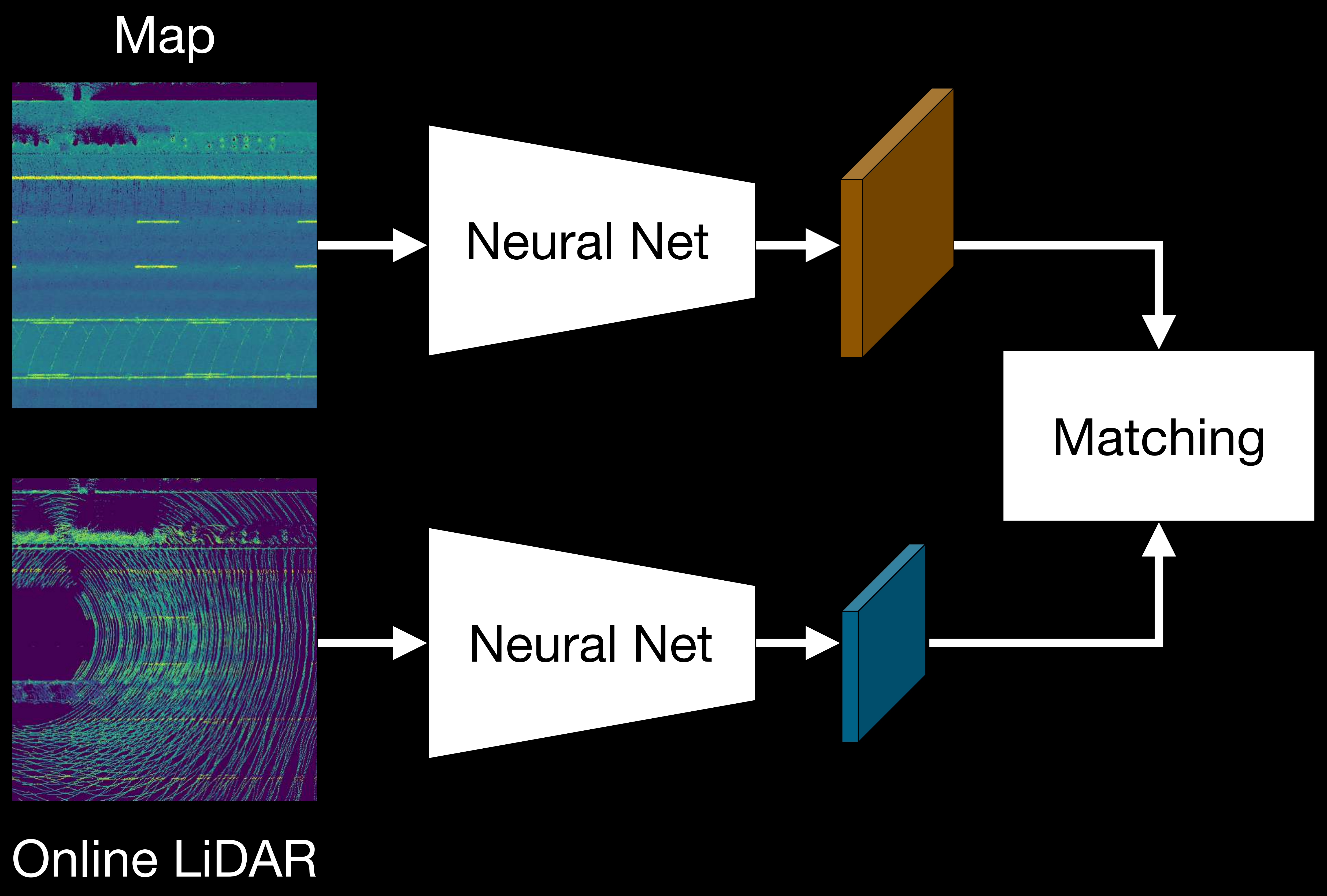
Template Matching Idea



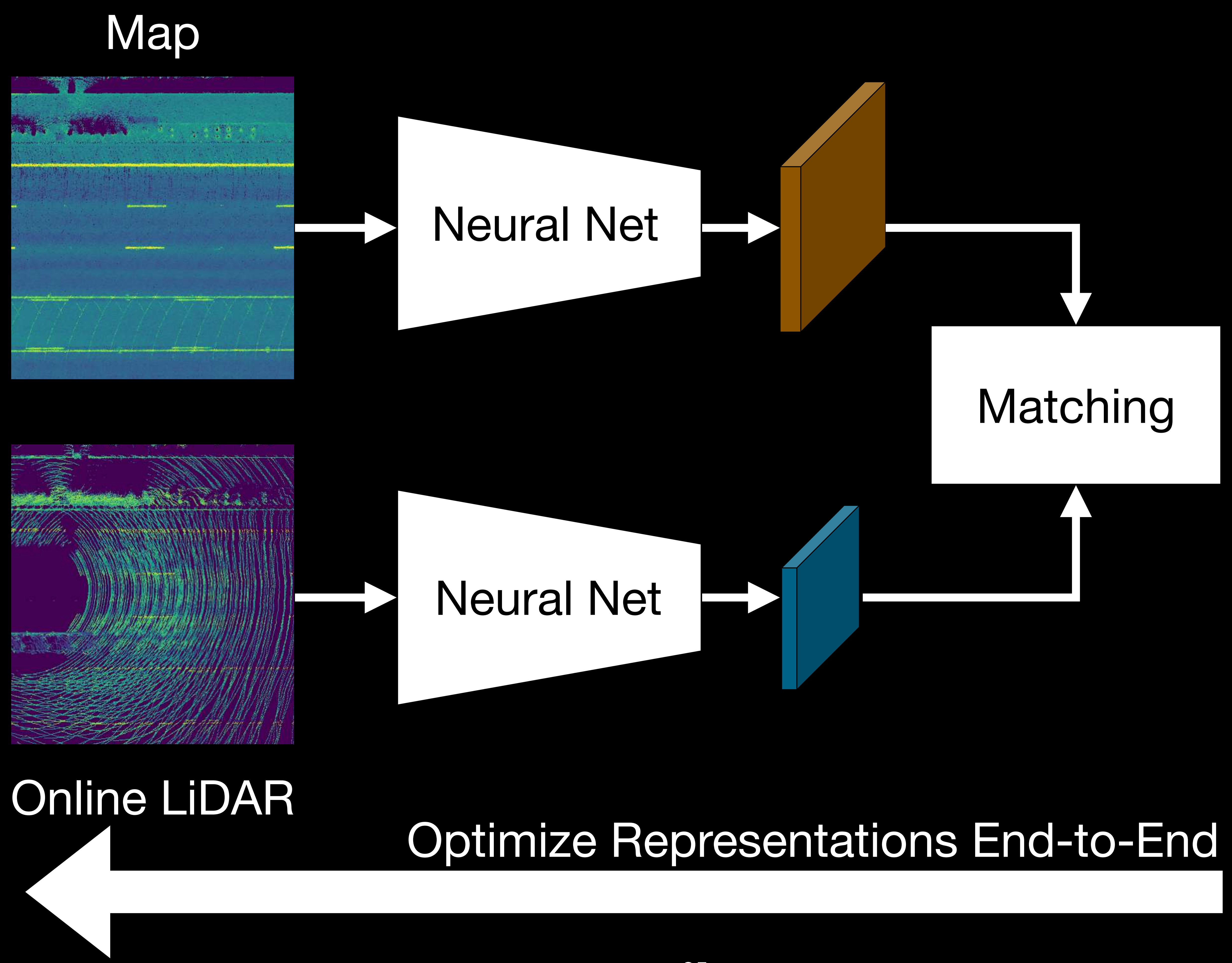
Template Matching Idea



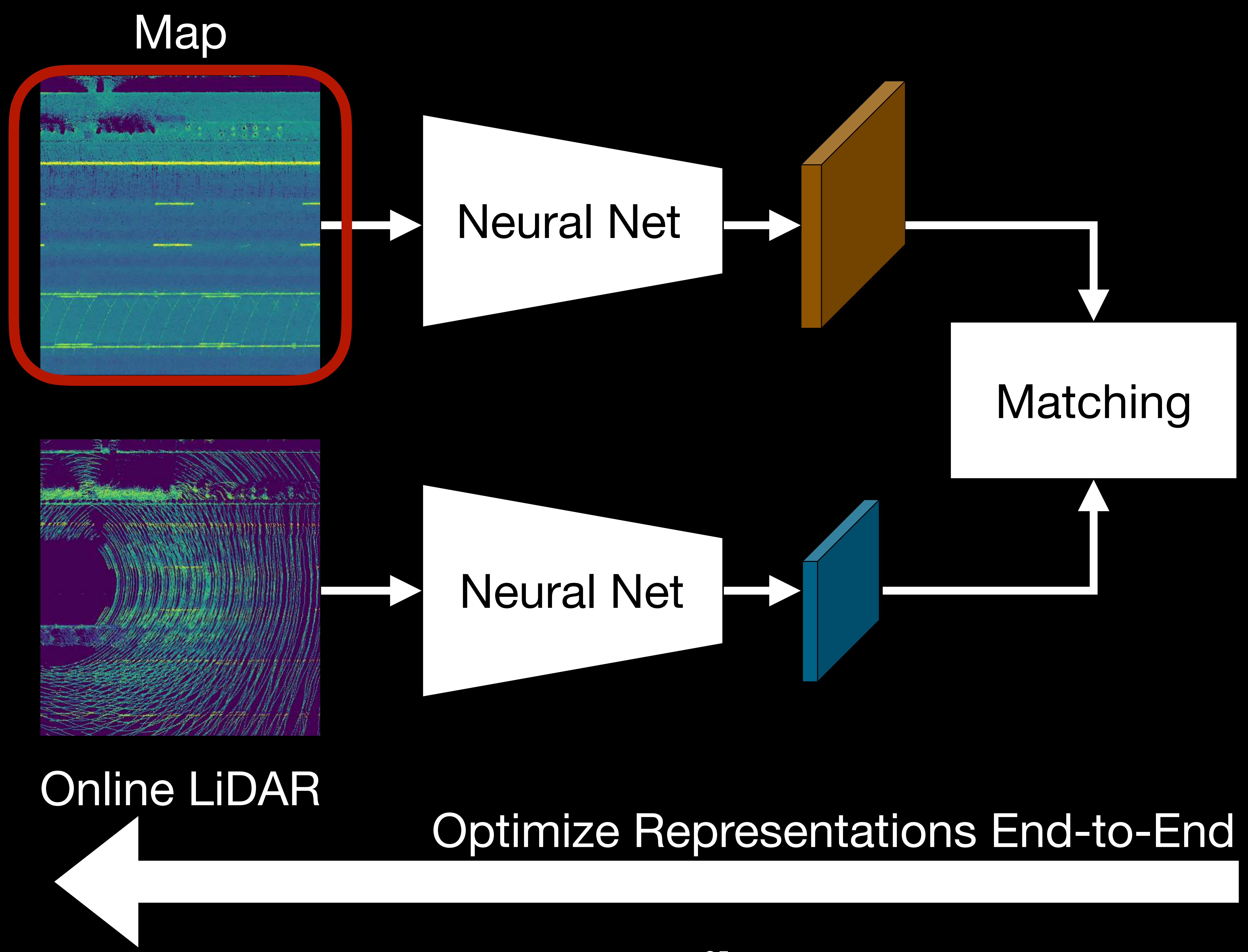
Learning to Match



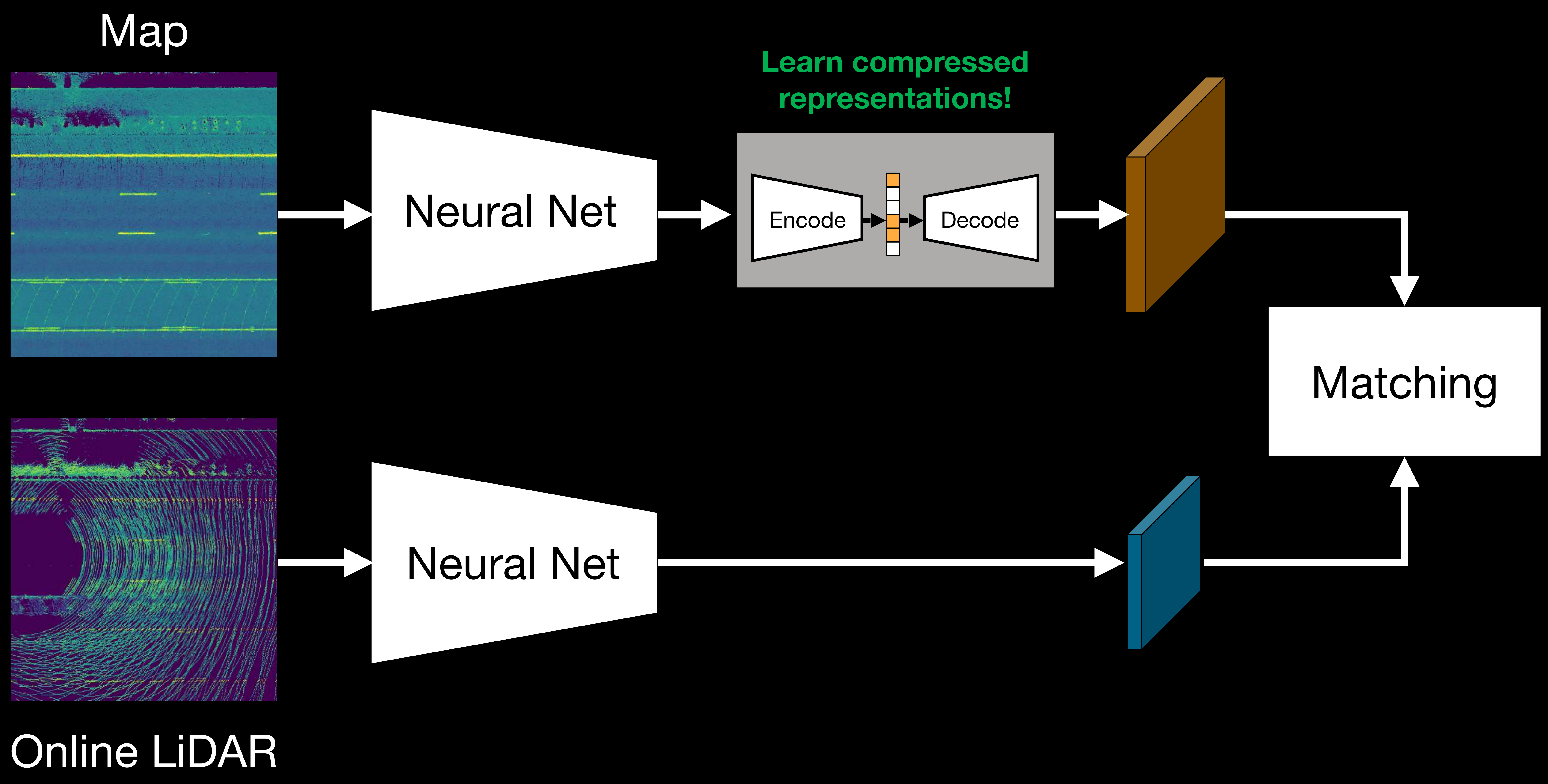
Learning to Match



Learning to Match



Learning to Compress Maps

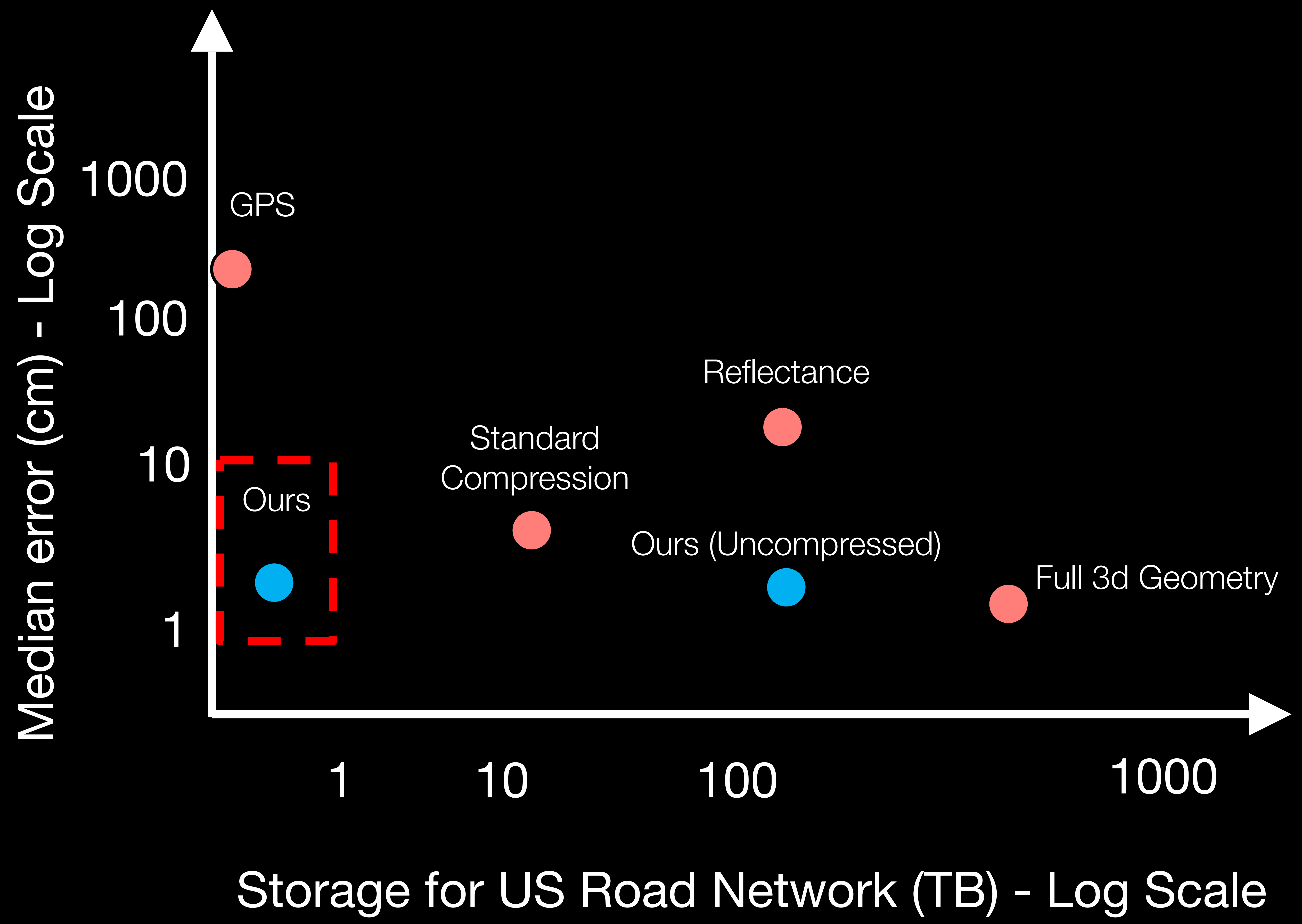


Metrics

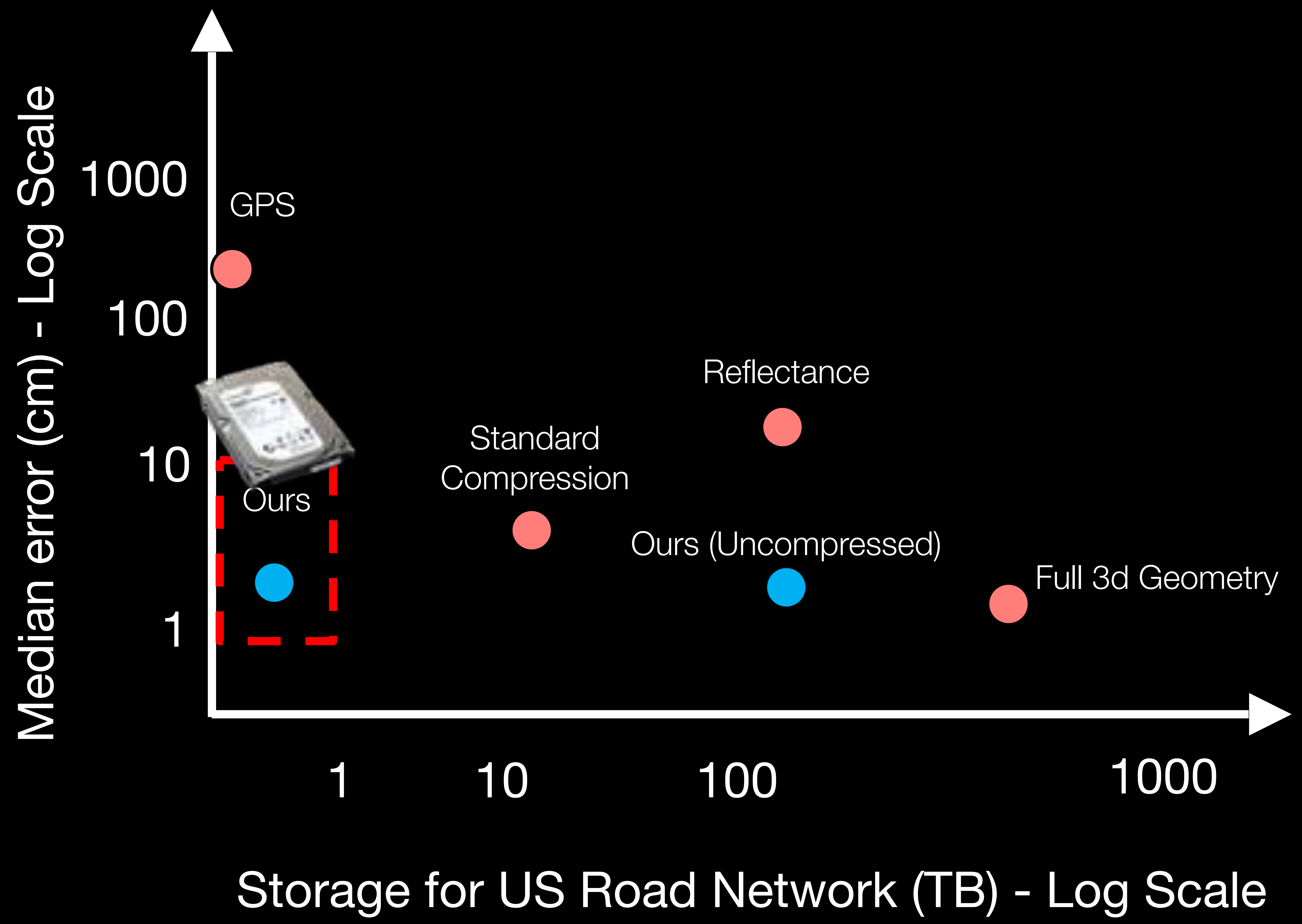
How good is my localizer?

- Localization accuracy
  - Euclidean distance between computed and ground-truth pose
- Map storage
  - Approx. size in TB to store entire US road network @ 5cm / px

Localization & Map Compression Results

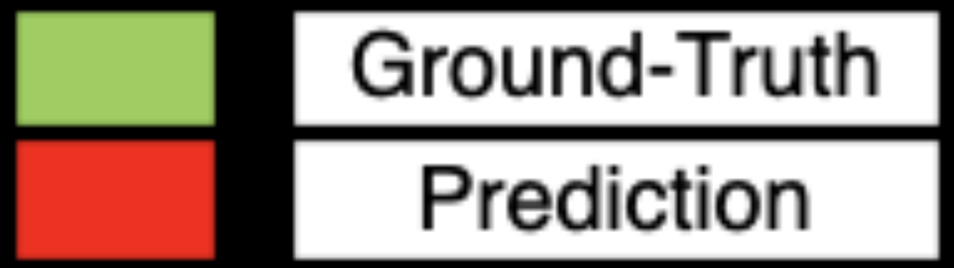


Localization & Map Compression Results



Localizer Result

Recent LiDAR Sweeps | **Dense Reflectivity Map** | **Localization Result**



Localizer Result

Recent LiDAR Sweeps | **Dense Reflectivity Map** | **Localization Result**

Ground-Truth

Prediction

Take-Home Message

- HD Maps are powerful but rely on precise localization
- LiDAR matching is effective for precise (online) localization
- Learning can dramatically improve the robustness of LiDAR matching
- When compressing data, **think!** Who or what will be using this data, and how?
    - If the data is very specialized, then it makes sense to specialize compression

How Good Does Localization Need to Be?

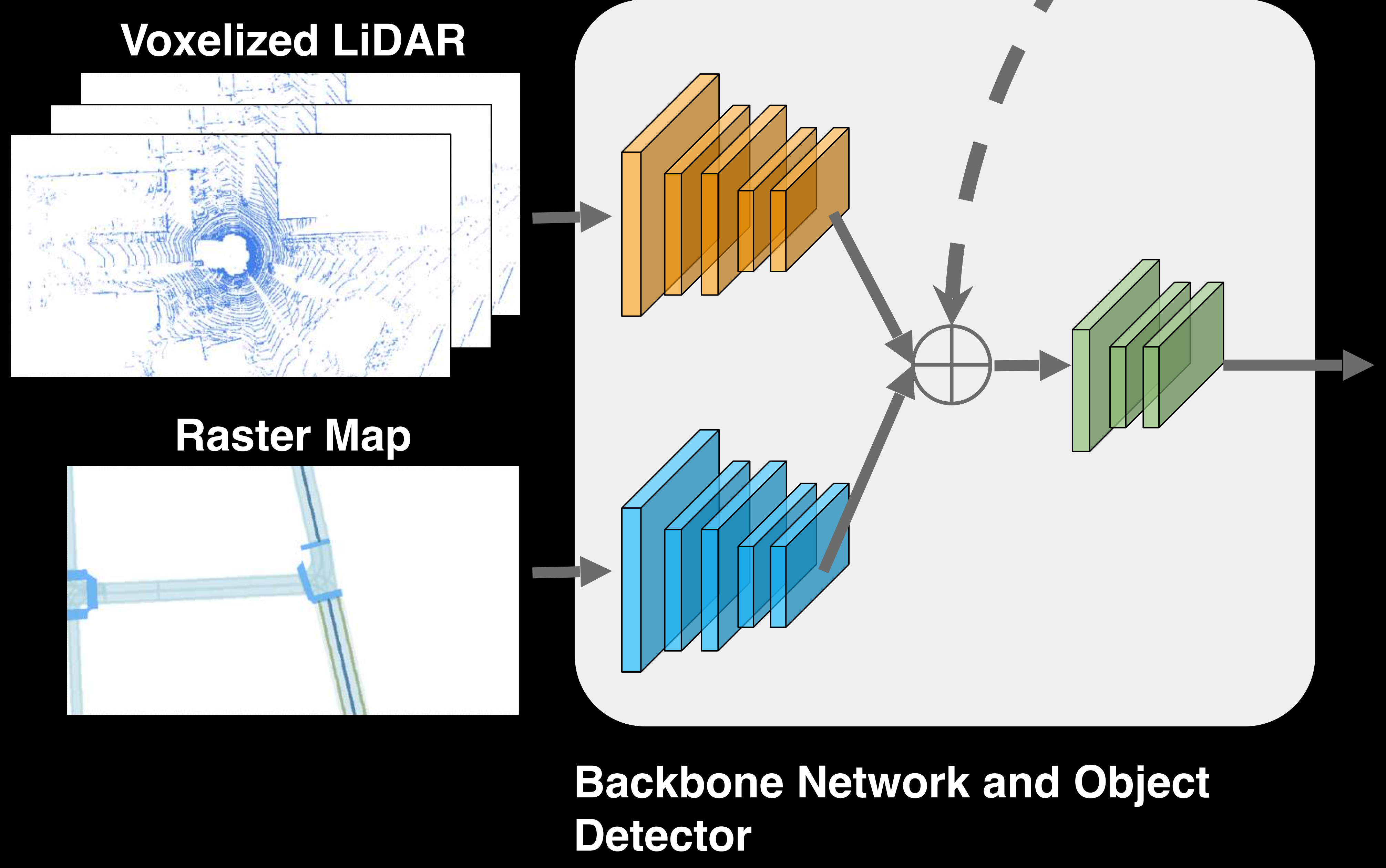
Based on joint work with John Phillips, Julieta Martinez, Sergio Casas, Abbas Sadat and Raquel Urtasun
Deep Multi-Task Learning for Joint Localization, Perception, and Prediction (CVPR 2021)

HD Map Limitations

- Expensive to build and maintain => automation & new sensing modalities
- Can go out of date => change detection, mapless driving, live updates
- Reliant on precise localization => **how much?**

Input Fusion

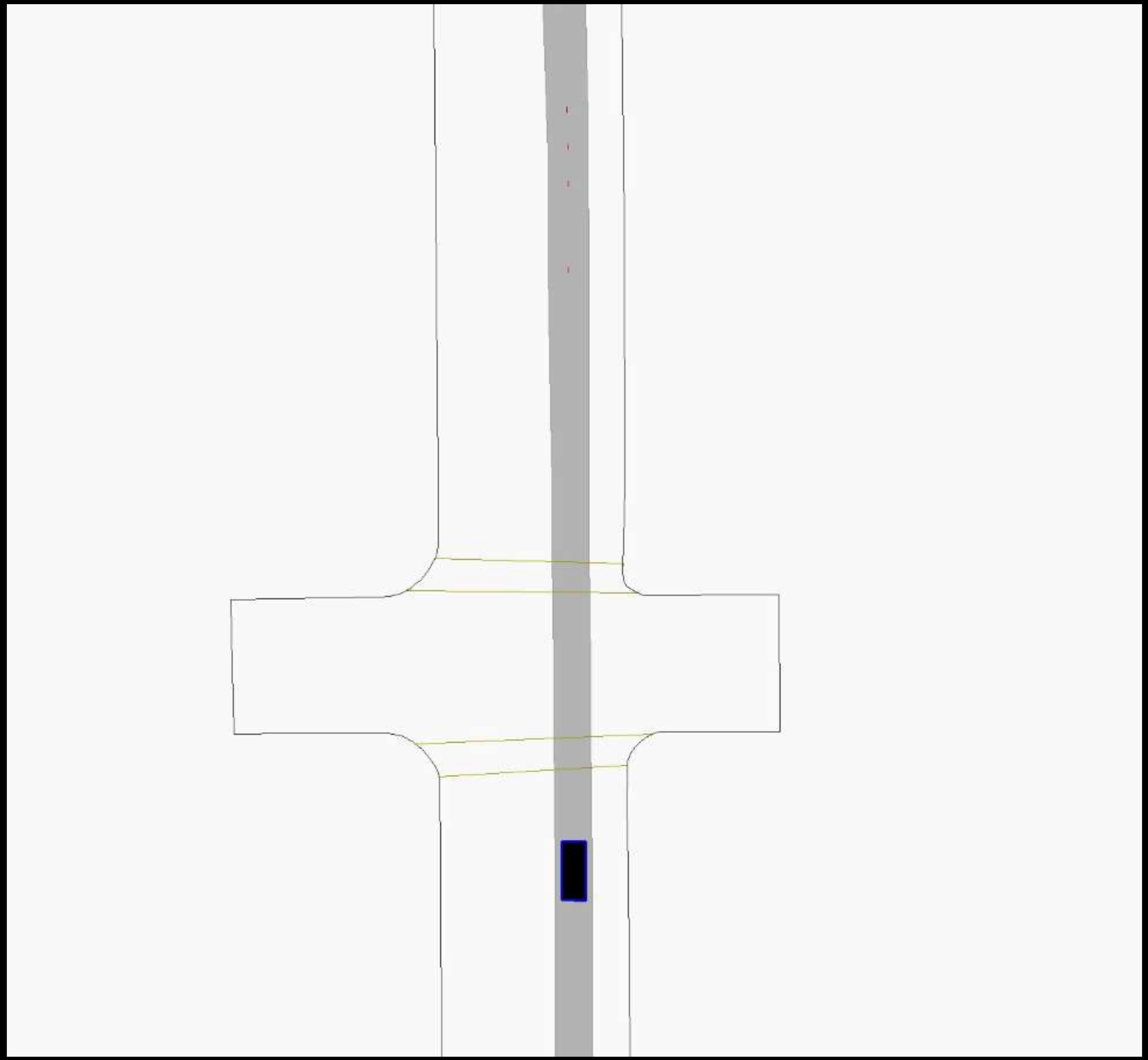
⚠️ Element-wise addition assumes good LiDAR-map alignment!



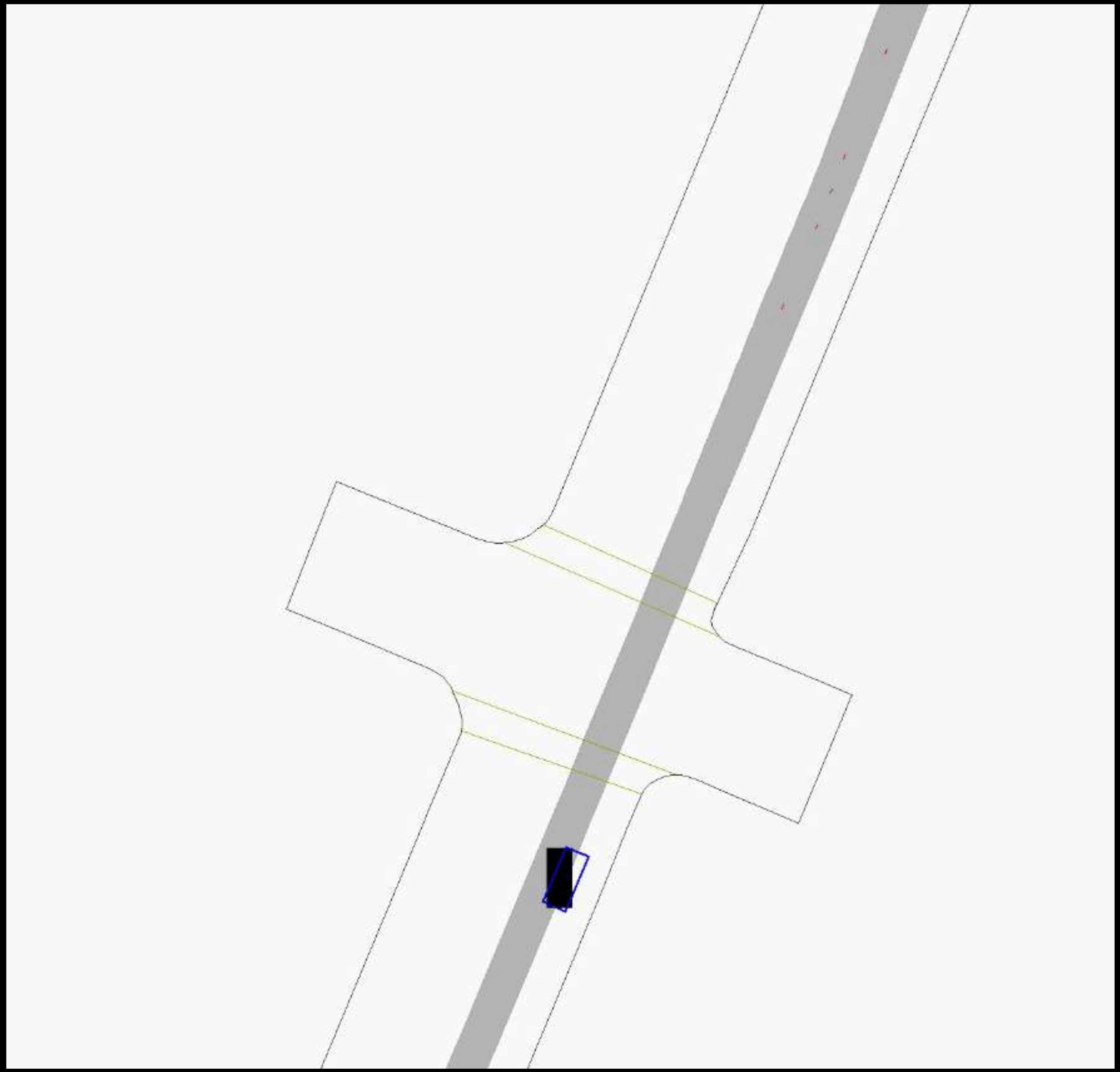
Detection,
Forecasting,
Motion Planning,
etc.

Diagram source: Implicit Latent Variable Model for Scene-Consistent Motion Forecasting [Casas et al., ECCV 2020]

What Is the Impact of Localization Errors?



Correct Pose

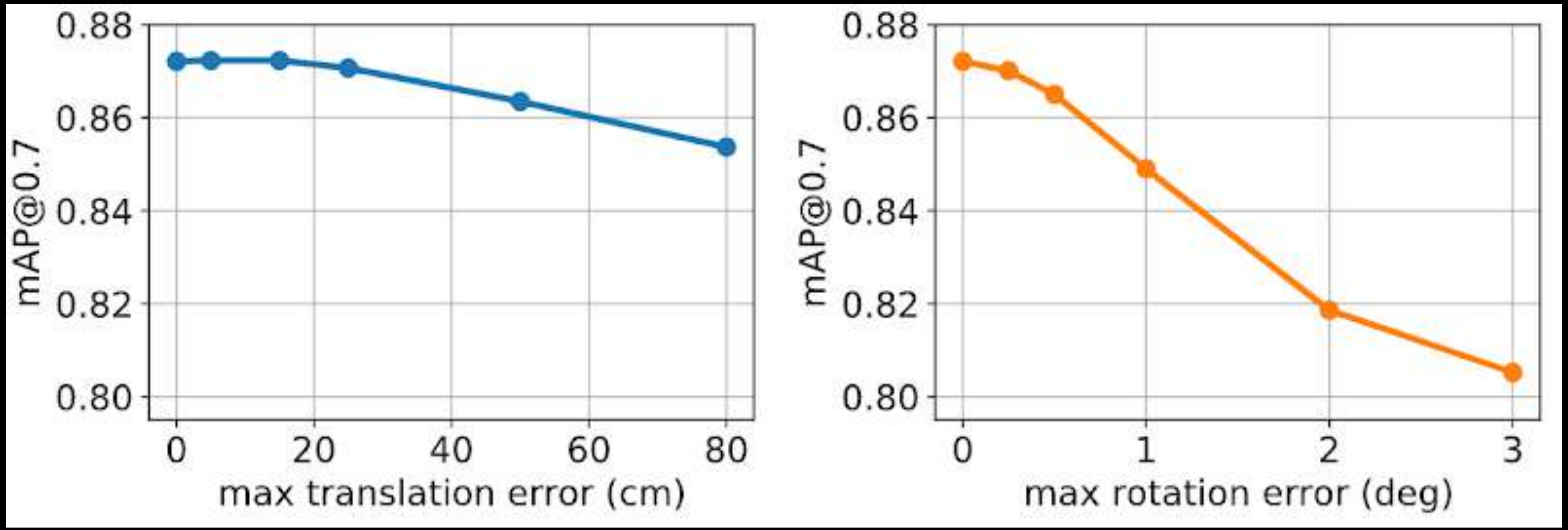


Localization Failure

The Effects of Localization Error

The Effects of Localization Error

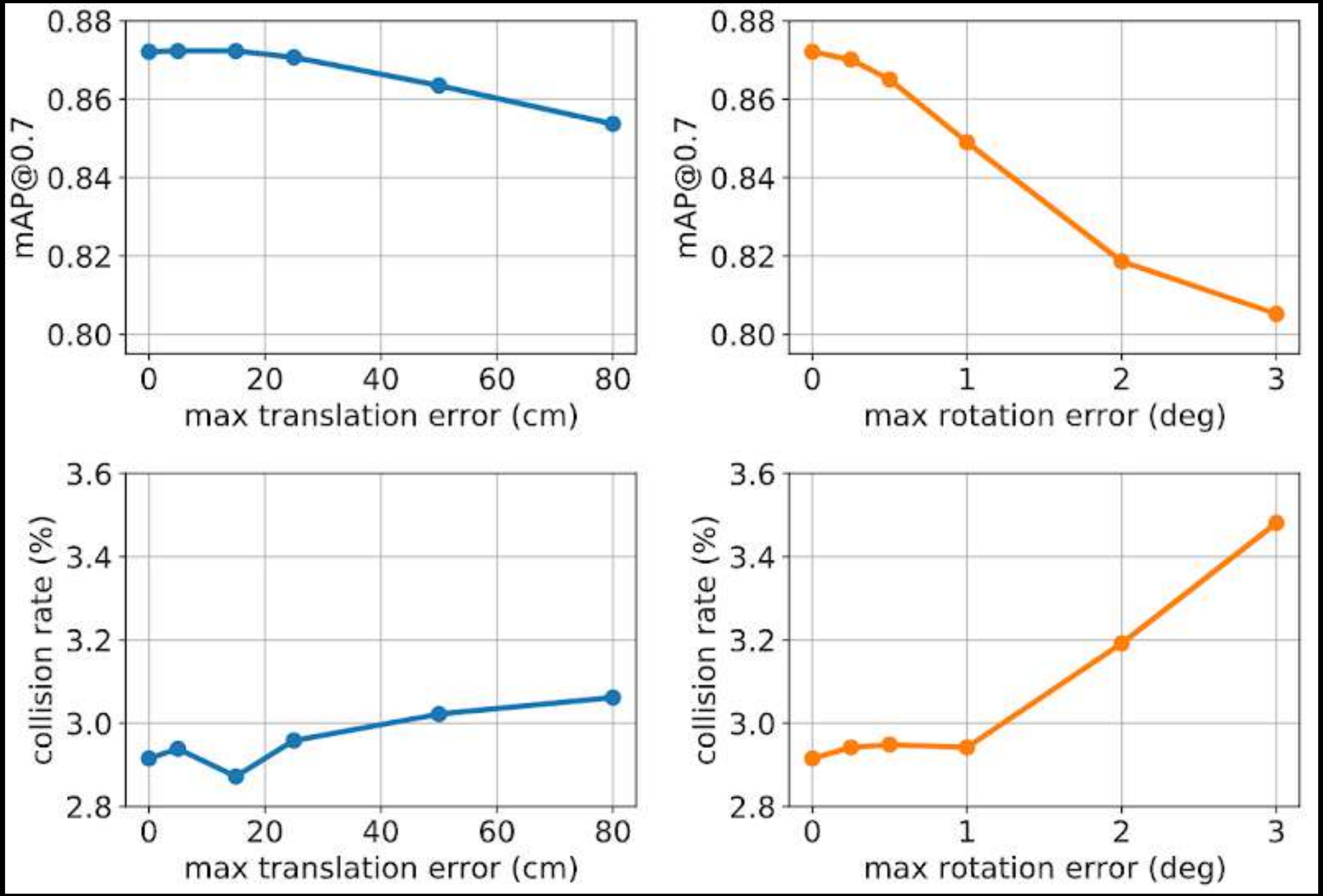
Perception



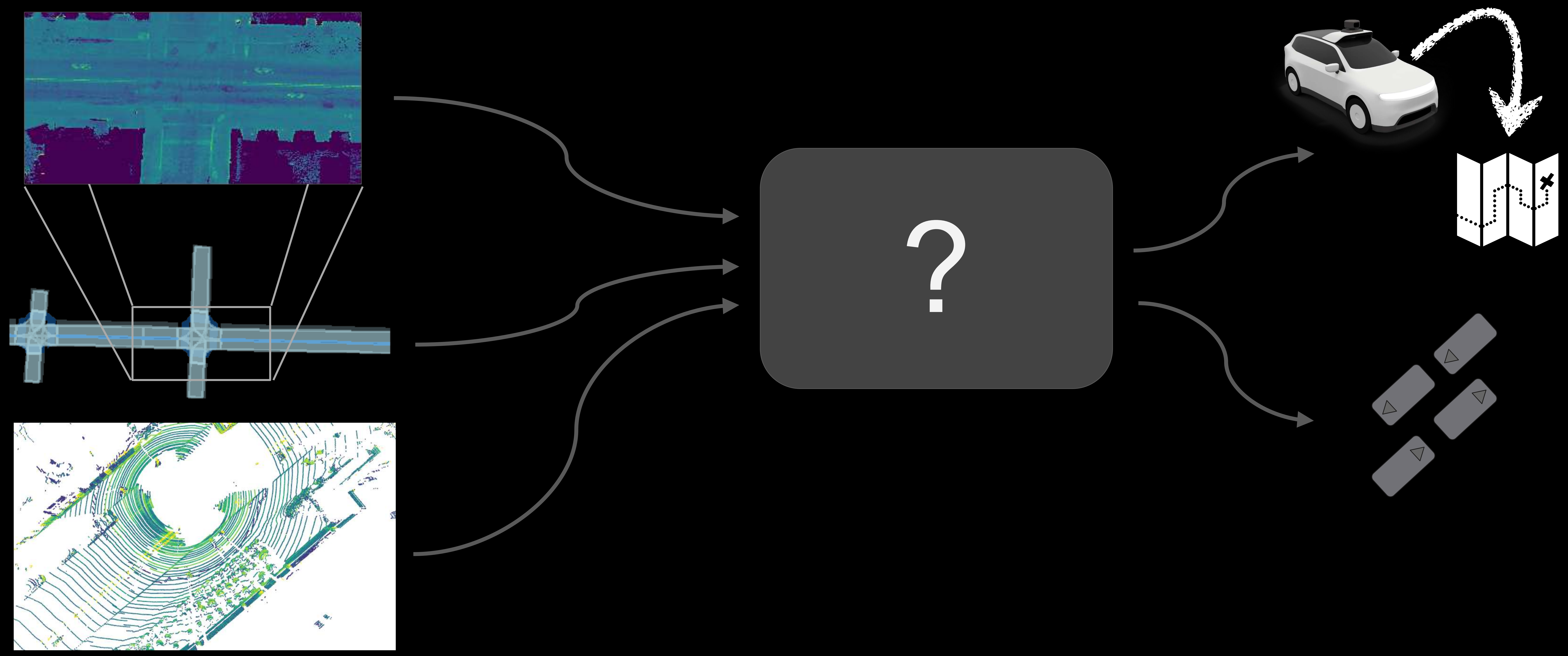
The Effects of Localization Error

Perception

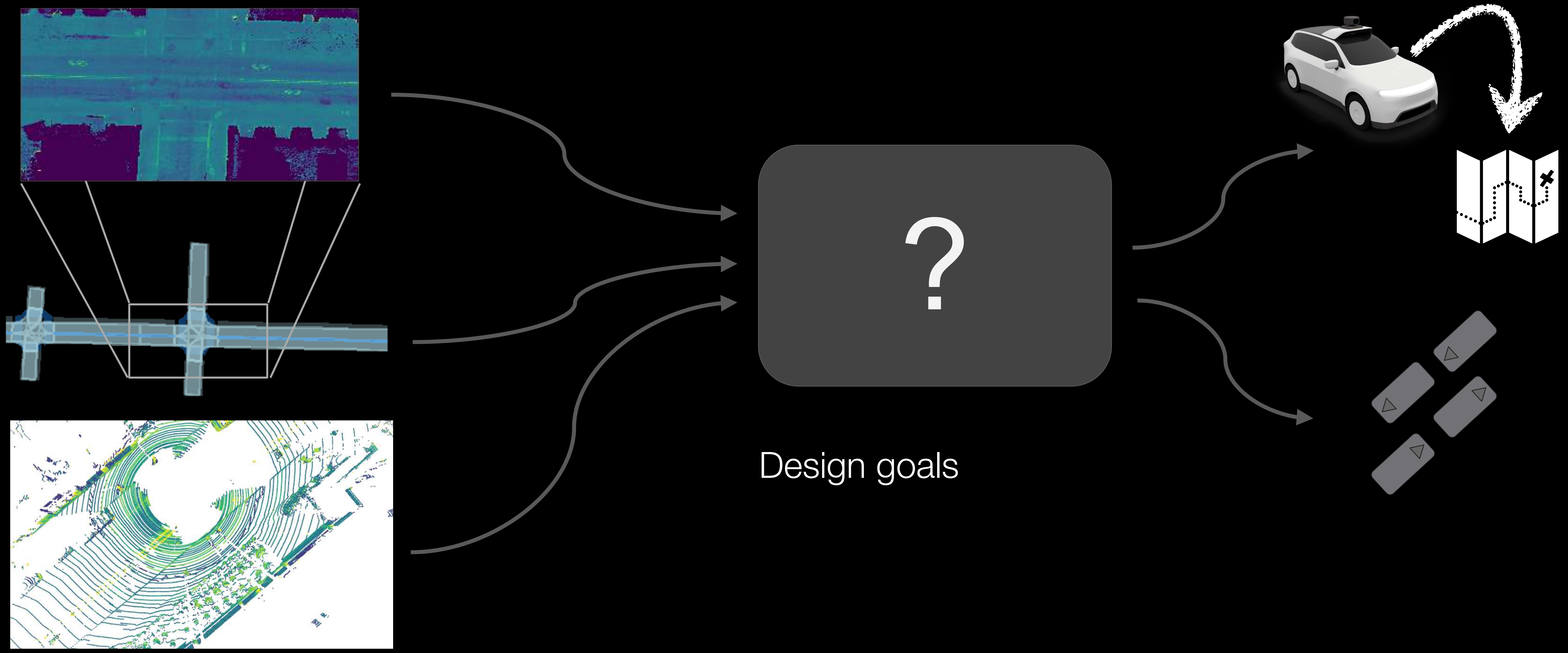
Motion Planning



Joint Localization, Perception, and Prediction

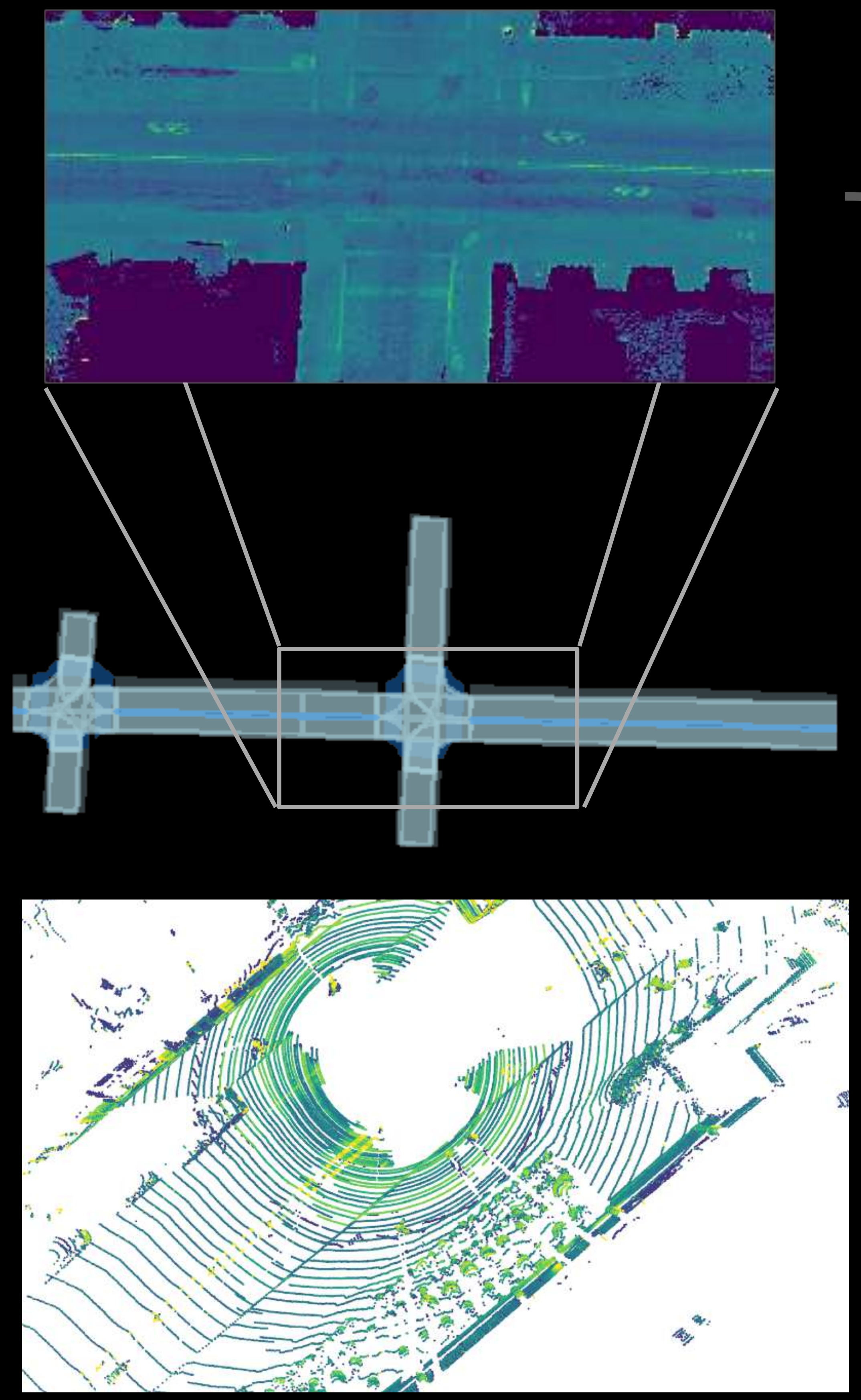


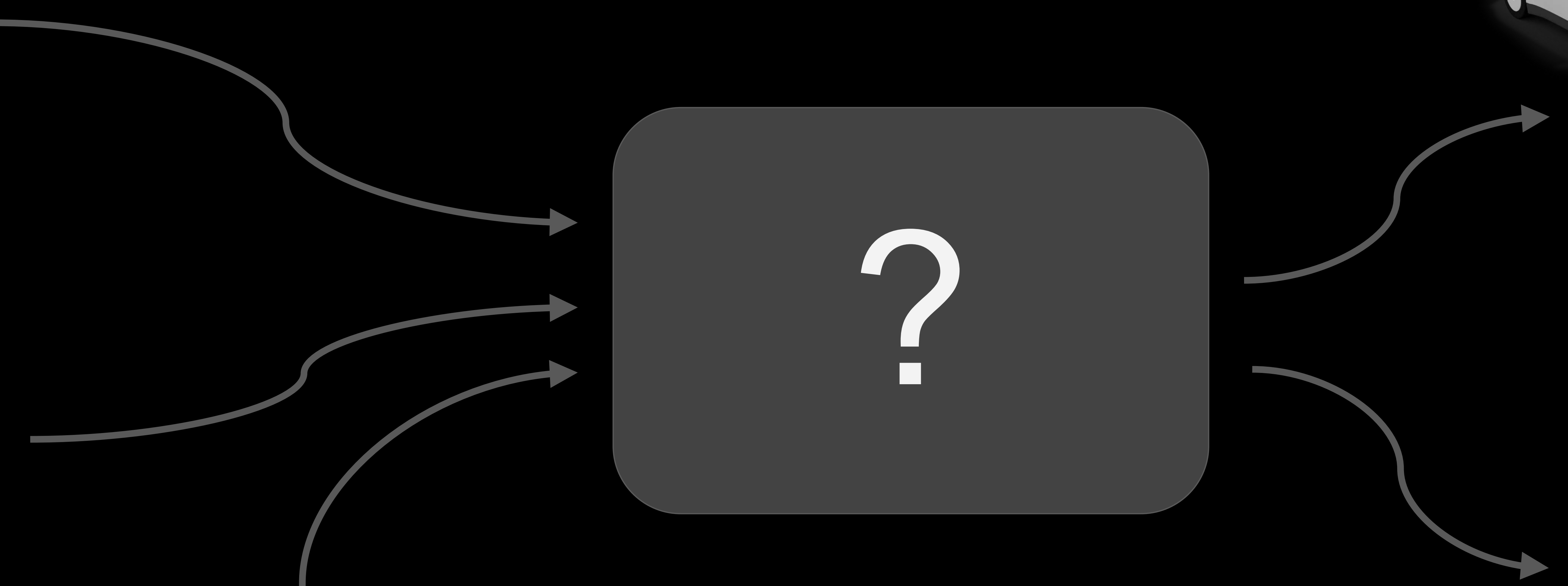
Joint Localization, Perception, and Prediction

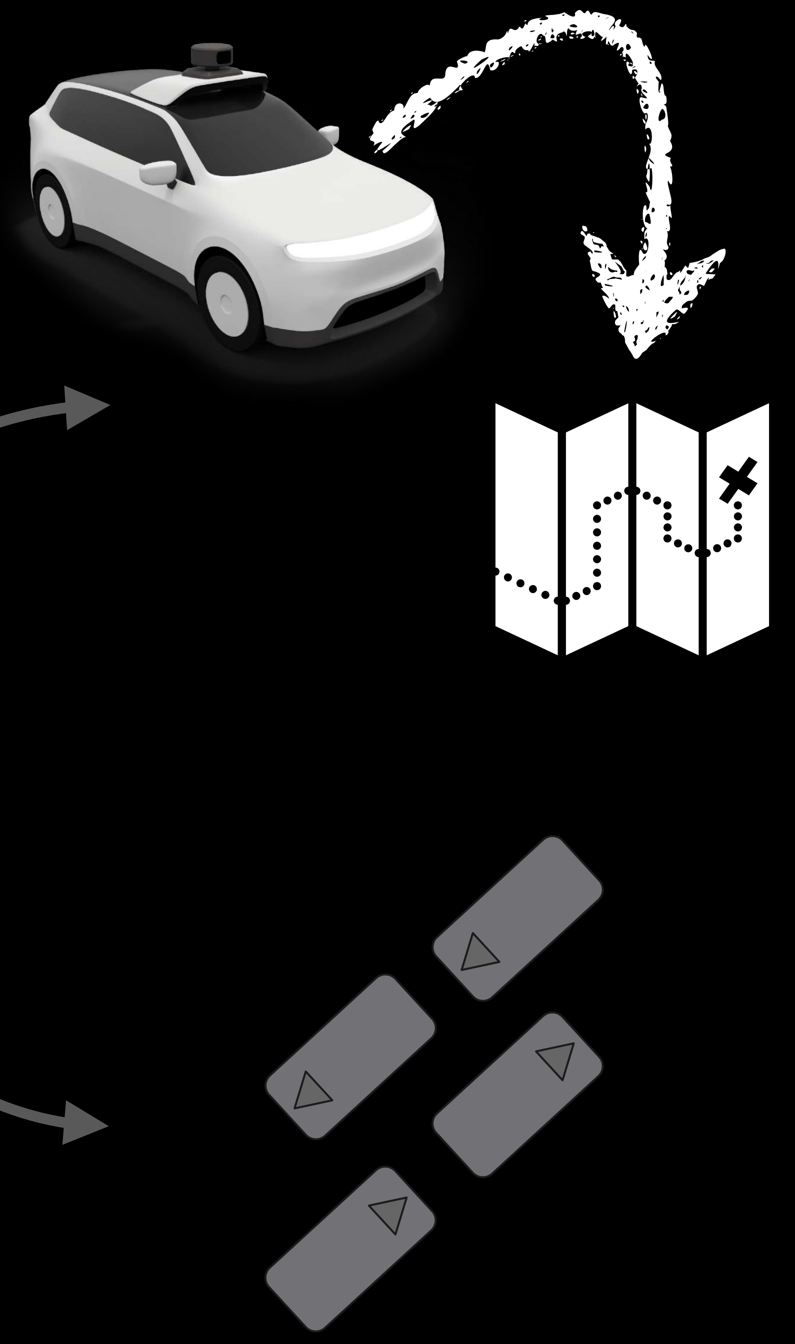


Design goals

Joint Localization, Perception, and Prediction



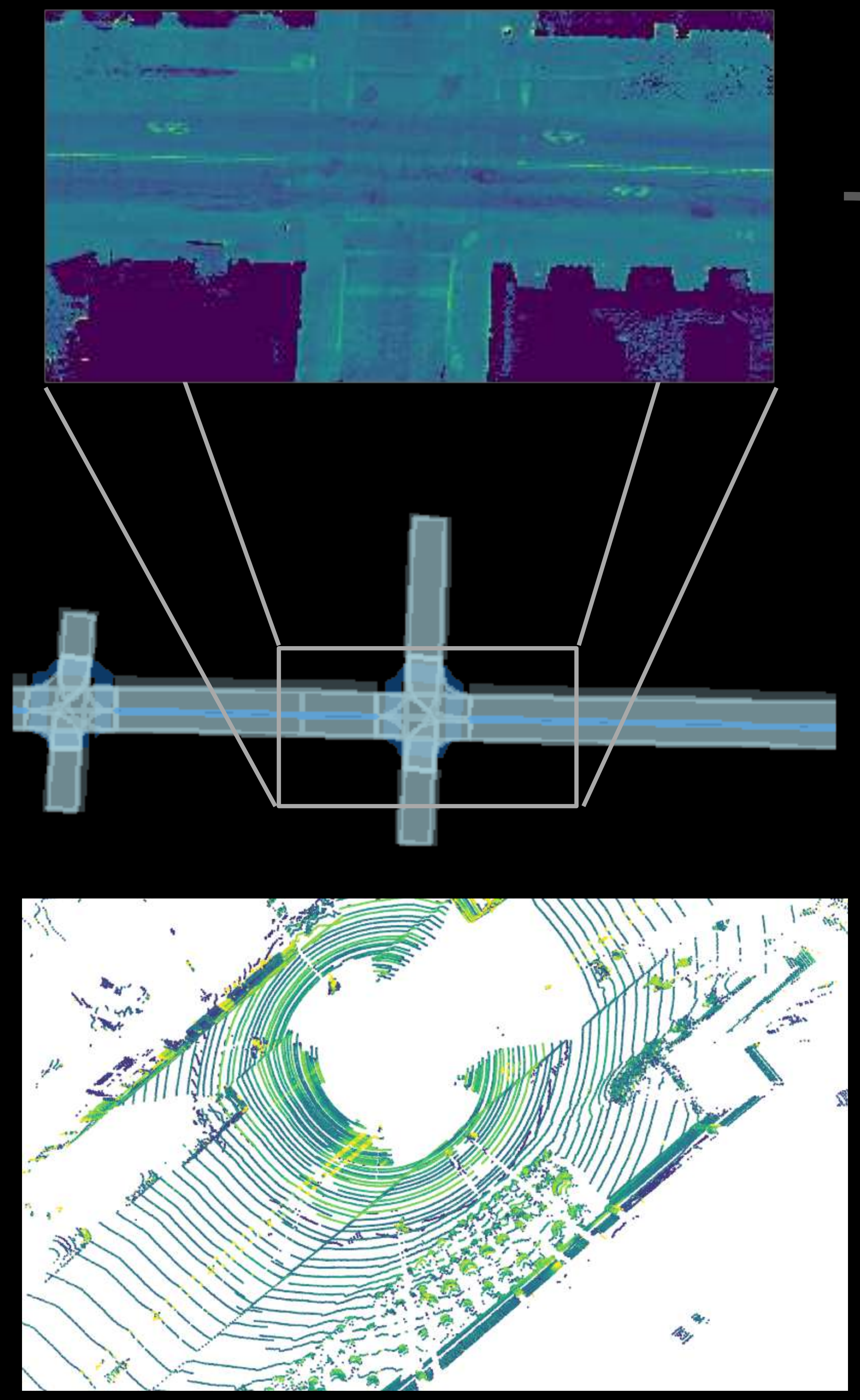


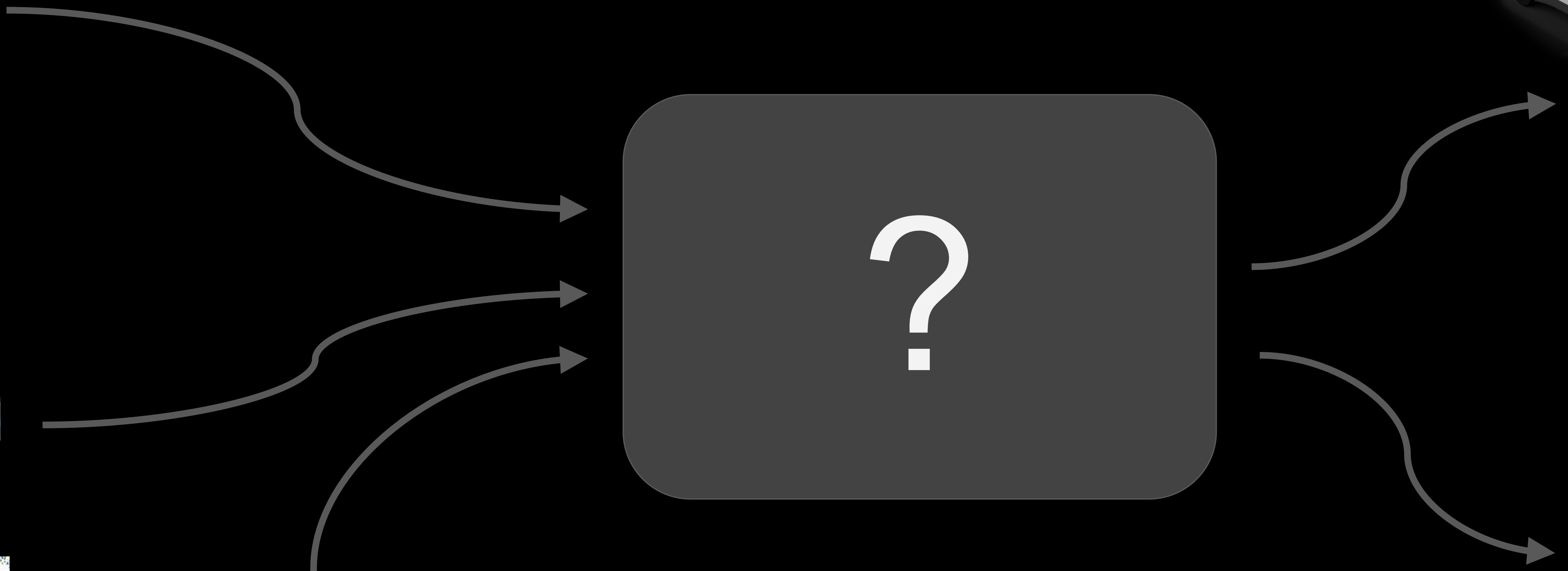


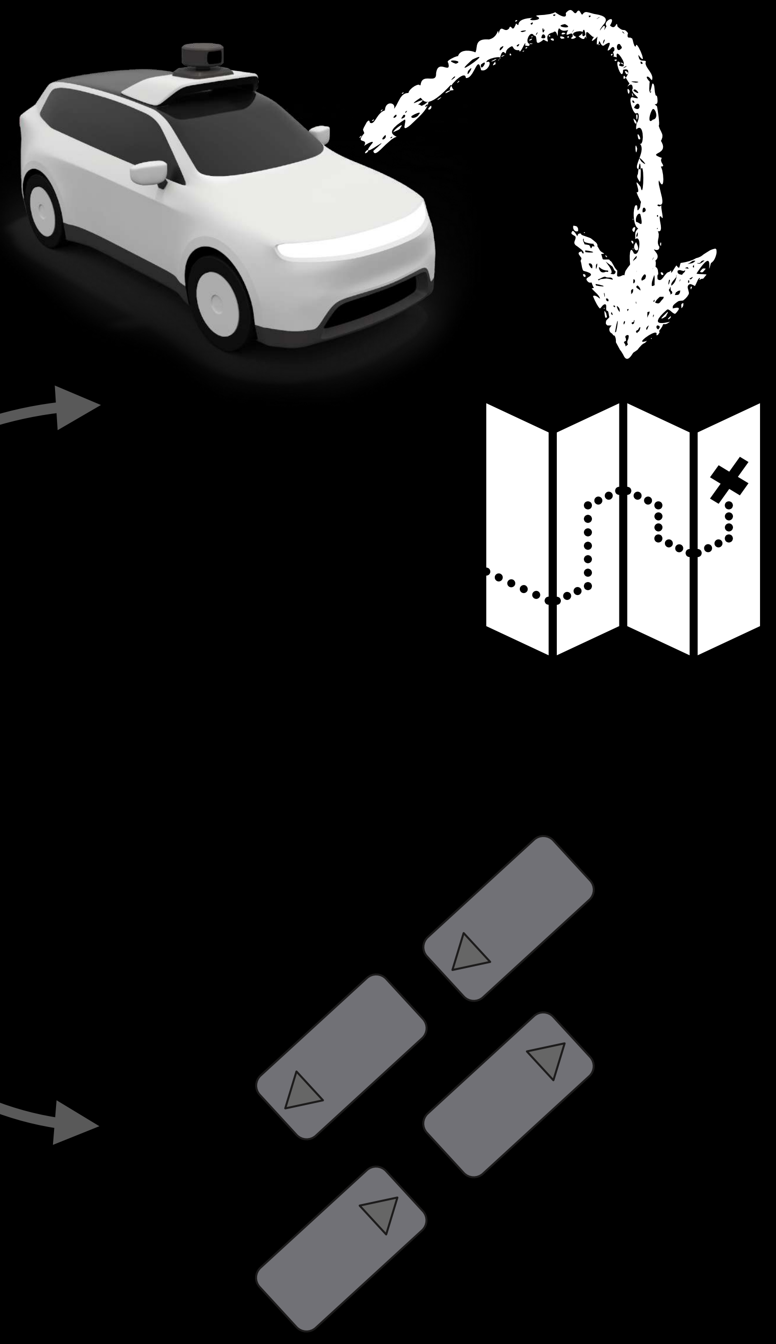
Design goals

- Low latency ⏱️

Joint Localization, Perception, and Prediction



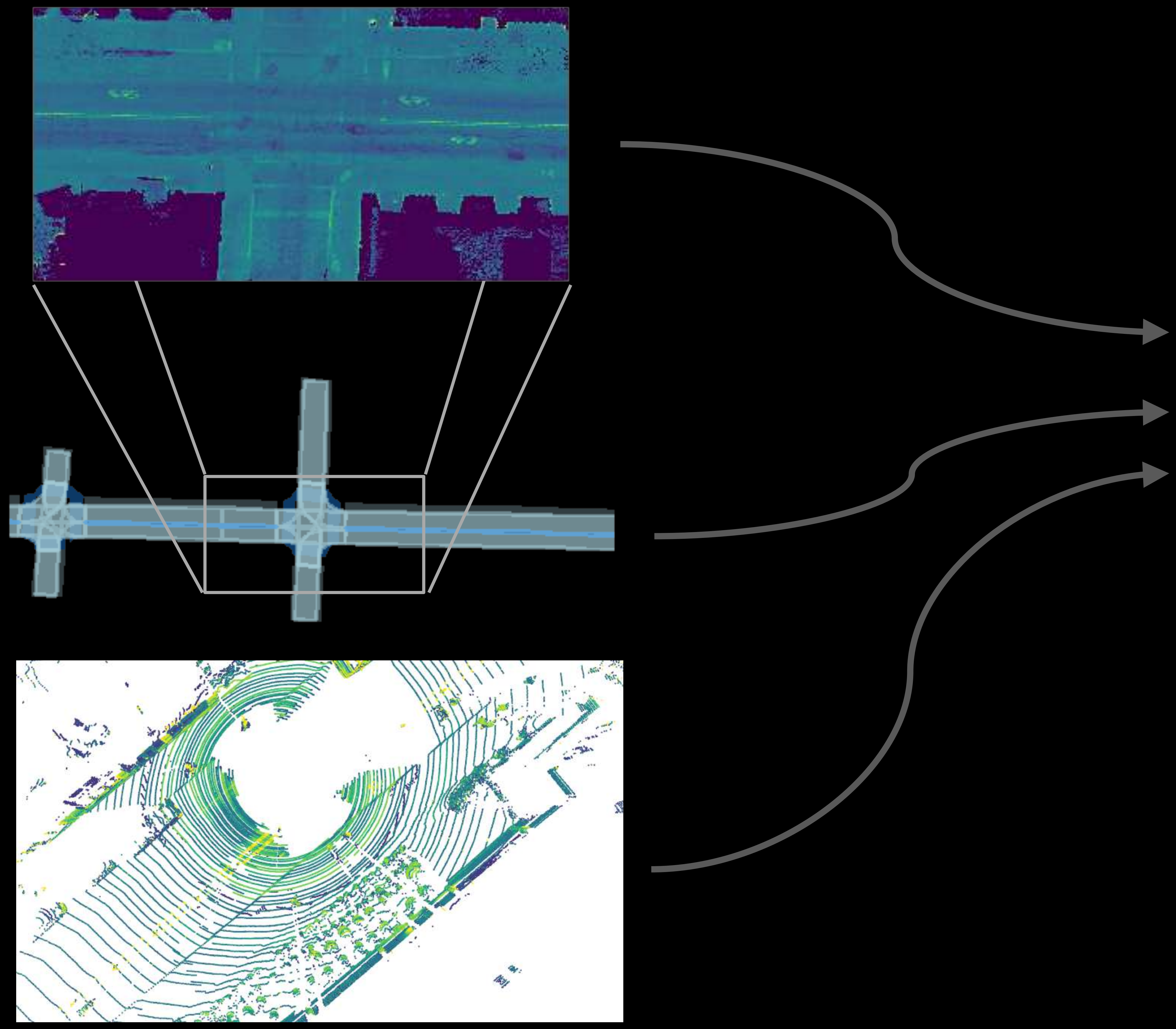


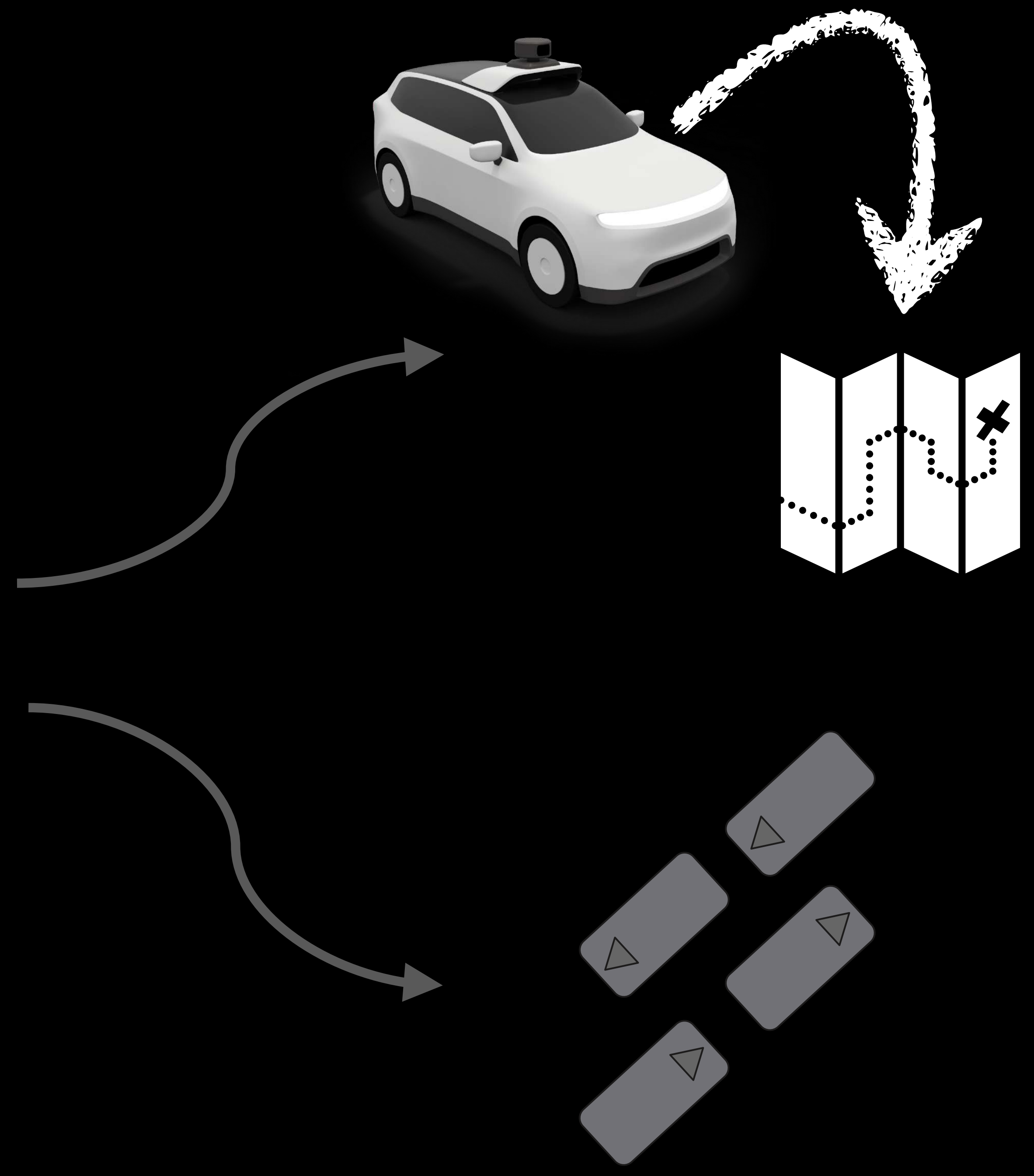


Design goals

- Low latency ⏱️
- Learning-based localization 🤖

Joint Localization, Perception, and Prediction

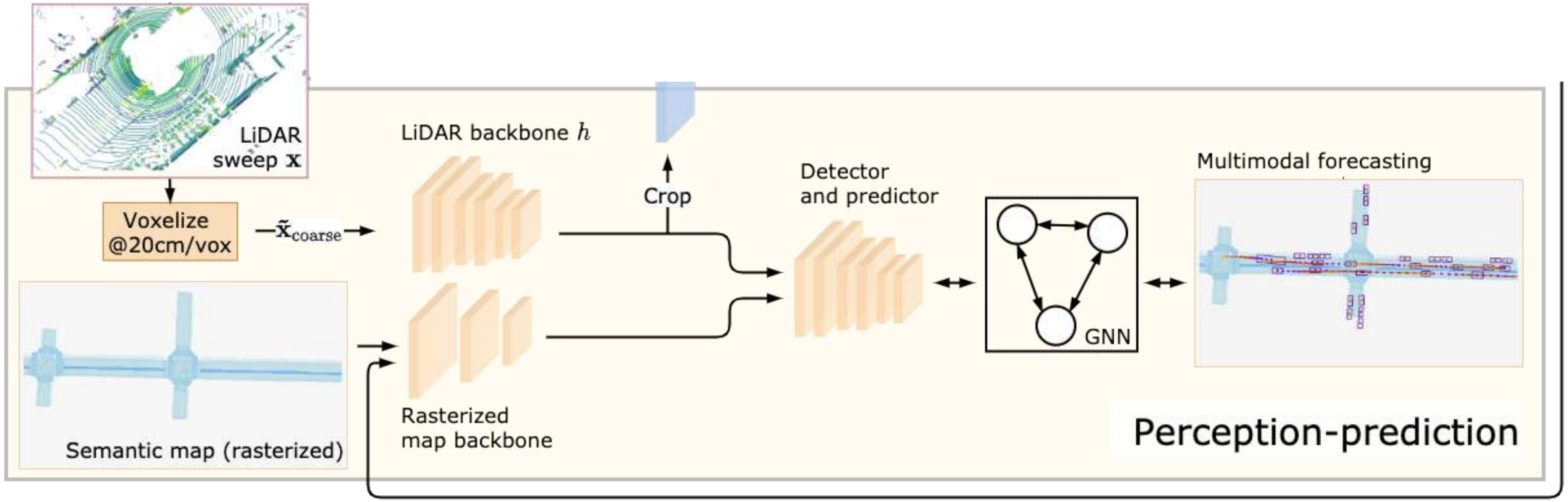




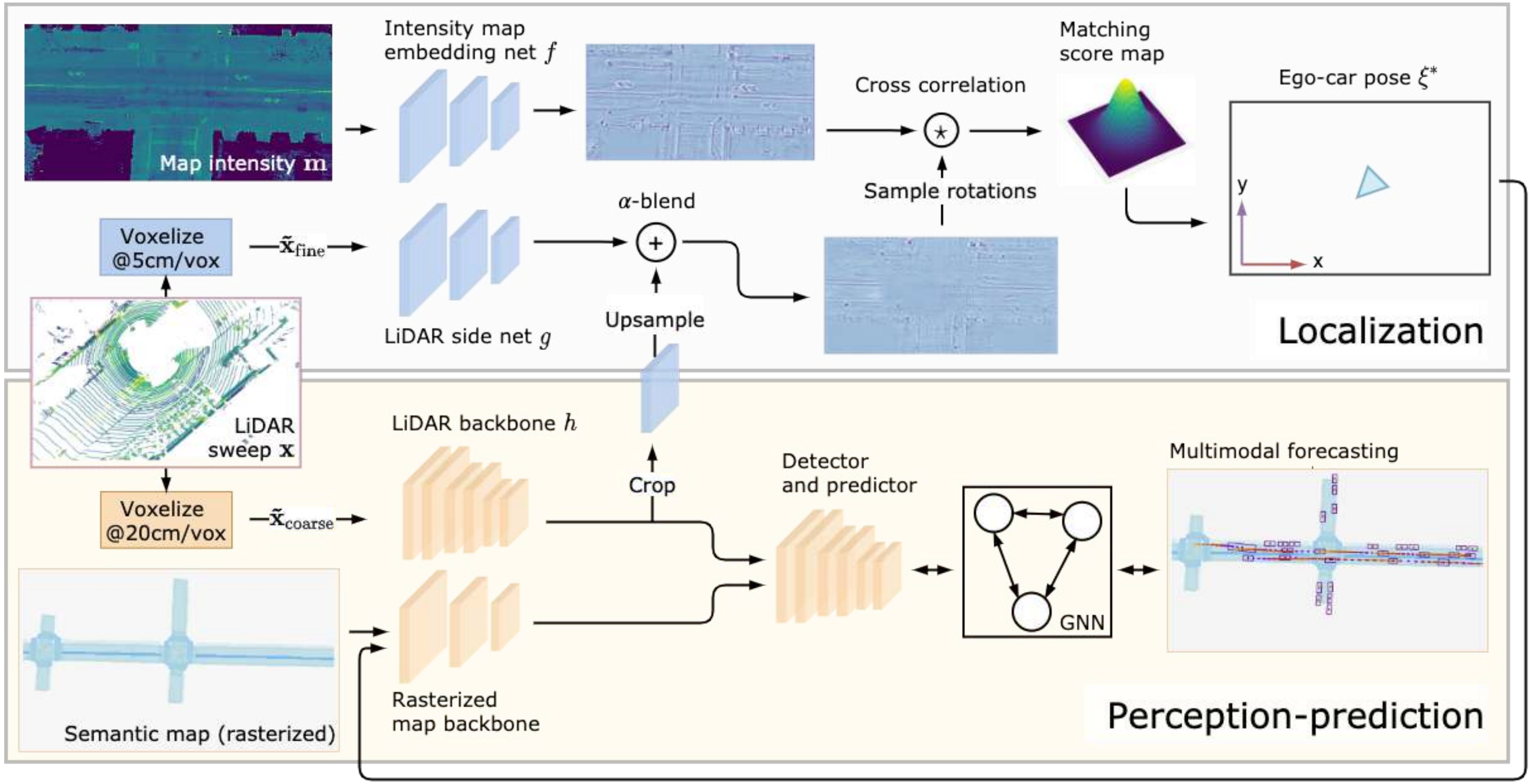
Design goals

- Low latency ⏱️
- Learning-based localization 🤖
- Easy to train and deploy 📉

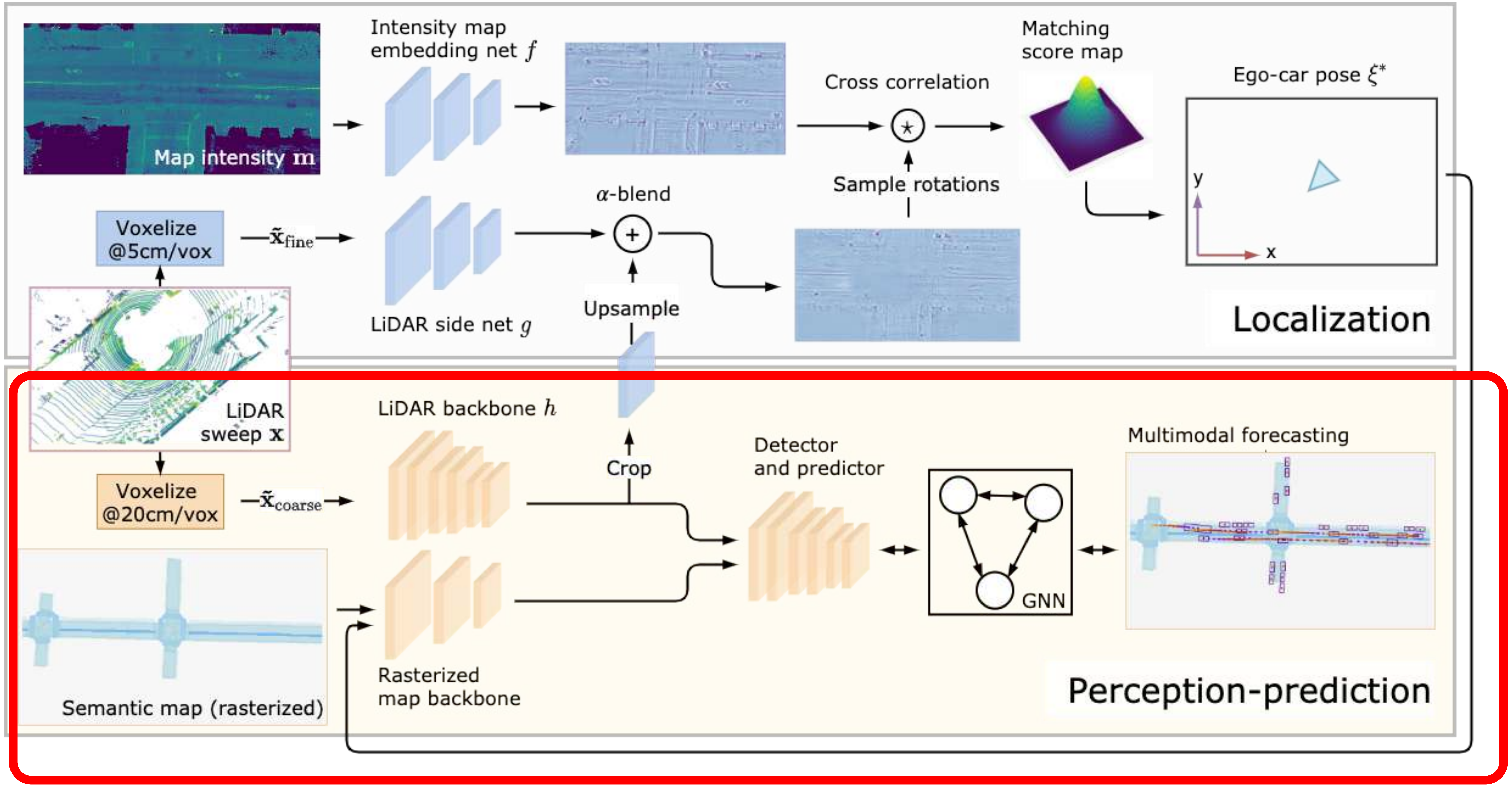
Localization + Perception + Prediction



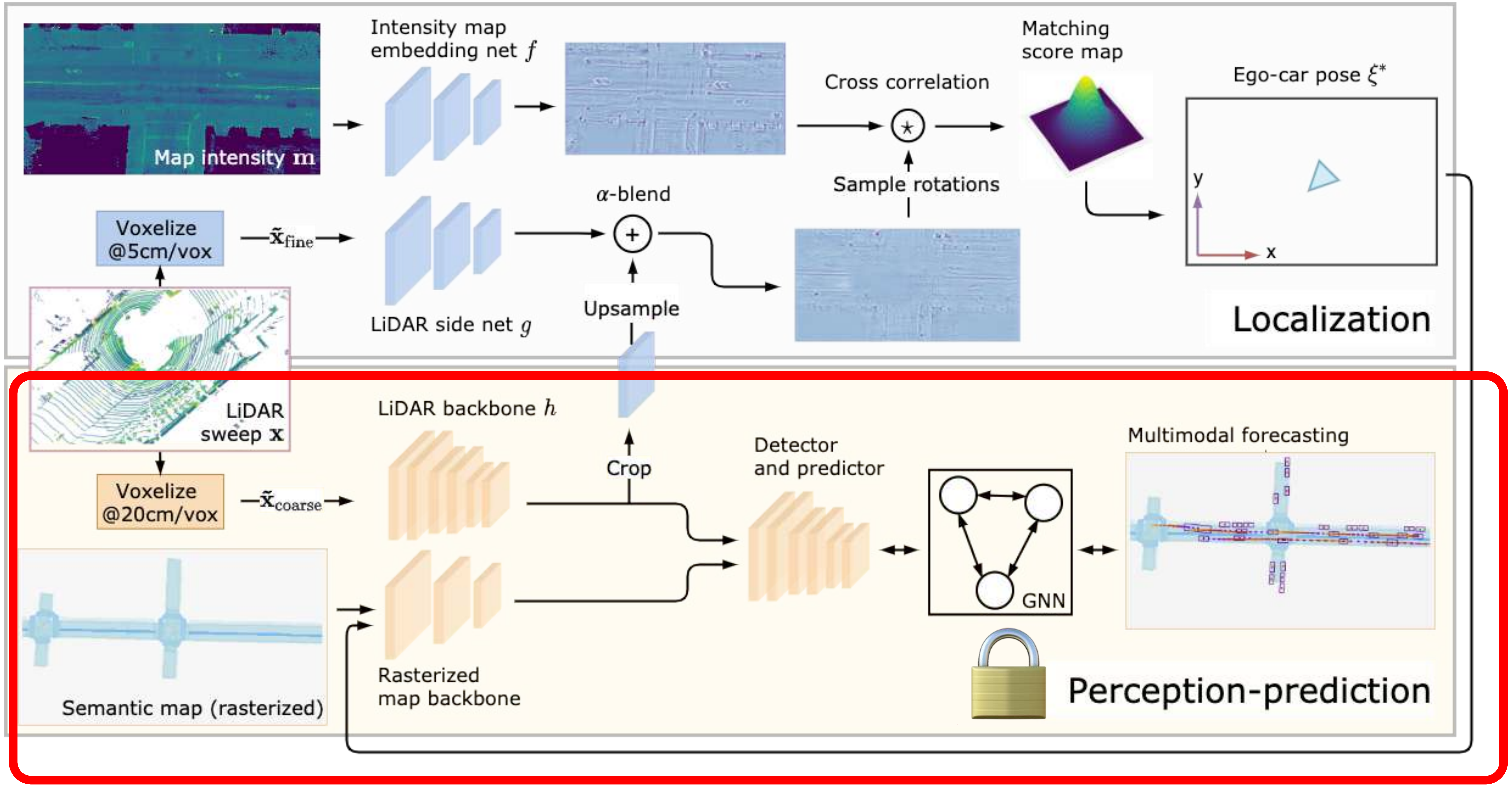
Localization + Perception + Prediction



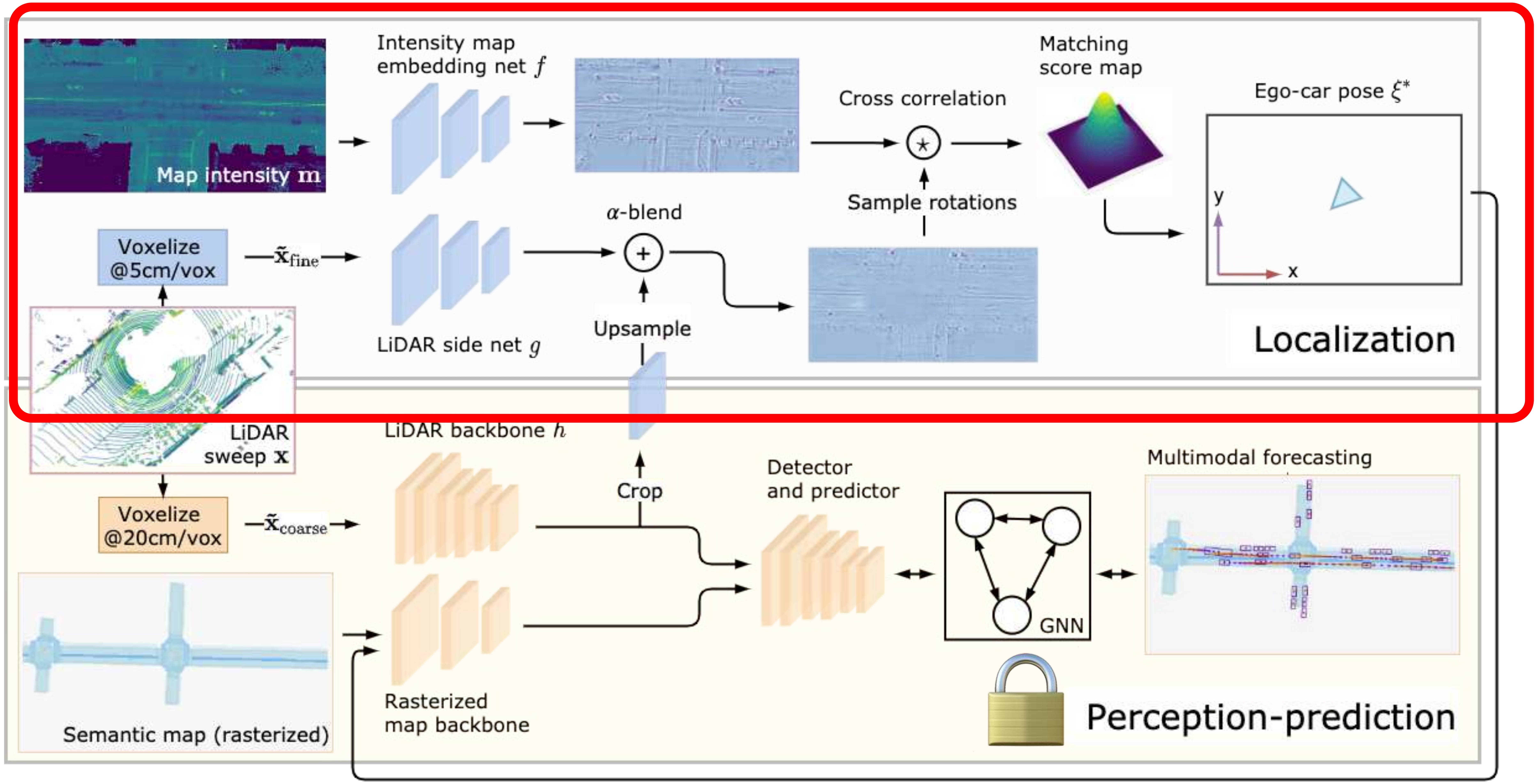
Localization + Perception + Prediction



Localization + Perception + Prediction

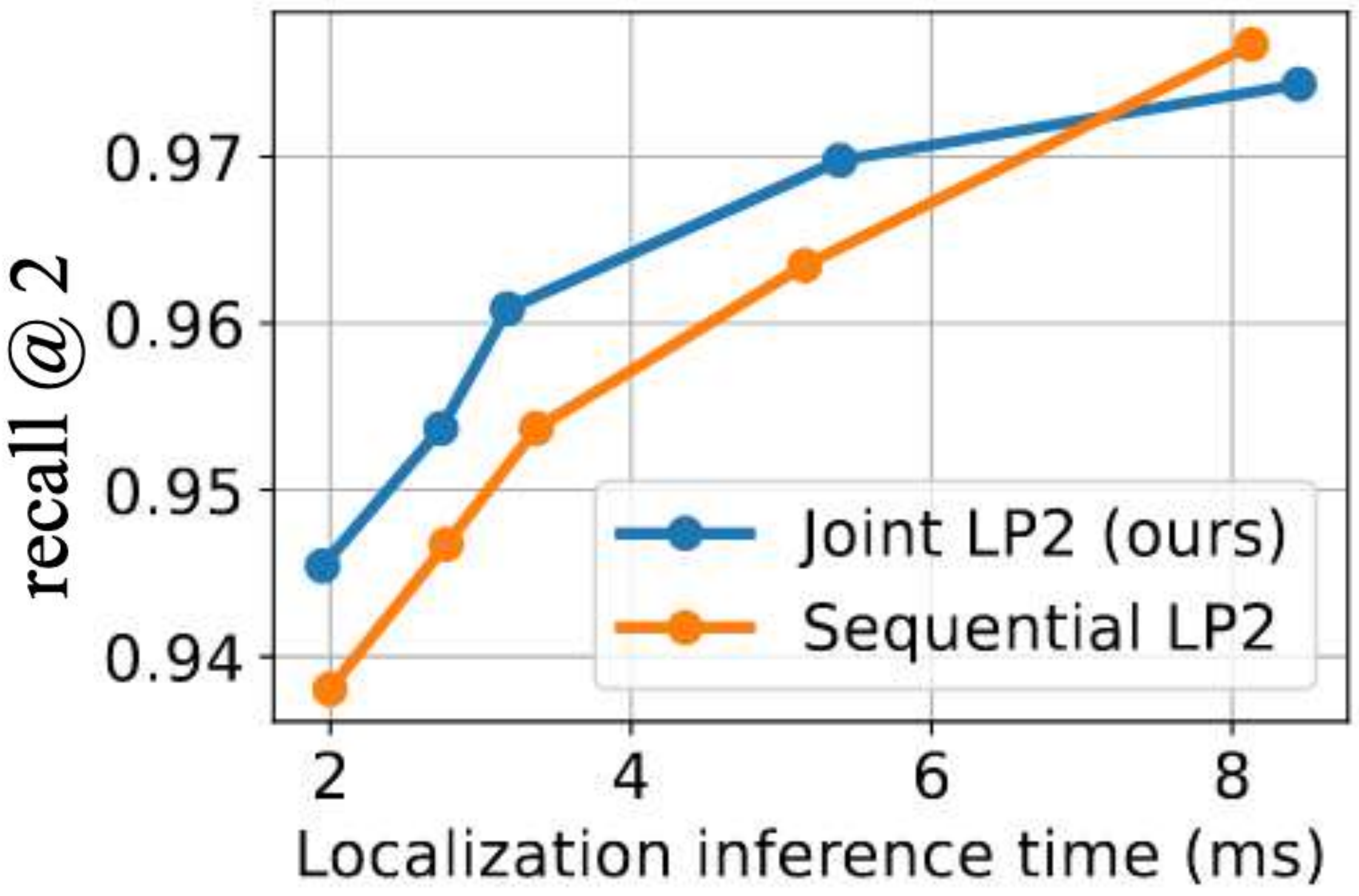


Localization + Perception + Prediction



Key Result

Fast localization without sacrificing perception quality.



Take-Home

- Localization error
  - Sub-20cm = little impact on perception and planning
  - Larger errors affect motion planning more
- Multi-task learning
  - Can significantly reduce inference time
  - Seemingly unrelated tasks like localization and detection can benefit from each other
- Incremental training
  - Helps manage model complexity
  - Avoids catastrophic forgetting

Further Reading

- See the website (andreibarsan.github.io/multi-task-lp3/) for:
  - Paper PDF (Phillips et al., CVPR 2021)
  - 5-min video with more details
  - See you at our CVPR 2021 poster if you're attending!

Project Website

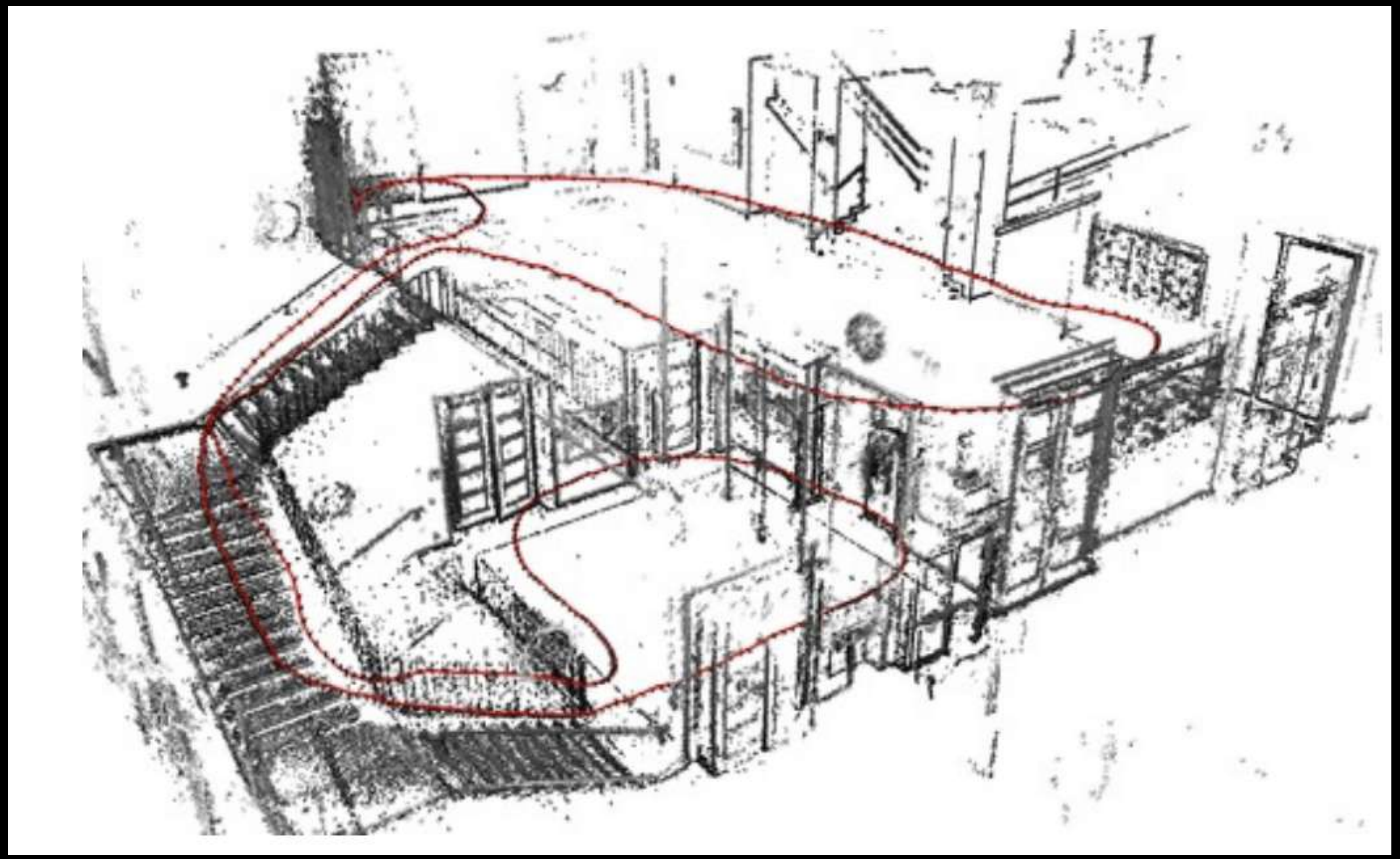
Simultaneous Localization and Mapping

Based on joint work with Anqi Joyce Yang, Can Cui, Raquel Urtasun, and Shenlong Wang
Asynchronous Multi-View SLAM (ICRA 2021)

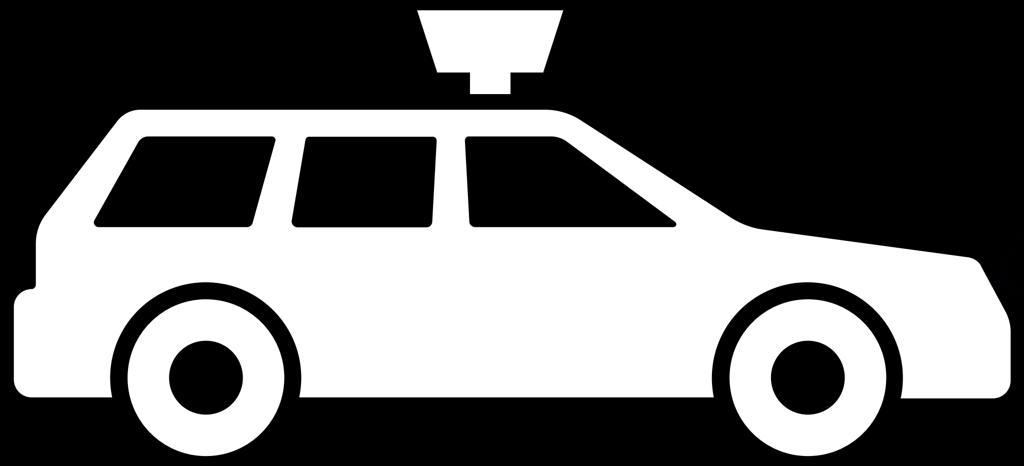
Skipped during the talk in the interest of time.
Check out the paper for more details!

Simultaneous Localization and Mapping (SLAM)

- Localize by building a map at the same time
- Applications:
  - Navigation in unknown areas without prior maps
  - Building HD maps
  - Augmented & virtual reality
- Focus on **camera-based** SLAM (visual SLAM)

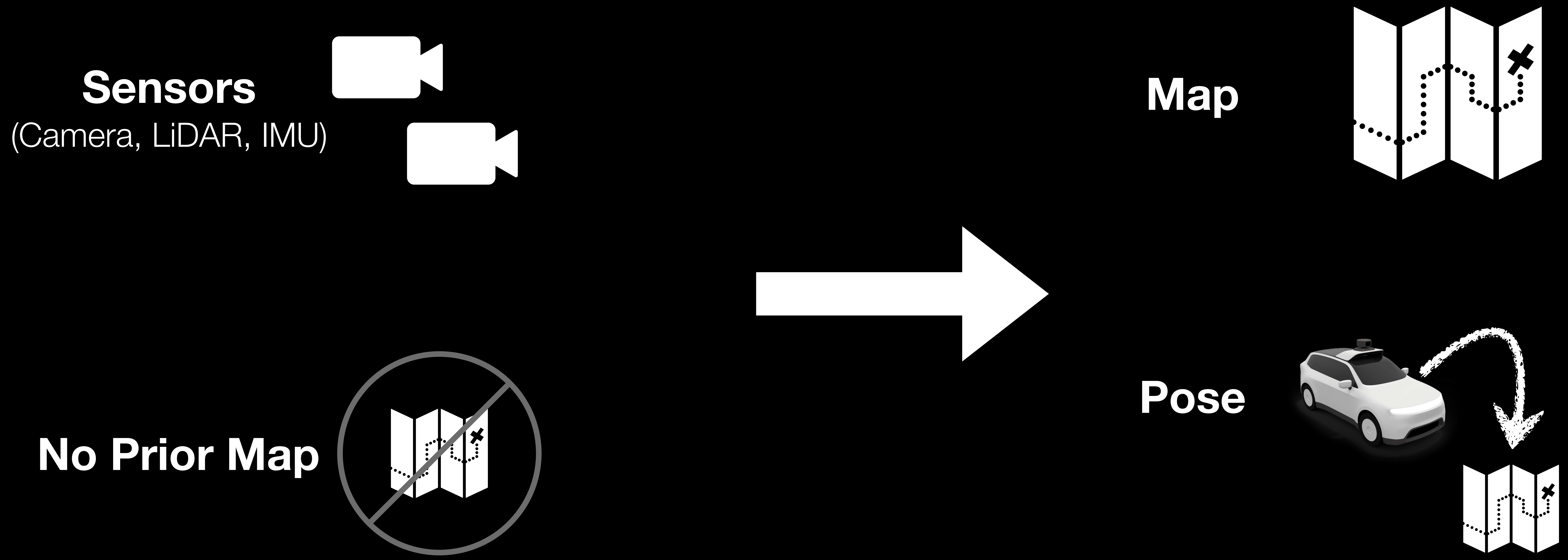


Visualization from Engel et al., 2017

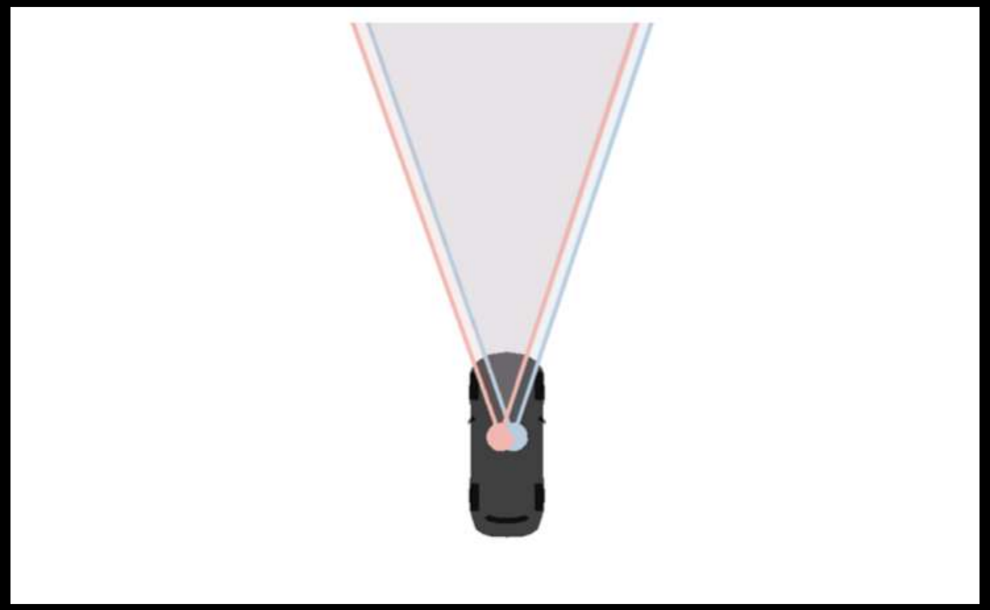


Drone image from Skydio

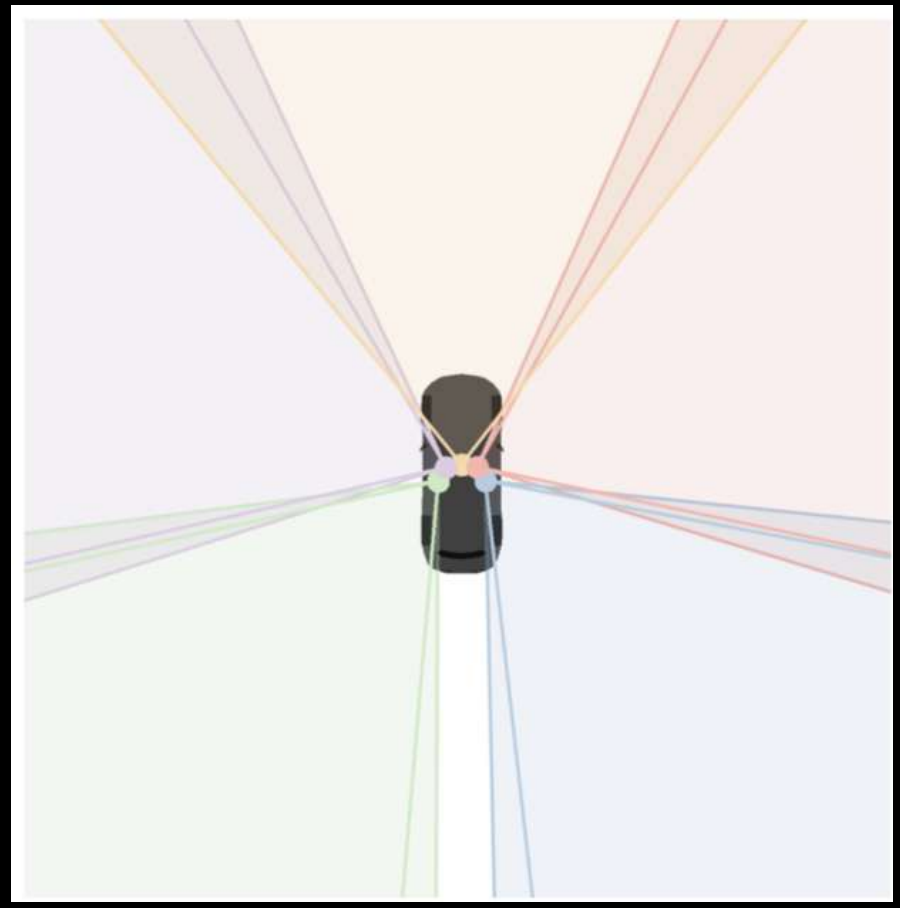
SLAM Problem Statement



Camera Rigs in Visual SLAM



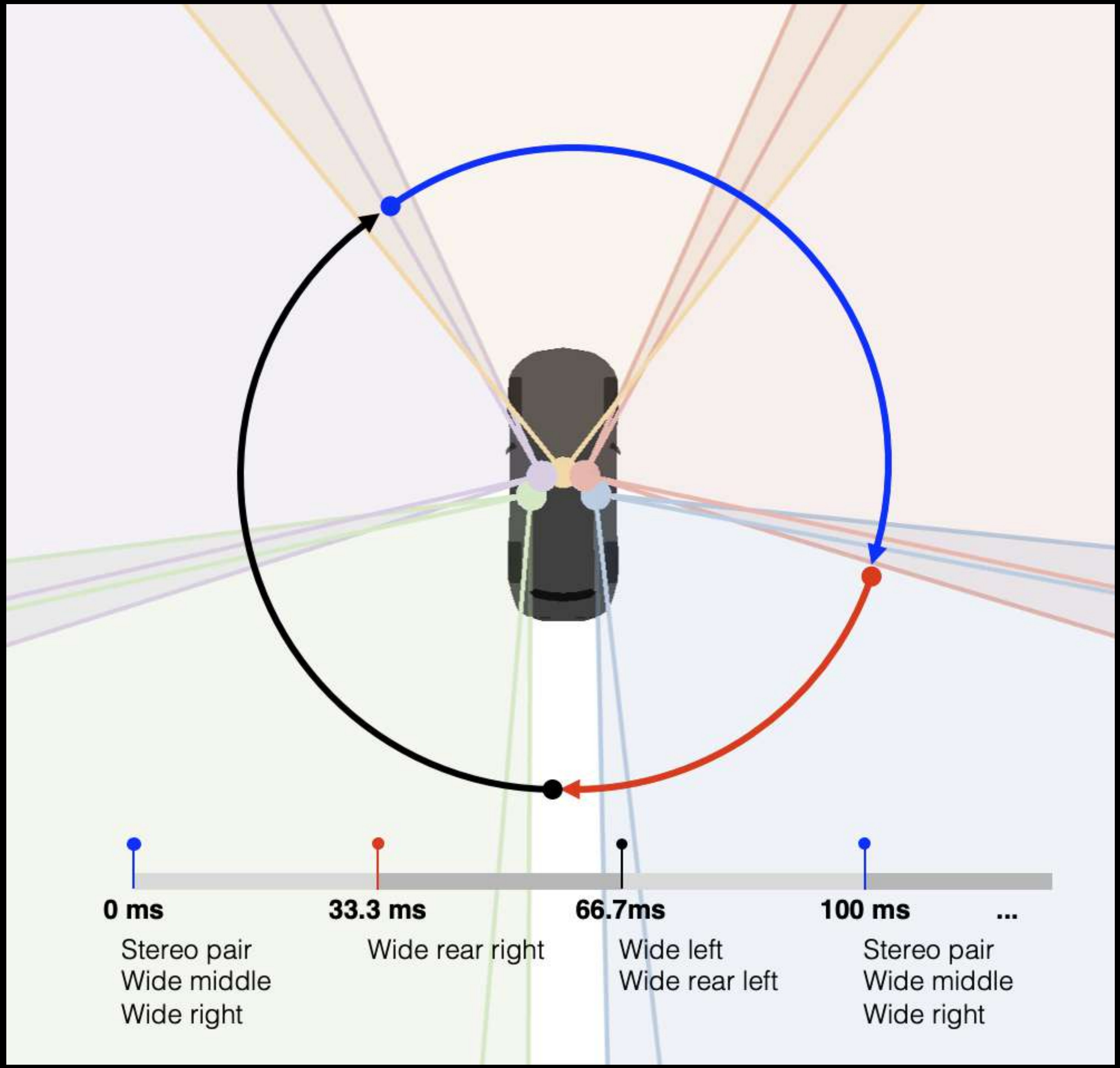
FoV of a stereo camera pair



FoV of five wide cameras

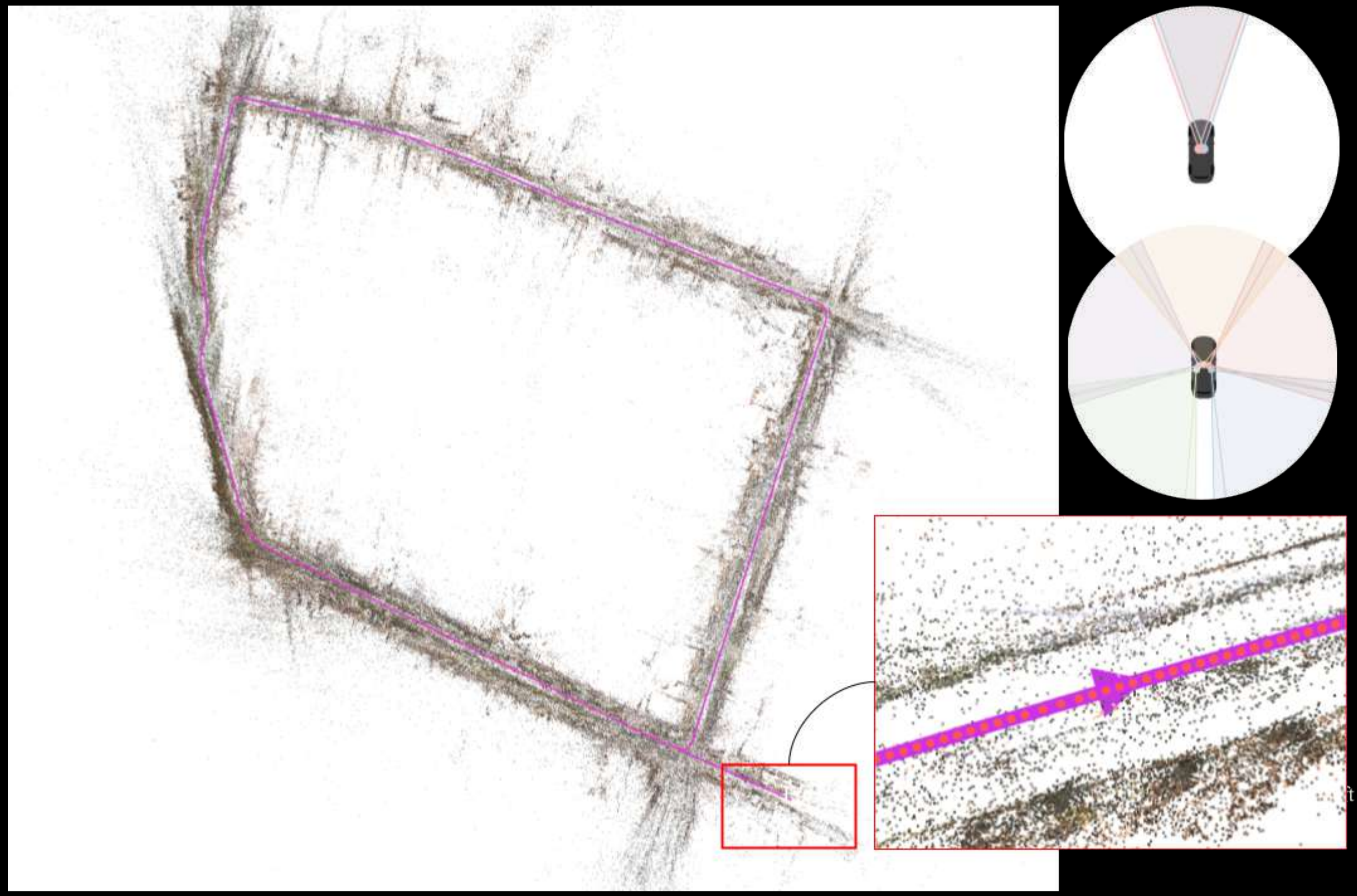
Asynchronous Modeling

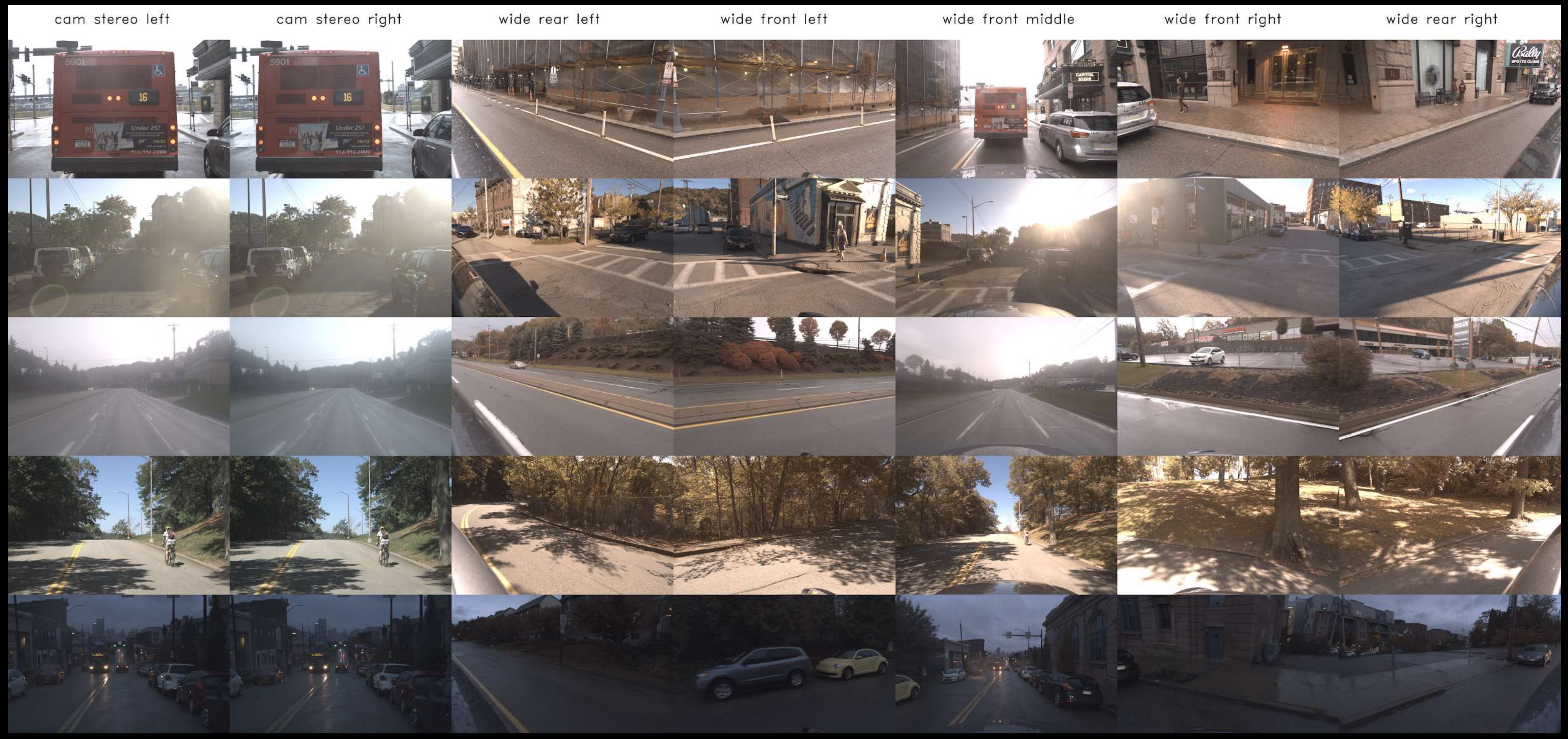
- Existing multi-view SLAM systems all assume **synchronous** camera shutters
- In practice cameras can be asynchronous due to technical limitations, or by design, e.g. synced to a LiDAR



Studying Asynchronous SLAM

- General multi-view **SLAM framework** agnostic to camera firing times
- A large-scale outdoor **SLAM dataset** with multiple cameras, diverse environments, and accurate ground-truth for evaluation

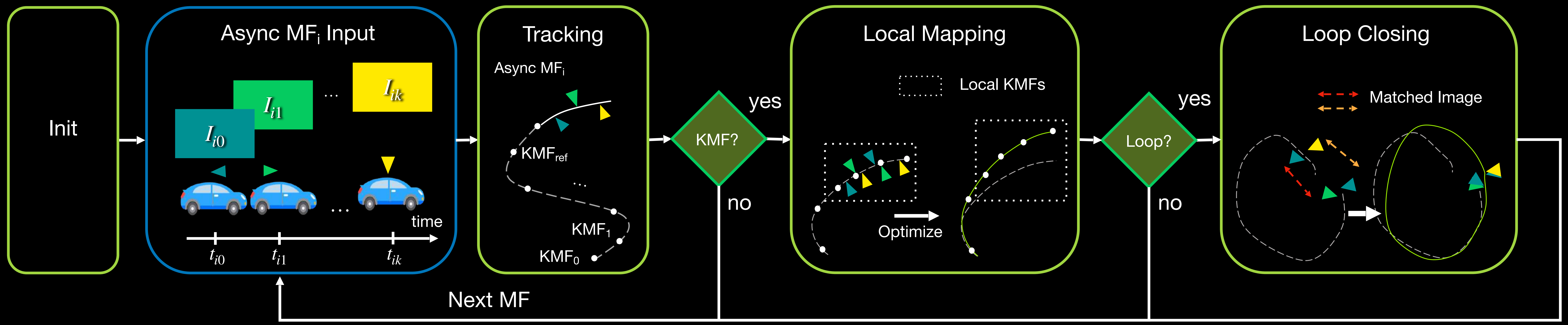




Asynchronous SLAM

- Camera images come at different times
- Group nearby images into **multi-**frames
- Continuous time trajectory estimation allows async information fusion
  - In practice, we use B-splines

Asynchronous SLAM Pipeline



Asynchronous SLAM vs. Baselines

Method	median RPE (translation, m/m)	median RPE (rotation, mrad/m)	ATE (m)	Success Rate
DSO Mono	42.72	0.802	594.39	62.67%
ORB-SLAM Mono	32.00	0.549	694.37	64.00%
ORB-SLAM Stereo	1.85	0.329	30.74	77.33%
Sync-Stereo	<u>1.30</u>	<u>0.291</u>	<u>24.53</u>	<u>80.00%</u>
Sync-All	2.15	0.347	58.18	74.67%
Async-All (Ours)	**0.35**	**0.113**	**6.13**	**92.00%**

Asynchronous SLAM vs. Baselines

Method	median RPE (translation, m/m)	median RPE (rotation, mrad/m)	ATE (m)	Success Rate
DSO Mono	42.72	0.802	594.39	62.67%
ORB-SLAM Mono	32.00	0.549	694.37	64.00%
ORB-SLAM Stereo	1.85	0.329	30.74	77.33%
Sync-Stereo	<u>1.30</u>	<u>0.291</u>	<u>24.53</u>	<u>80.00%</u>
Sync-All	2.15	0.347	58.18	74.67%
Async-All (Ours)	**0.35**	**0.113**	**6.13**	**92.00%**

Asynchronous SLAM vs. Baselines

Method	median RPE (translation, m/m)	median RPE (rotation, mrad/m)	ATE (m)	Success Rate
DSO Mono	42.72	0.802	594.39	62.67%
ORB-SLAM Mono	32.00	0.549	694.37	64.00%
ORB-SLAM Stereo	1.85	0.329	30.74	77.33%
Sync-Stereo	<u>1.30</u>	<u>0.291</u>	<u>24.53</u>	<u>80.00%</u>
Sync-All	2.15	0.347	58.18	74.67%
Async-All (Ours)	**0.35**	**0.113**	**6.13**	**92.00%**

Asynchronous SLAM vs. Baselines

Method	median RPE (translation, m/m)	median RPE (rotation, mrad/m)	ATE (m)	Success Rate
DSO Mono	42.72	0.802	594.39	62.67%
ORB-SLAM Mono	32.00	0.549	694.37	64.00%
ORB-SLAM Stereo	1.85	0.329	30.74	77.33%
Sync-Stereo	<u>1.30</u>	<u>0.291</u>	<u>24.53</u>	<u>80.00%</u>
Sync-All	2.15	0.347	58.18	74.67%
Async-All (Ours)	**0.35**	**0.113**	**6.13**	**92.00%**

Asynchronous SLAM vs. Baselines

Method	median RPE (translation, m/m)	median RPE (rotation, mrad/m)	ATE (m)	Success Rate
DSO Mono	42.72	0.802	594.39	62.67%
ORB-SLAM Mono	32.00	0.549	694.37	64.00%
ORB-SLAM Stereo	1.85	0.329	30.74	77.33%
Sync-Stereo	1.30	0.291	24.53	80.00%
Sync-All	2.15	0.347	58.18	74.67%
Async-All (Ours)	**0.35**	**0.113**	**6.13**	**92.00%**

Asynchronous SLAM vs. Baselines

Method	**median RPE (translation, m/m)**	**median RPE (rotation, mrad/m)**	**ATE (m)**	**Success Rate**
DSO Mono	42.72	0.802	594.39	62.67%
ORB-SLAM Mono	32.00	0.549	694.37	64.00%
ORB-SLAM Stereo	1.85	0.329	30.74	77.33%
Sync-Stereo	1.30	0.291	24.53	80.00%
Sync-All	2.15	0.347	58.18	74.67%
Async-All (Ours)	**0.35**	**0.113**	**6.13**	**92.00%**


PIADA


PIADA

The Future

Limitations & Mapless Driving

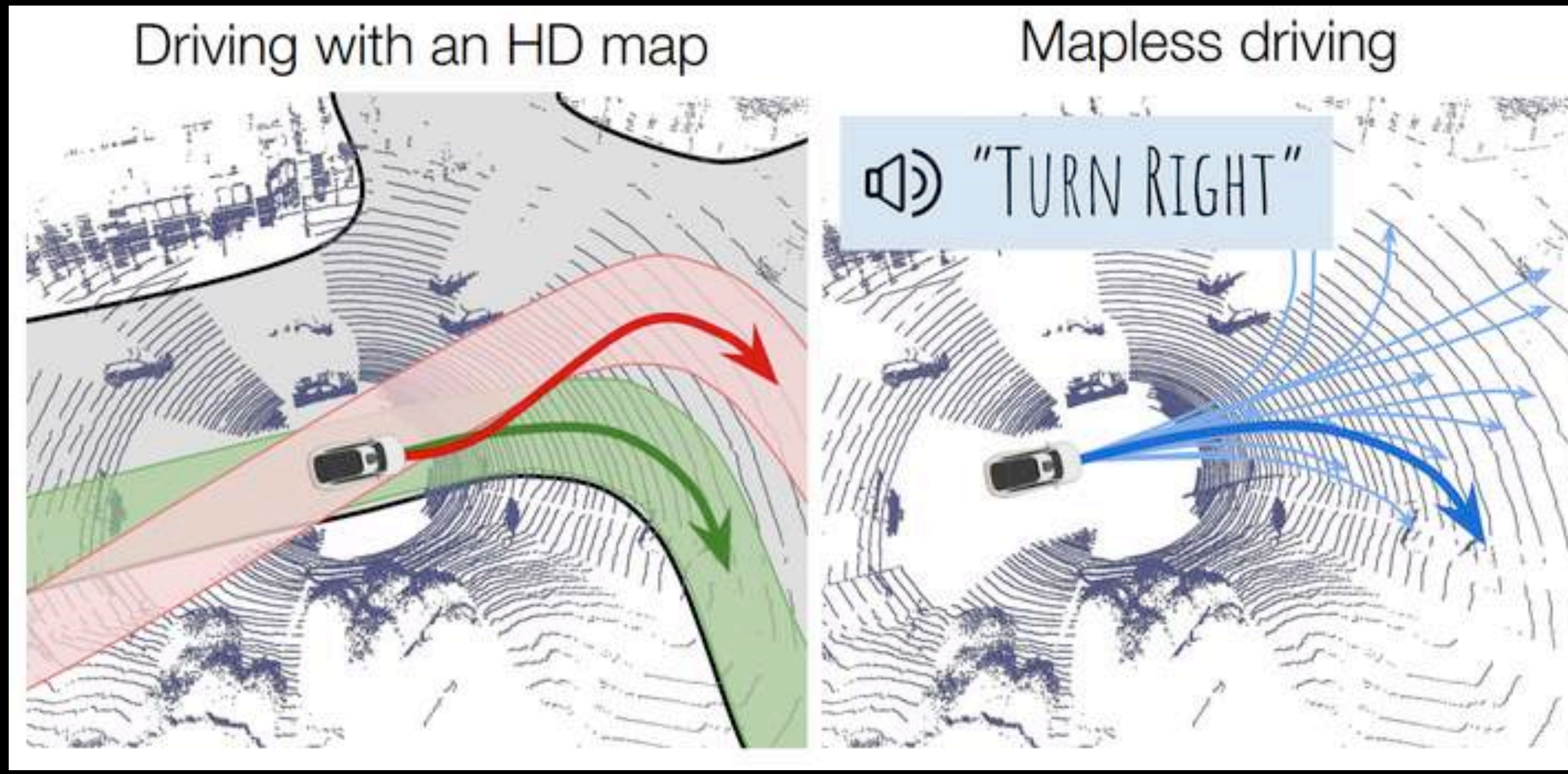


Image source: MP3 - A Unified Model to Map, Perceive, Predict and Plan
by Casas, Sadat, and Urtasun (CVPR 2021)

Limitations & Mapless Driving

- HD Maps can provide **rich prior** knowledge to autonomous agents

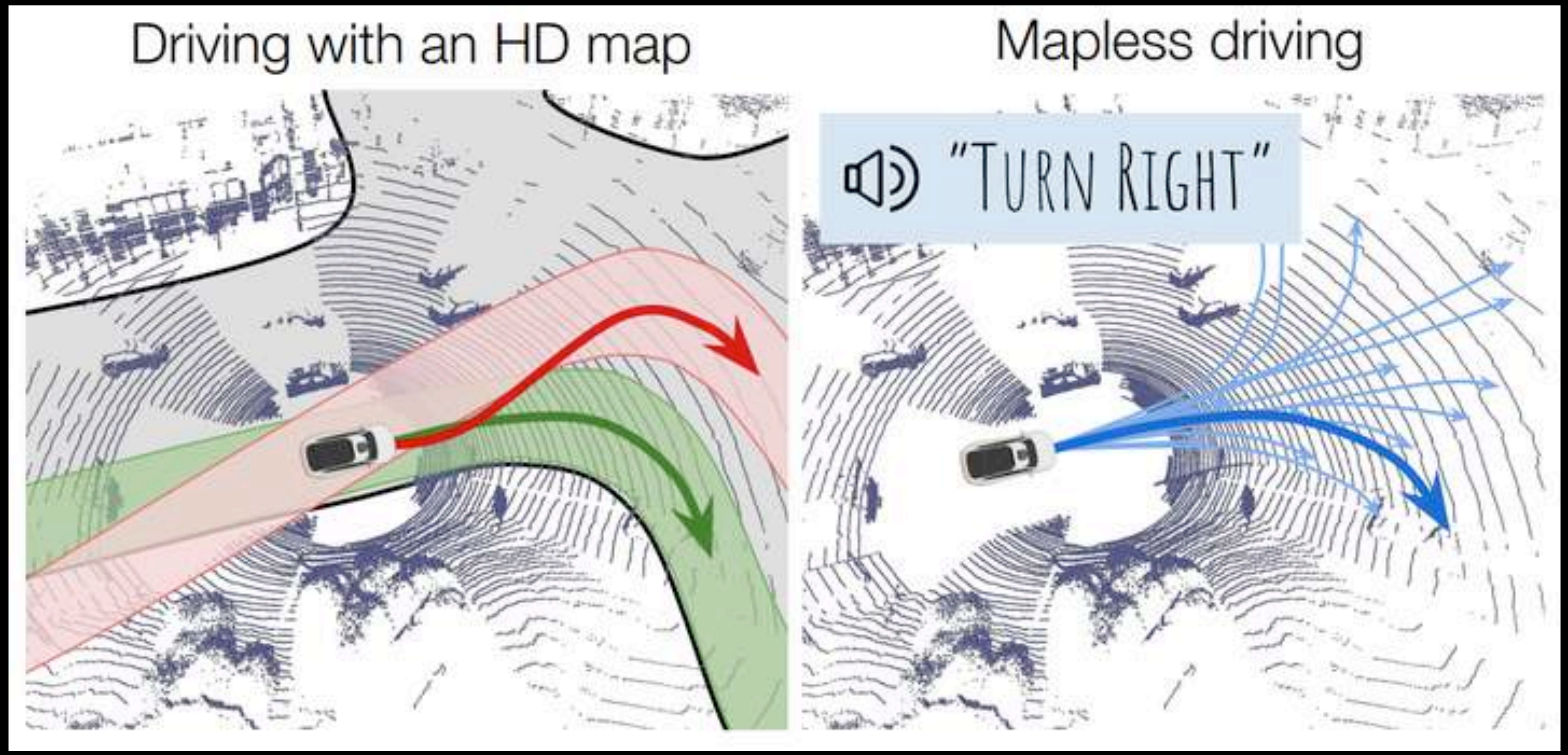


Image source: [MP3 - A Unified Model to Map, Perceive, Predict and Plan]
by Casas, Sadat, and Urtasun (CVPR 2021)

Limitations & Mapless Driving

- HD Maps can provide **rich prior** knowledge to autonomous agents
- But: We need to account for **inaccurate poses** and **outdated maps**

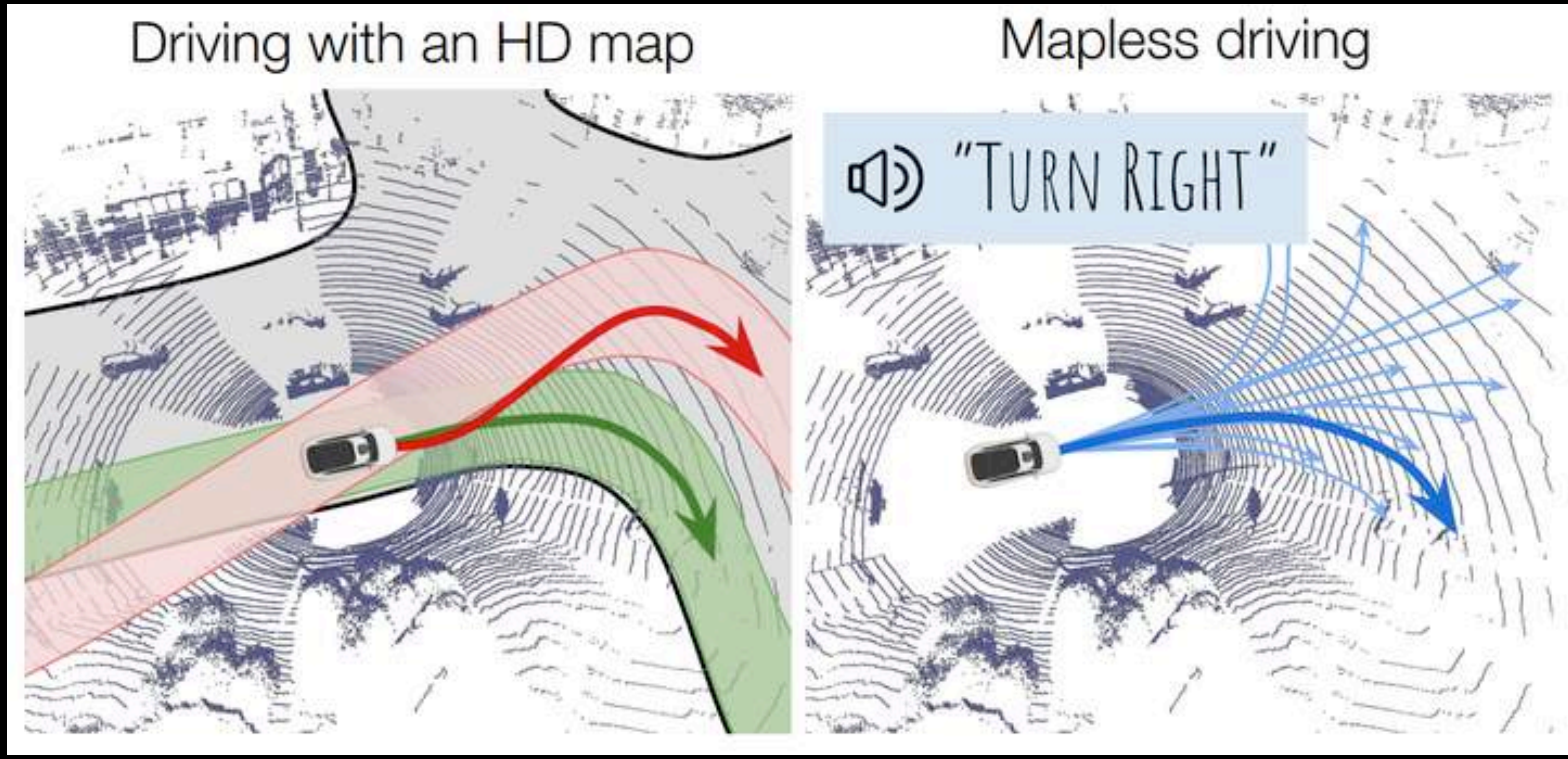


Image source: [MP3 - A Unified Model to Map, Perceive, Predict and Plan]
by Casas, Sadat, and Urtasun (CVPR 2021)

Limitations & Mapless Driving

- HD Maps can provide **rich prior** knowledge to autonomous agents
- But: We need to account for **inaccurate poses** and **outdated maps**
- Robust **mapless driving** is gaining traction

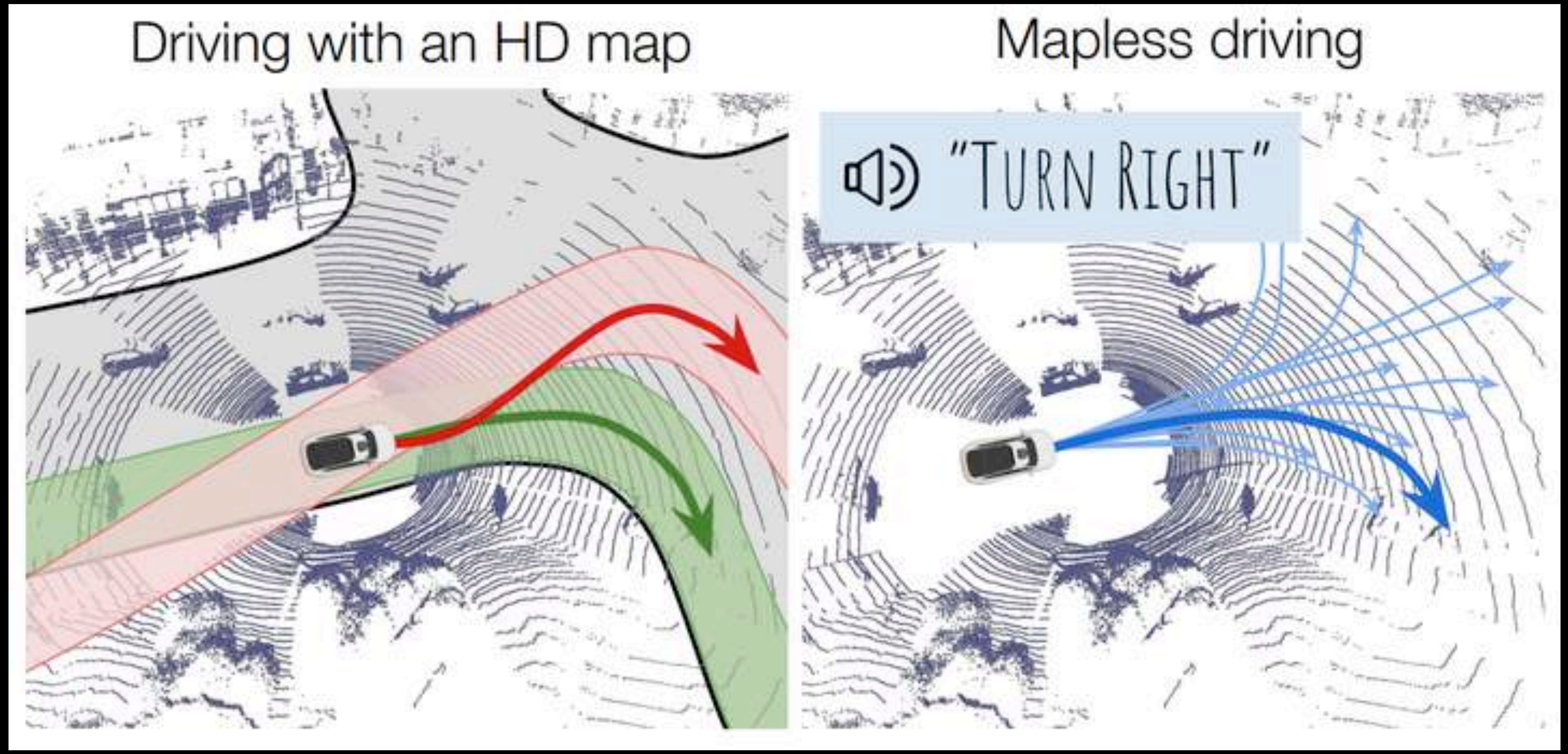
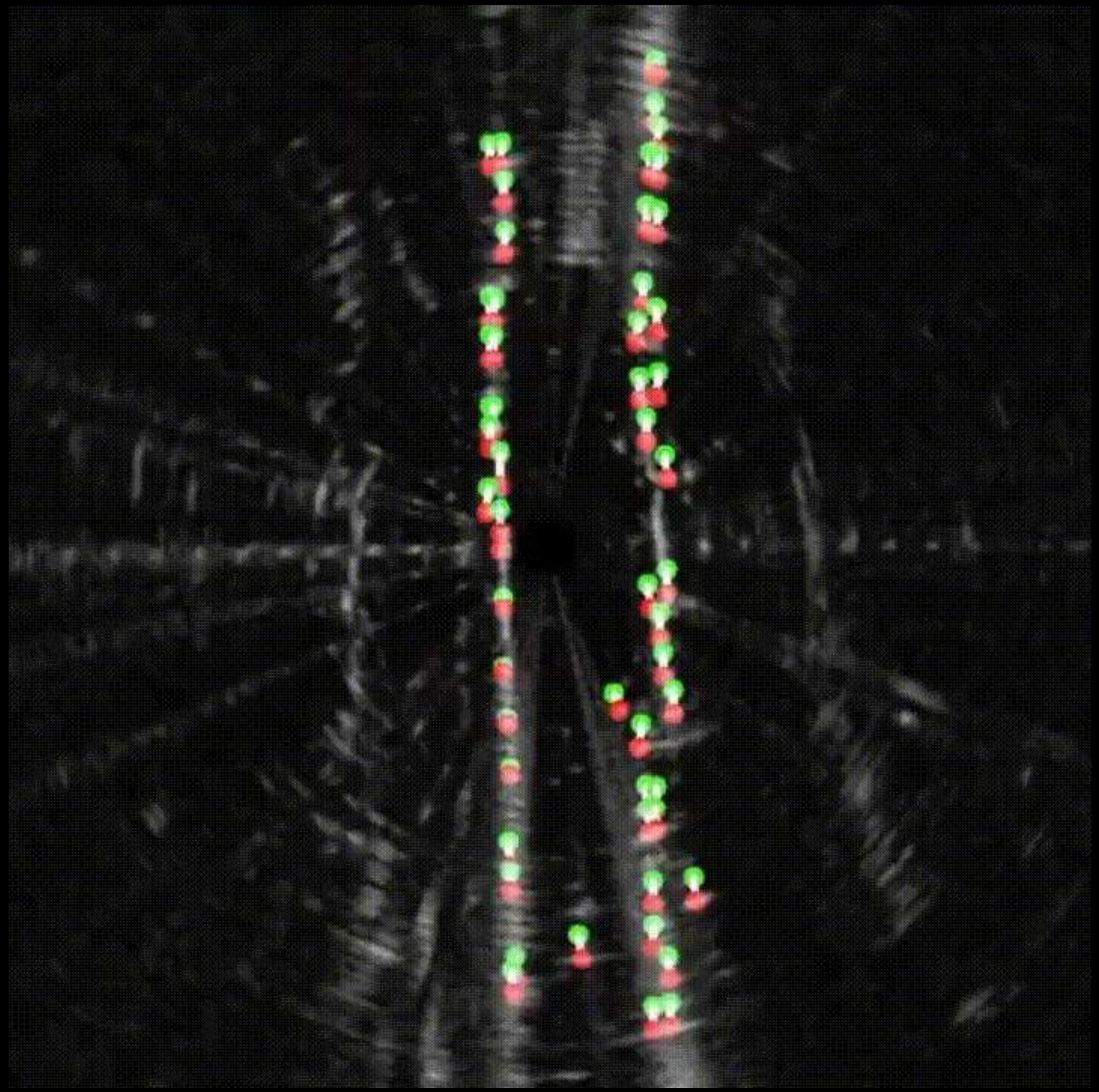


Image source: [MP3 - A Unified Model to Map, Perceive, Predict and Plan]
by Casas, Sadat, and Urtasun (CVPR 2021)

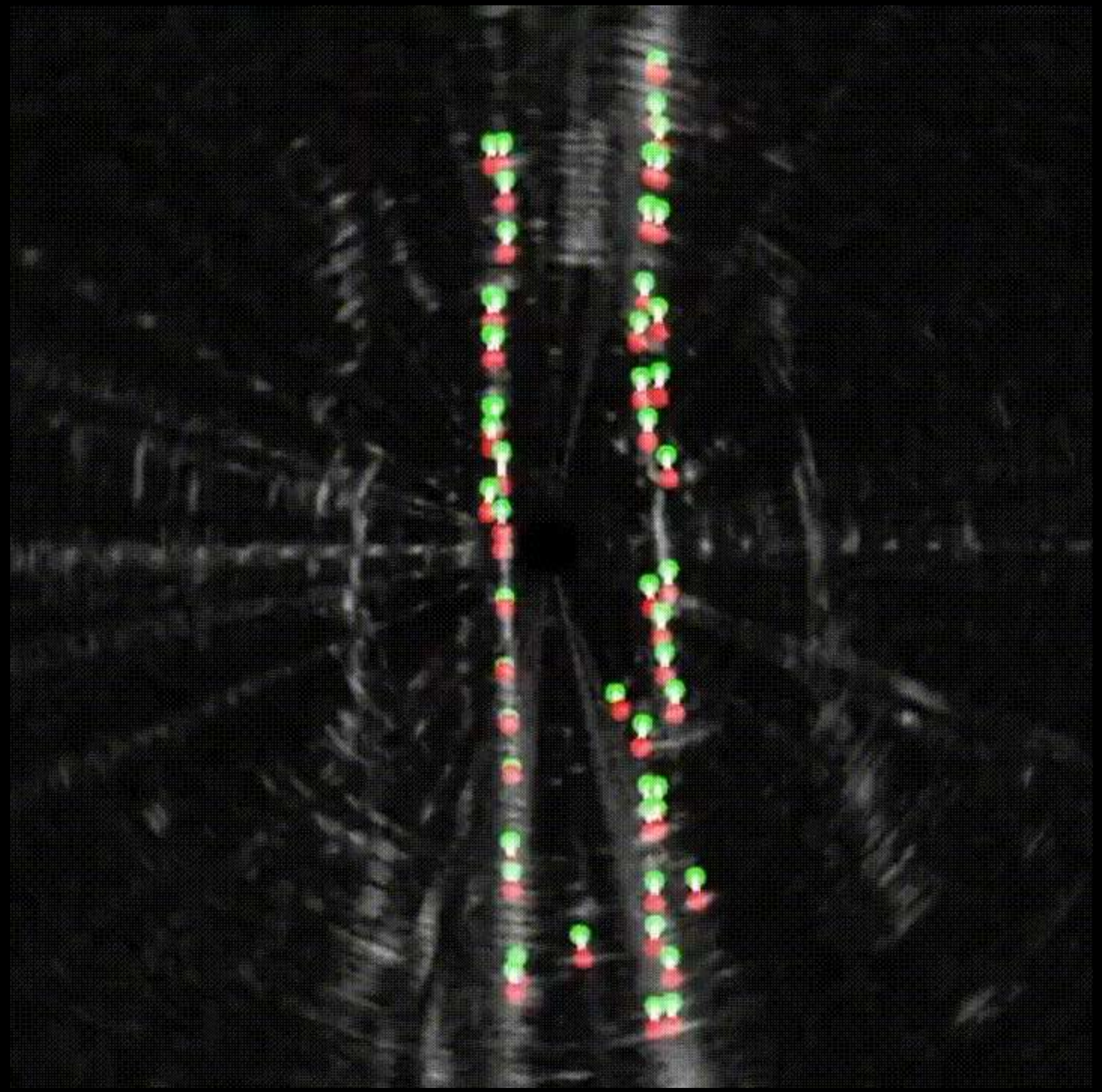
New Sensors and Infrastructure

New Sensors and Infrastructure



Imaging RADAR for Maps,
Localization & Perception
Image credit: Barnes & Posner, 2020
(Oxford RobotCar RADAR)

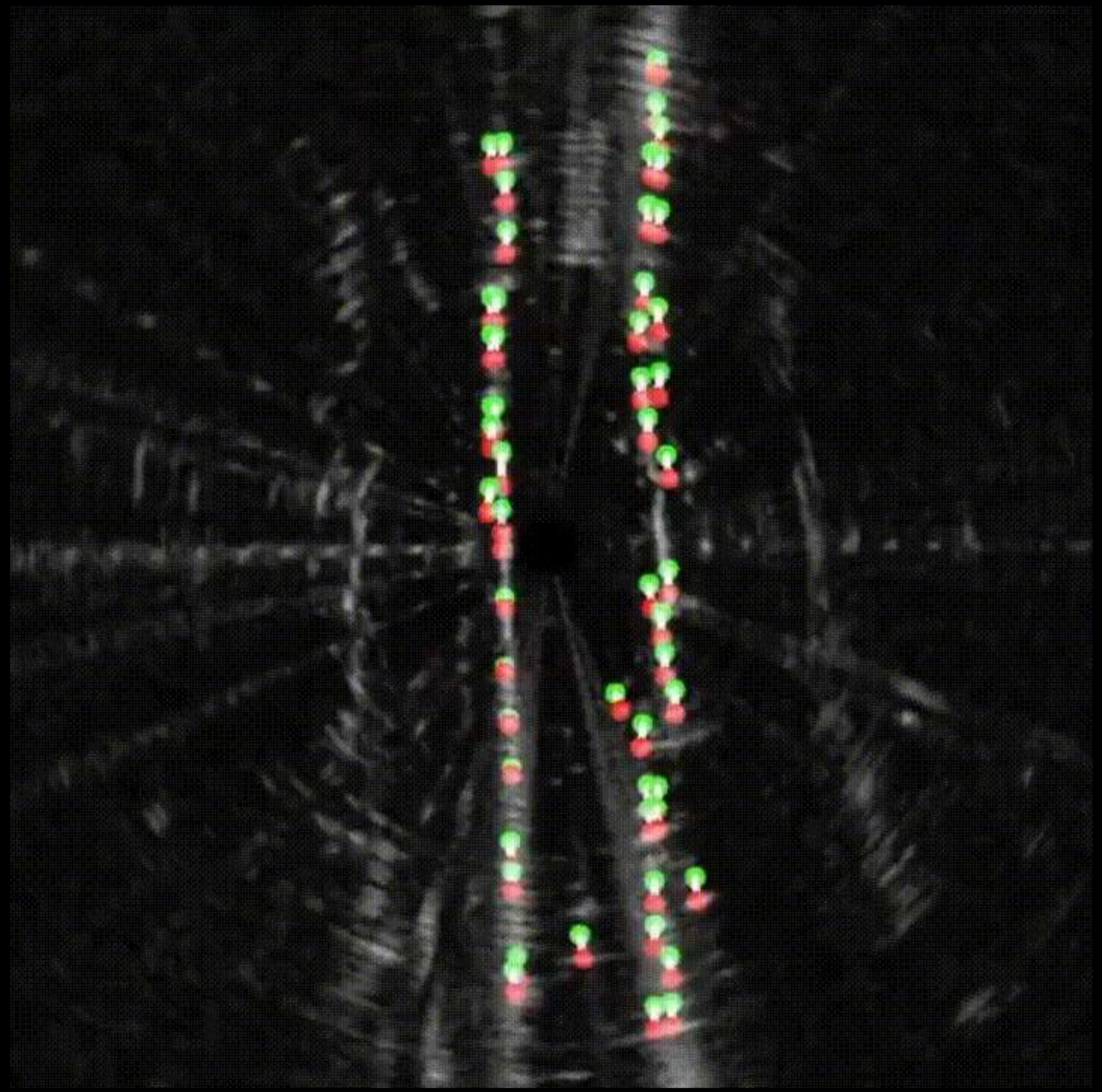
New Sensors and Infrastructure



Imaging RADAR for Maps,
Localization & Perception
Image credit: Barnes & Posner, 2020
(Oxford RobotCar RADAR)

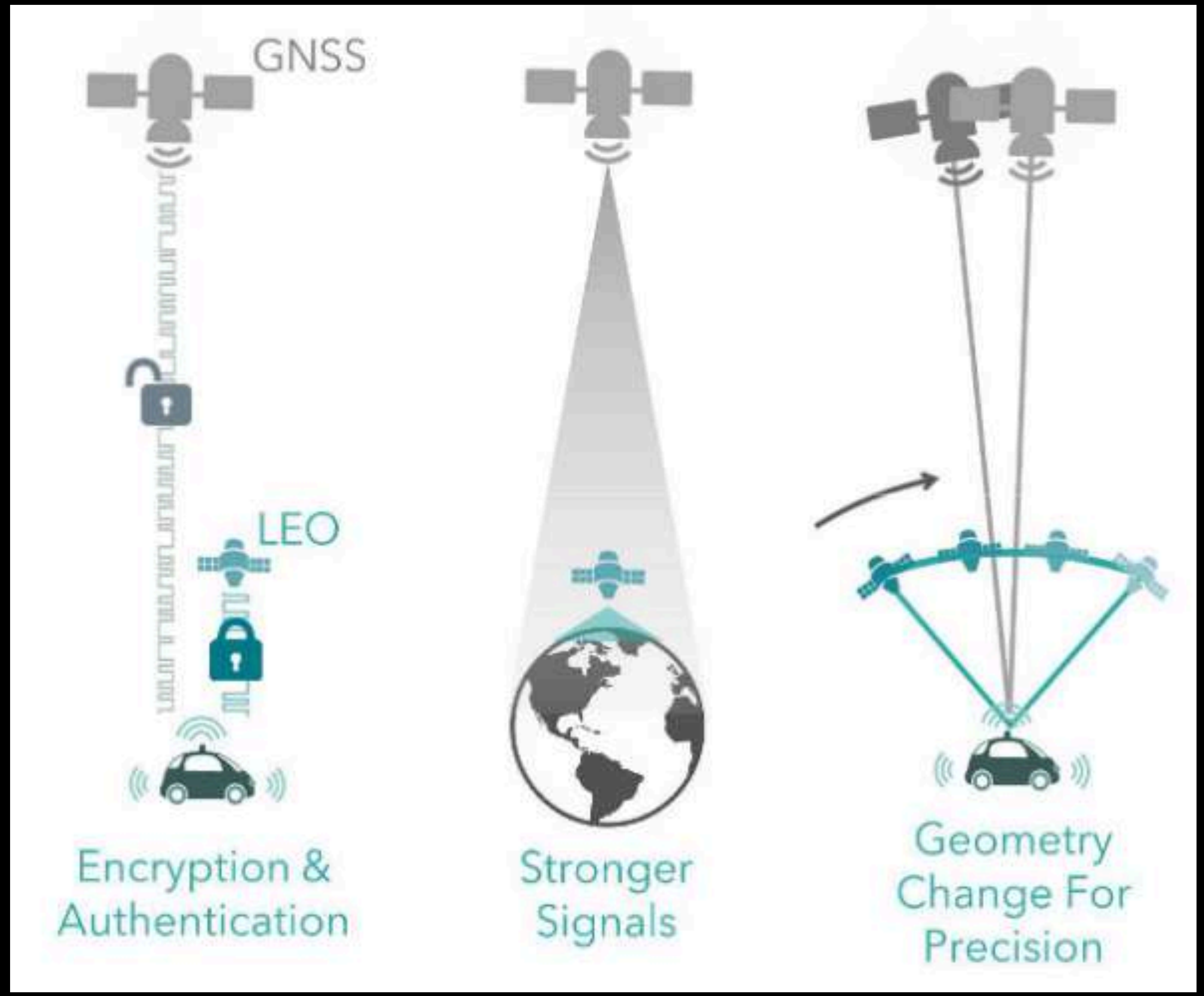
Doppler LiDAR
3D points + velocity
Blackmore, Aeva Inc.
Image credit: Blackmore

New Sensors and Infrastructure



Imaging RADAR for Maps,
Localization & Perception
Image credit: Barnes & Posner, 2020
(Oxford RobotCar RADAR)

Doppler LiDAR
3D points + velocity
Blackmore, Aeva Inc.
Image credit: Blackmore



Microsat Constellation GNSS
Next-gen GPS with cubesats
Image credit: Xona Space Systems

Conclusions

Conclusions

1. Autonomous robots are a huge field — cars are just one aspect

Conclusions

1. Autonomous robots are a huge field — cars are just one aspect
2. Huge potential for saving lives and reducing logistics costs

Conclusions

1. Autonomous robots are a huge field — cars are just one aspect
2. Huge potential for saving lives and reducing logistics costs
3. HD Maps can empower robust autonomy — maps as “another sensor”

Conclusions

1. Autonomous robots are a huge field — cars are just one aspect
2. Huge potential for saving lives and reducing logistics costs
3. HD Maps can empower robust autonomy — maps as “another sensor”
4. LiDAR localization benefits from learning

Conclusions

1. Autonomous robots are a huge field — cars are just one aspect
2. Huge potential for saving lives and reducing logistics costs
3. HD Maps can empower robust autonomy — maps as “another sensor”
4. LiDAR localization benefits from learning
5. Localization errors can affect perception & plan. but **can** be corrected

Conclusions

1. Autonomous robots are a huge field — cars are just one aspect
2. Huge potential for saving lives and reducing logistics costs
3. HD Maps can empower robust autonomy — maps as “another sensor”
4. LiDAR localization benefits from learning
5. Localization errors can affect perception & plan. but **can** be corrected
6. Multi-task learning can simplify training and deployment

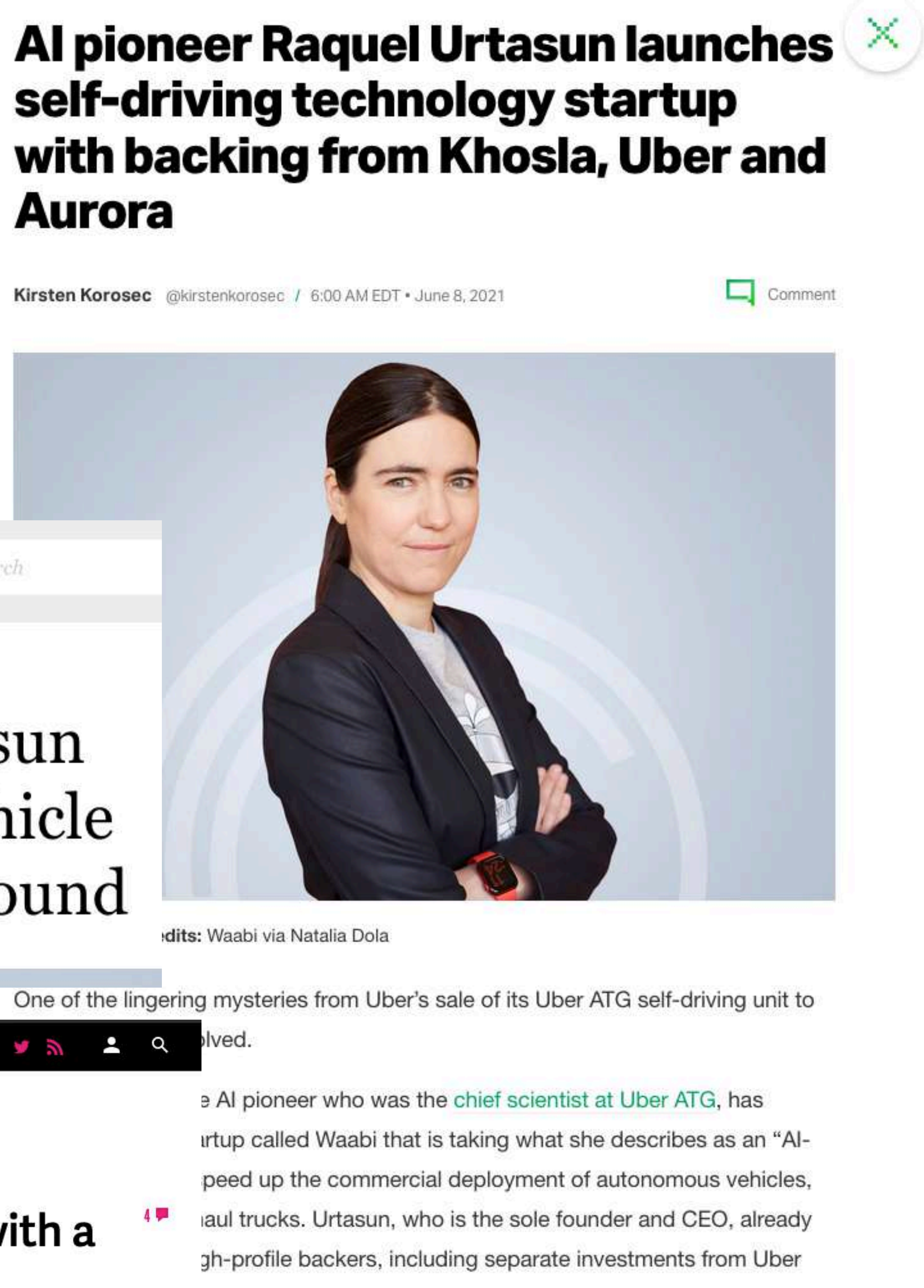
Conclusions

1. Autonomous robots are a huge field — cars are just one aspect
2. Huge potential for saving lives and reducing logistics costs
3. HD Maps can empower robust autonomy — maps as “another sensor”
4. LiDAR localization benefits from learning
5. Localization errors can affect perception & plan. but **can** be corrected
6. Multi-task learning can simplify training and deployment
7. New sensors and infrastructure can accelerate autonomy rollout

Come work with us @ waabi

#ad

- waabi.ai — out of stealth **~24h ago!**
- Work to solve self-driving at scale!
- Research & Innovation DNA
- US$83.5M Series A
- https://jobs.lever.co/waabi



AI pioneer Raquel Urtasun launches self-driving technology startup with backing from Khosla, Uber and Aurora

Kirsten Korosec @kirstenkorosec / 6:00 AM EDT • June 8, 2021

Comment

edits: Waabi via Natalia Dola

One of the lingering mysteries from Uber's sale of its Uber ATG self-driving unit to ...lved.

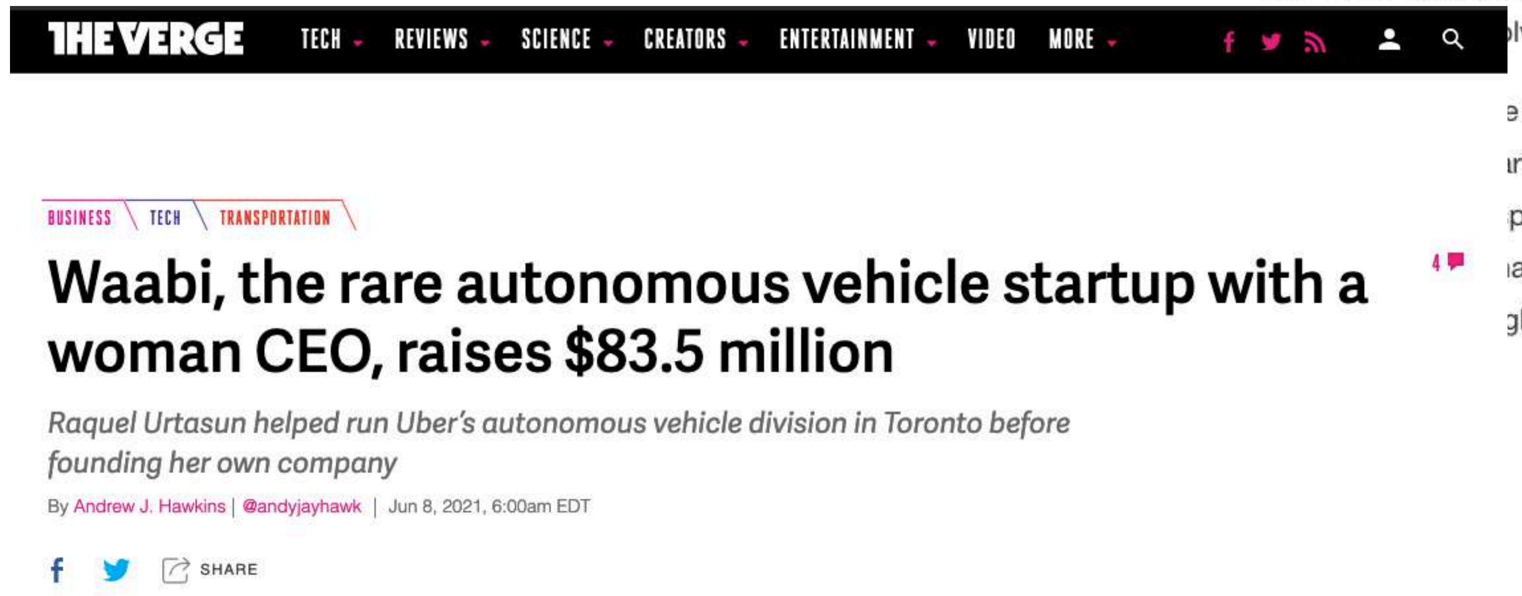
...e AI pioneer who was the chief scientist at Uber ATG, has ...rtup called Waabi that is taking what she describes as an "AI- ...peed up the commercial deployment of autonomous vehicles, ...aul trucks. Urtasun, who is the sole founder and CEO, already ...gh-profile backers, including separate investments from Uber

The Logic IN-DEPTH REPORTING ON THE INNOVATION ECONOMY

Search

News

AI superstar Raquel Urtasun launches autonomous-vehicle startup with US$83.5M round



THE VERGE TECH REVIEWS SCIENCE CREATORS ENTERTAINMENT VIDEO MORE

BUSINESS TECH TRANSPORTATION

Waabi, the rare autonomous vehicle startup with a woman CEO, raises $83.5 million

Raquel Urtasun helped run Uber's autonomous vehicle division in Toronto before founding her own company

By Andrew J. Hawkins | @andyjayhawk | Jun 8, 2021, 6:00am EDT

SHARE

Global NEWS World Canada Local Politics Money Health Entertainment Lifes

CANADA

Raquel Urtasun starts Toronto self-driving company Waabi, after leaving Uber

By Tara Deschamps · The Canadian Press

Posted June 8, 2021 6:43 am · Updated June 8, 2021 10:42 am

Forbes

EDITORS' PICK | Jun 8, 2021, 06:00am EDT | 995 views

Uber Veteran Launches Her 'AI Mindset' Self-Driving Startup With $83.5 Million Round

Alan Ohnsman Forbes Staff

Transportation

I follow technology-driven changes reshaping transportation.

Thank you!

See you later in the networking area if you want to chat!

Andrei Bârsan — andreibarsan.github.io — @andreib

References

- Resources
  - andreibarsan.github.io for the main highlighted papers
  - All About Self-Driving CVPR 2020 Tutorial (I'll be contributing to the updated 2021 version at CVPR in 1.5 weeks!!)
- Papers & Websites
  - Introduction:
    - US Road Deaths (NHTSA for Death Count, this Stanford Law Report for 90%+ human error estimate)
    - IntentNet (Casas et al., CoRL '18)
  - Scalable LiDAR Localization:
    - Map-based precision vehicle localization in urban environments (Levinson, Montemerlo & Thrun, RSS '07)
    - Learning to Localize using a LiDAR Intensity Map (Barsan, CoRL '18)
    - Learning to Localize through Compressed Binary Maps (Wei, CVPR '19) (Also contains sources for how to estimate the storage for the US road network.)
  - How Good Does Localization Need to Be?
    - The Implicit Latent Variable Model for Scene-Consistent Motion Forecasting (Casas et al., ECCV '20)
    - Deep Multi-Task Learning for Joint Localization, Perception, and Prediction (Phillips et al., CVPR '21)
  - Future:
    - Cen and Newman (ICRA '18 — one of the first modern RADAR localization papers from Oxford), https://dbarnes.github.io/ (Dan Barnes's papers for RADAR Localization)
    - https://www.aeva.ai/ (for Doppler LiDAR)
    - Satellite Navigation for the Age of Autonomy (Reid et al., '20 — Xona Space Systems)
    - MP3 (Casas, Sadat, and Urtasun CVPR '21 — for mapless driving)