Lex Fridman

Lecture 2:

Self-Driving Cars

Self-Driving Cars

(aka driverless cars, autonomous cars, robocars)

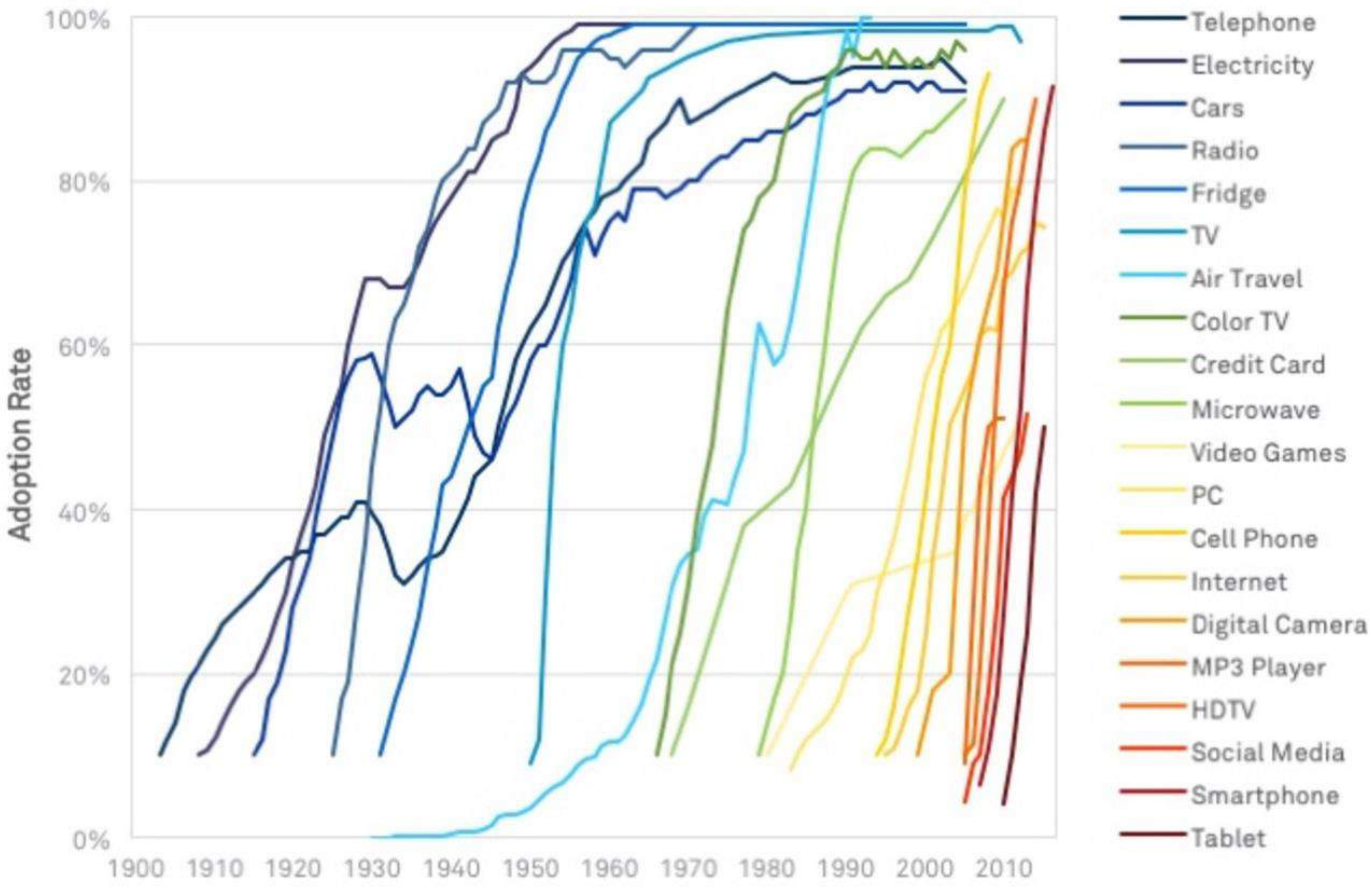
- Utopian view
  - Save lives (1.3 million die every year in manual driving)
    - 4D's of human folly: drunk, drugged, distracted, drowsy driving
  - Eliminate car ownership
    - Increase mobility and access
    - Save money
  - Make transportation personalized, efficient, and reliable

- Dystopian view
  - Eliminate jobs in the transportation sector
  - Failure (even if much rarer) may not depend on factors that are human interpretable or under human control
  - Artificial intelligence systems may be biased in ways that do not coincide with social norms or be **ethically grounded**
  - Security

Self-Driving Cars: Grain of Salt

- Our intuition about what is hard or easy for AI is flawed *(see first lecture)*
- Carefully differentiate between:
  - **Doubtful:** Promises for future vehicles (in 2+ years)
  - **Skeptical:** Promises for future vehicles (in 1 year)
  - **Possible:** Actively testing vehicles on public roads at scale
  - **Real:** Available for consumer purchase today
- Rodney brooks prediction in "My Dated Predictions":
  - **>2032:** A driverless "taxi" service in a major US city with arbitrary pick and drop off locations, even in a restricted geographical area.
  - **>2045:** The majority of US cities have the majority of their downtown under such rules.

Self-Driving Cars

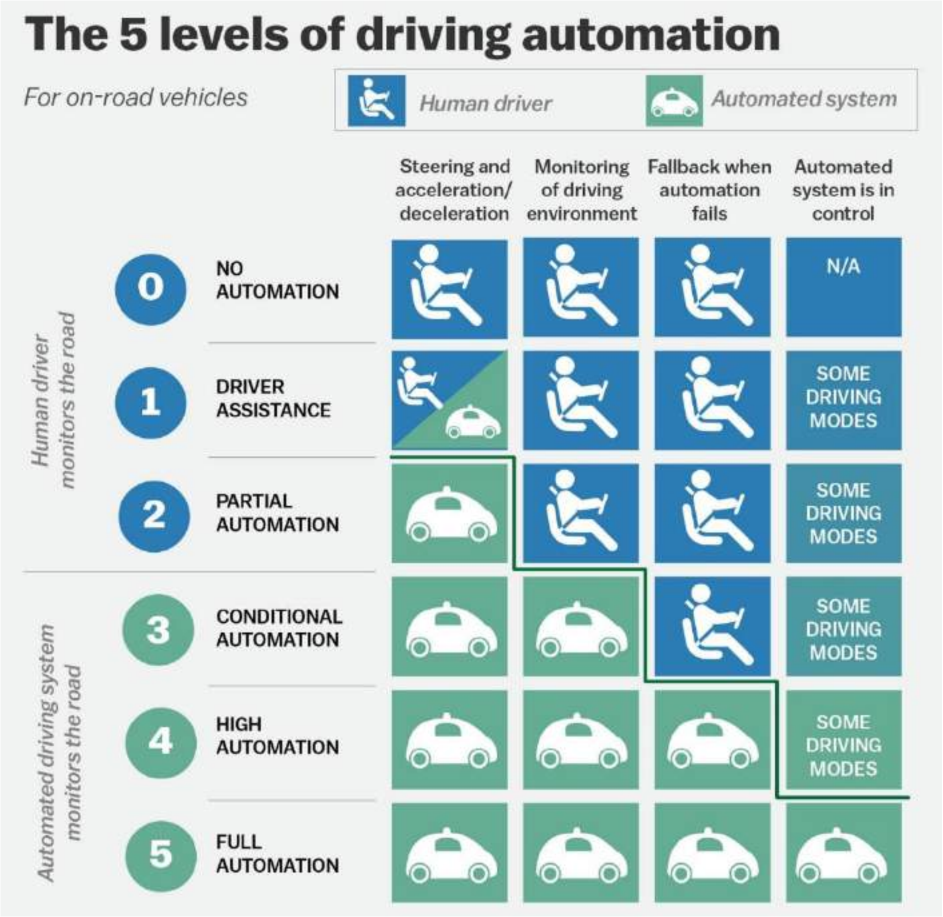


Overview

- **Different approaches autonomy**
- Sensors
- Companies doing it
- Opportunities for AI and deep learning

Levels of Automation (SAE J3016)

- Useful for initial discussion (especially for policy making), but not useful for design and engineering of the underlying intelligence and the holistic system performance:



Beyond Traditional Levels: **Two** AI Systems

- **Starting point:**
  - All cars are manually controlled until the AI system shows itself to be available and is elected to be turned on by the human.
- **A1:** Human-Centered Autonomy
  - **Definition:** AI is not fully responsible
  - Feature axis:
    - Where/how often is it "available"? (traffic, highway, sensor-based, etc.)
    - How many seconds for take-over? (0, 1, 10, etc)
    - Teleoperation support
- **A2:** Full Autonomy
  - **Definition:** AI is fully responsible
  - Notes:
    - No teleoperation
    - No 10-second rule: It's allowed to ask for human help, but not guaranteed to ever receive it.
    - Arrive to a **safe** destination or safe harbor.
    - Allow the human to take over **when they choose to**.

Beyond Traditional Levels: **Two** AI Systems

- **Starting point:**
  - All cars are manually controlled until the AI system shows itself to be available and is elected to be turned on by the human.
- **A1:** Human-Centered Autonomy
  - **Definition:** AI is not fully responsible
  - Feature axis:
    - Where/how often is it "available"? (traffic, highway, sensor-based, etc.)
    - How many seconds for take-over? (0, 1, 10, etc)
    - Teleoperation support
- **A2:** Full Autonomy
  - **Definition:** AI is fully responsible
  - Notes:
    - No teleoperation
    - No 10-second rule: It's allowed to ask for human help, but not guaranteed to ever receive it.
    - Arrive to a **safe** destination or safe harbor.
    - Allow the human to take over **when they choose to**.

Beyond Traditional Levels: **Two** AI Systems

L0 →
- **Starting point:**
  - All cars are manually controlled until the AI system shows itself to be available and is elected to be turned on by the human.

L1, L2, L3 →
- **A1:** Human-Centered Autonomy
  - **Definition:** AI is not fully responsible

L4, L5 →
- **A2:** Full Autonomy
  - **Definition:** AI is fully responsible

Two AI Systems:

A2: Full Autonomy

Two AI Systems:

A1: Human-Centered Autonomy

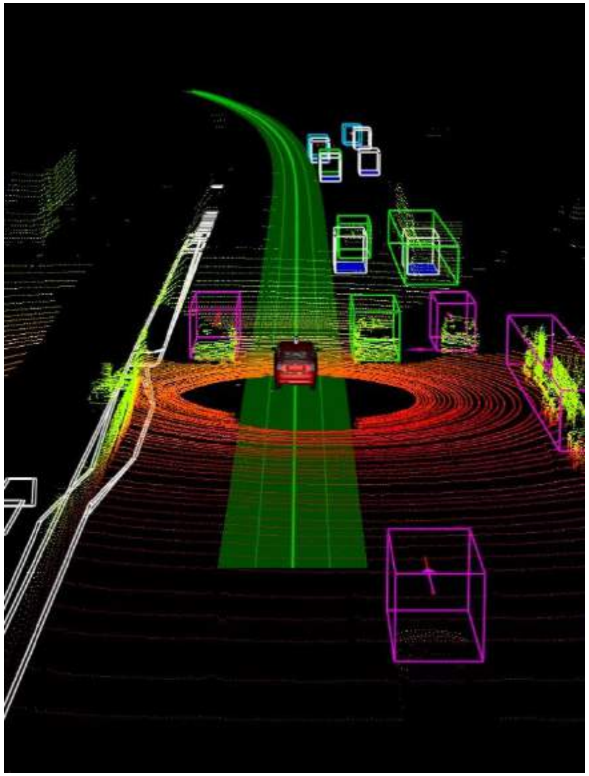
Two Paths to an Autonomous Future

A1:
Human-Centered Autonomy

- **Localization and Mapping:** Where am I?
- **Scene Understanding:** Where/who/what/why of everyone else?
- **Movement Planning:** How do I get from A to B?
- **Human-Robot Interaction:** What is the physical and mental state of the driver?
- **Communicate:** How to I convey intent to the driver and to the world?

Blue Text: Easier

Red Text: Harder



A2:
Full Autonomy

- **Localization and Mapping:** Where am I?
- **Scene Understanding:** Where/who/what/why of everyone else?
- **Movement Planning:** How do I get from A to B?
- **Human-Robot Interaction:** What is the physical and mental state of the driver?
- **Communicate:** How to I convey intent to the driver and to the world?

Is partially automated driving a bad idea? Observations from an on-road study

Article · April 2018 *with* 447 Reads
DOI: 10.1016/j.apergo.2017.11.010

Victoria Banks
14.44 · University of Southampton

Alexander Eriksson
11.13 · Swedish National Road and Transport Research Inst...

Jim O'donoghue

Neville A Stanton
43.23 · University of Southampton

WAYMO

Chris Urmson

Public Perception of What Drivers Do in Semi-Autonomous Vehicles

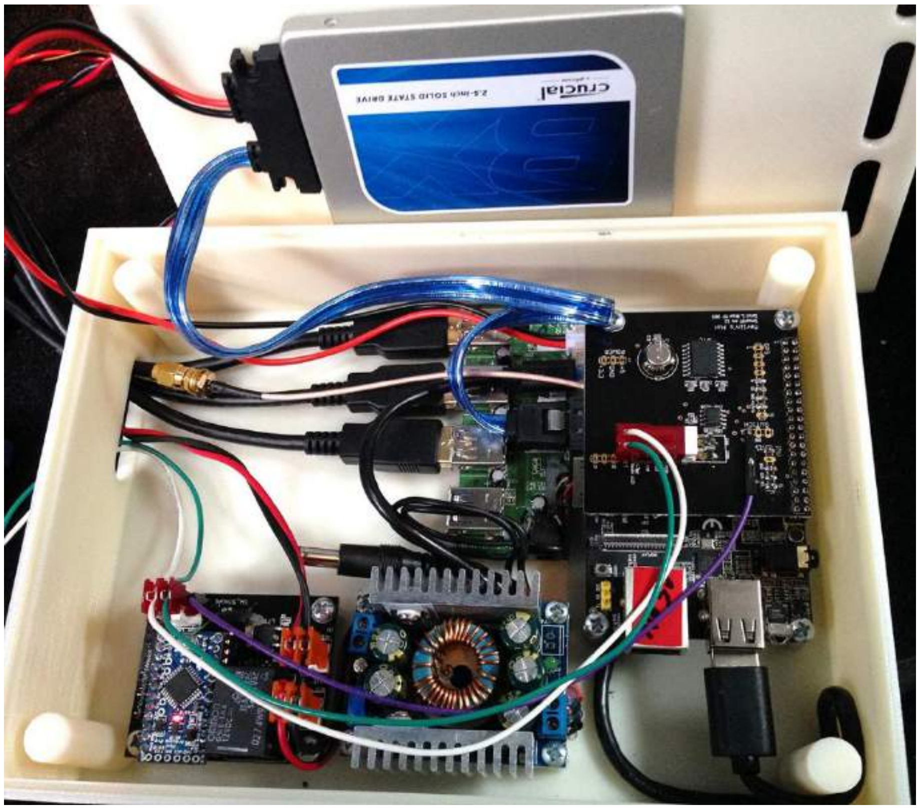
MIT-AVT Naturalistic Driving Dataset

Vehicles instrumented: 25

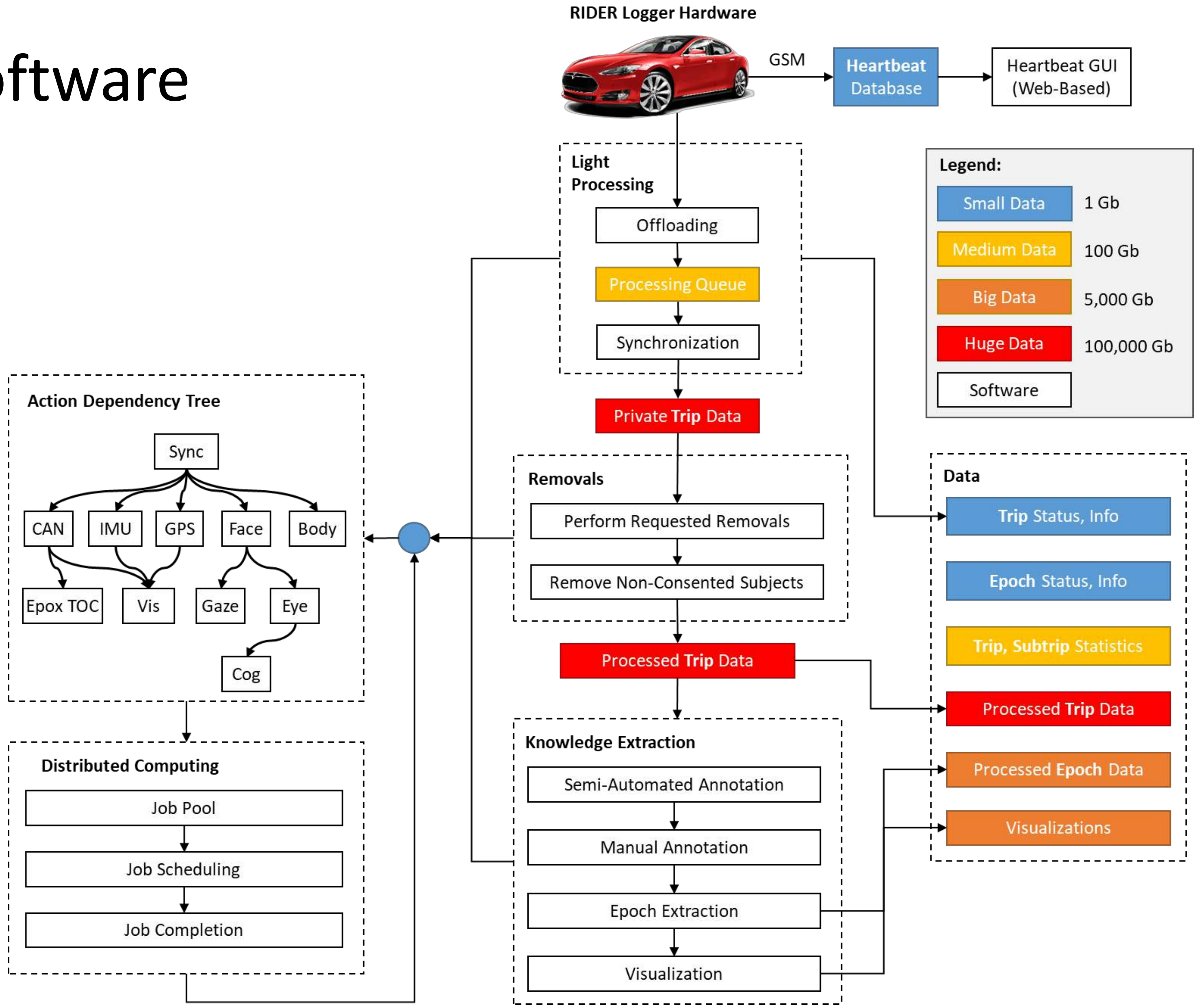
Distance traveled: 275,000+ miles

Video frames: 4.7+ billion

Hardware



Software



RIDER Logger Hardware
GSM
Heartbeat Database
Heartbeat GUI (Web-Based)
Legend:
Small Data
1 Gb
Medium Data
100 Gb
Big Data
5,000 Gb
Huge Data
100,000 Gb
Software
Light Processing
Offloading
Processing Queue
Synchronization
Private Trip Data
Removals
Perform Requested Removals
Remove Non-Consented Subjects
Processed Trip Data
Knowledge Extraction
Semi-Automated Annotation
Manual Annotation
Epoch Extraction
Visualization
Action Dependency Tree
Sync
CAN
IMU
GPS
Face
Body
Epox TOC
Vis
Gaze
Eye
Cog
Distributed Computing
Job Pool
Job Scheduling
Job Completion
Data
Trip Status, Info
Epoch Status, Info
Trip, Subtrip Statistics
Processed Trip Data
Processed Epoch Data
Visualizations


Autopilot: Unavailable
Total Time Driving: 0 mins
Autopilot Available: 0 mins
Autopilot Engaged: 0 mins

Human Behavior

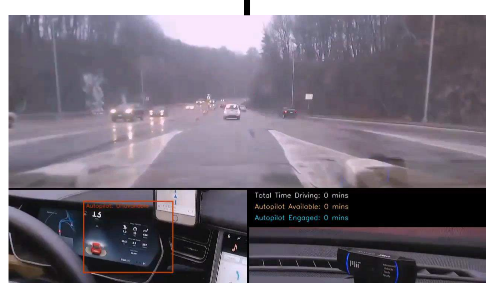
Shared Autonomy

Understand Behavior

Assist Behavior

Share Control

Semi-Supervised Learning



Large-Scale Naturalistic Data

MIT-AVT Naturalistic Driving Dataset

MIT Autonomous Vehicle Technology Study

Study months to-date: **21**
Participant days: **7,146**
Drivers: **78**
Vehicles: **25**
Miles driven: **275,589**
Video frames: **3.48 billion**

Study data collection is ongoing.
*Statistics updated on: **Oct 23, 2017**.*

Tesla Model S
24,657 miles
588 days in study

Tesla Model X
22,001 miles
421 days in study

Tesla Model S
18,896 miles
435 days in study

Tesla Model S
18,666 miles
353 days in study

Range Rover Evoque
18,130 miles
483 days in study

Tesla Model S
15,735 miles
322 days in study

Tesla Model X
15,074 miles
276 days in study

Range Rover Evoque
14,499 miles
440 days in study

Tesla Model S
14,410 miles
371 days in study

Tesla Model S
14,117 miles
248 days in study

Volvo S90
13,970 miles
325 days in study

Tesla Model S
12,353 miles
321 days in study

Volvo S90
11,072 miles
412 days in study

Tesla Model X
10,271 miles
366 days in study

Tesla Model S
9,188 miles
183 days in study

Tesla Model S
8,319 miles
374 days in study

Tesla Model S
6,720 miles
194 days in study

Tesla Model S
5,186 miles
91 days in study

Tesla Model X
5,111 miles
232 days in study

Tesla Model S
4,596 miles
132 days in study

Tesla Model X
4,587 miles
233 days in study

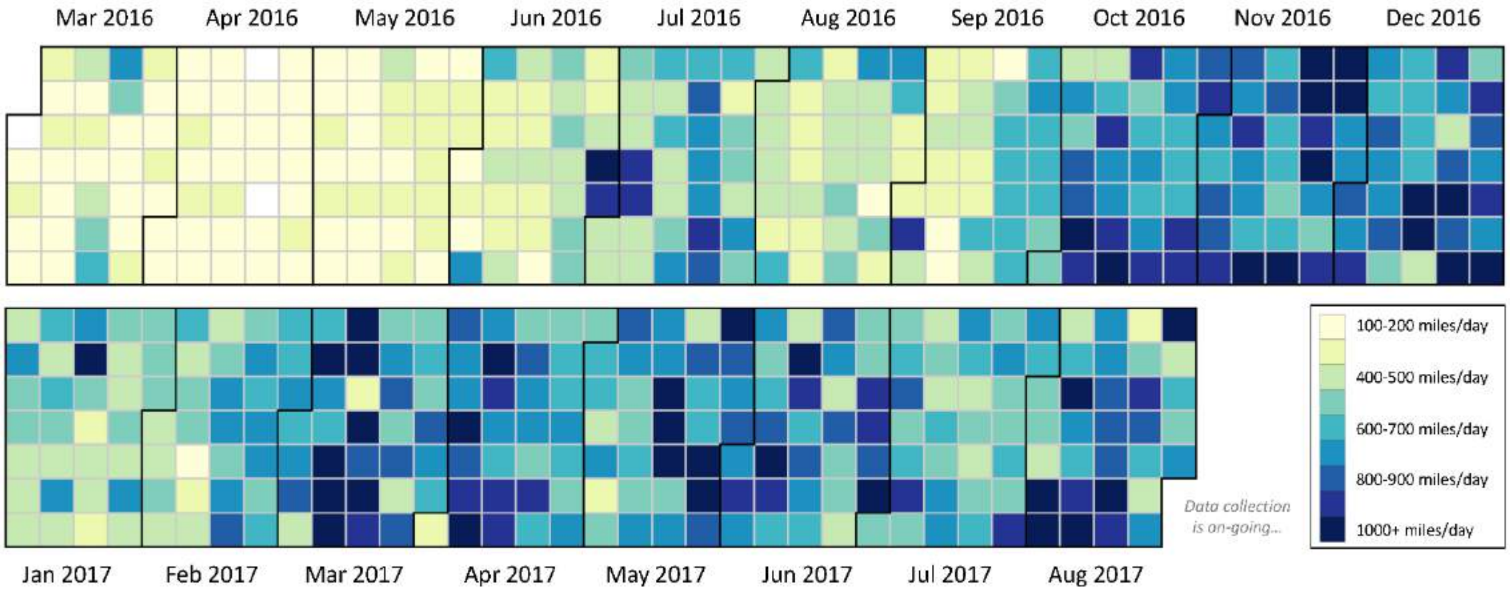
Tesla Model X
3,719 miles
133 days in study

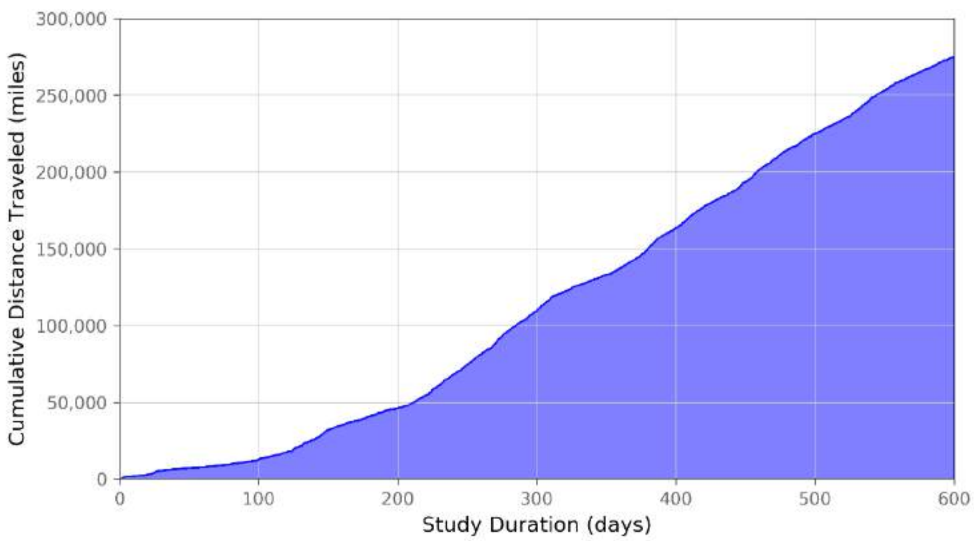
Tesla Model S
3,006 miles
144 days in study

Tesla Model X
1,306 miles
69 days in study

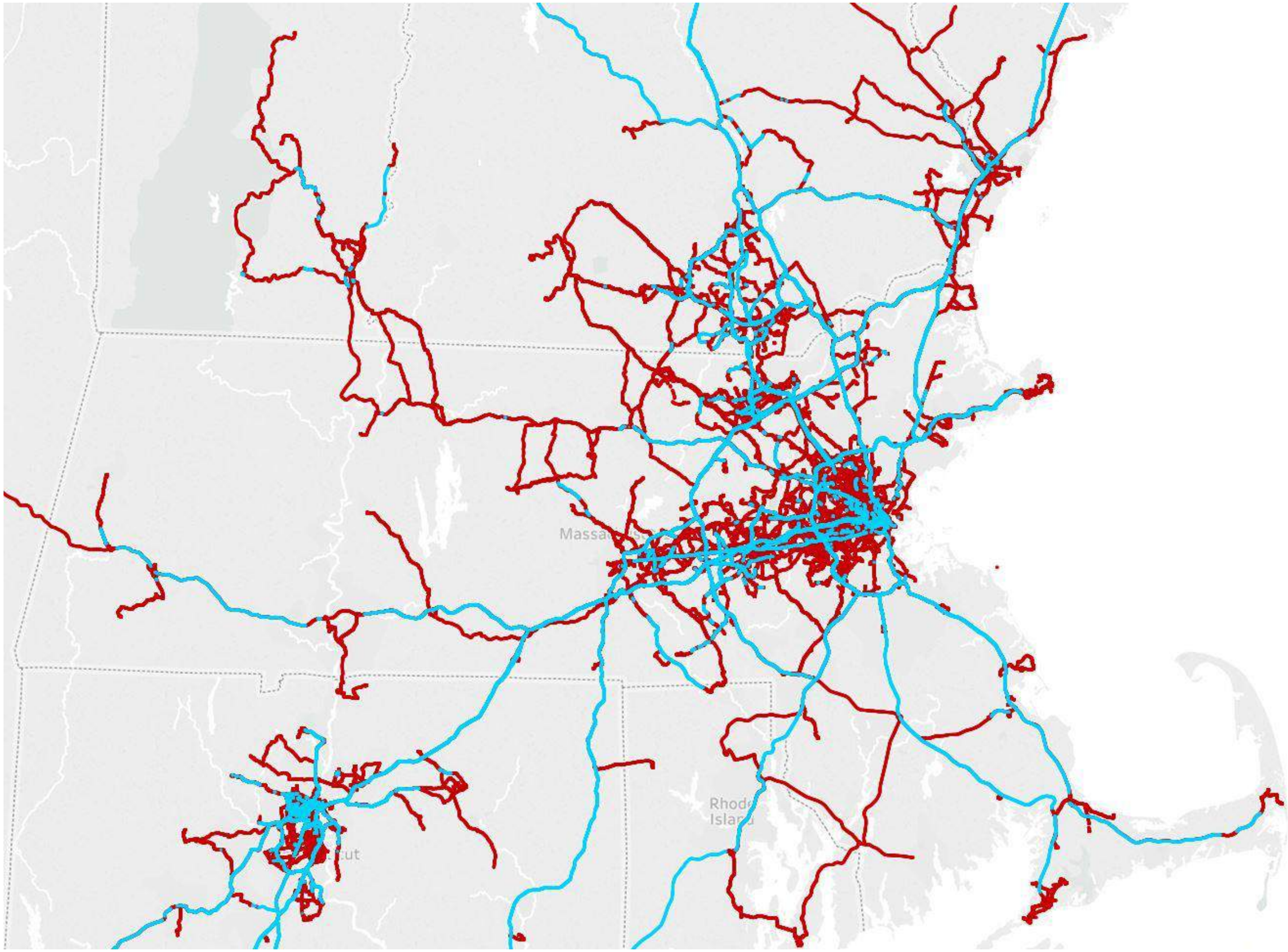
Tesla Model S
(Offload pending)

500+ Miles / Day and Growing





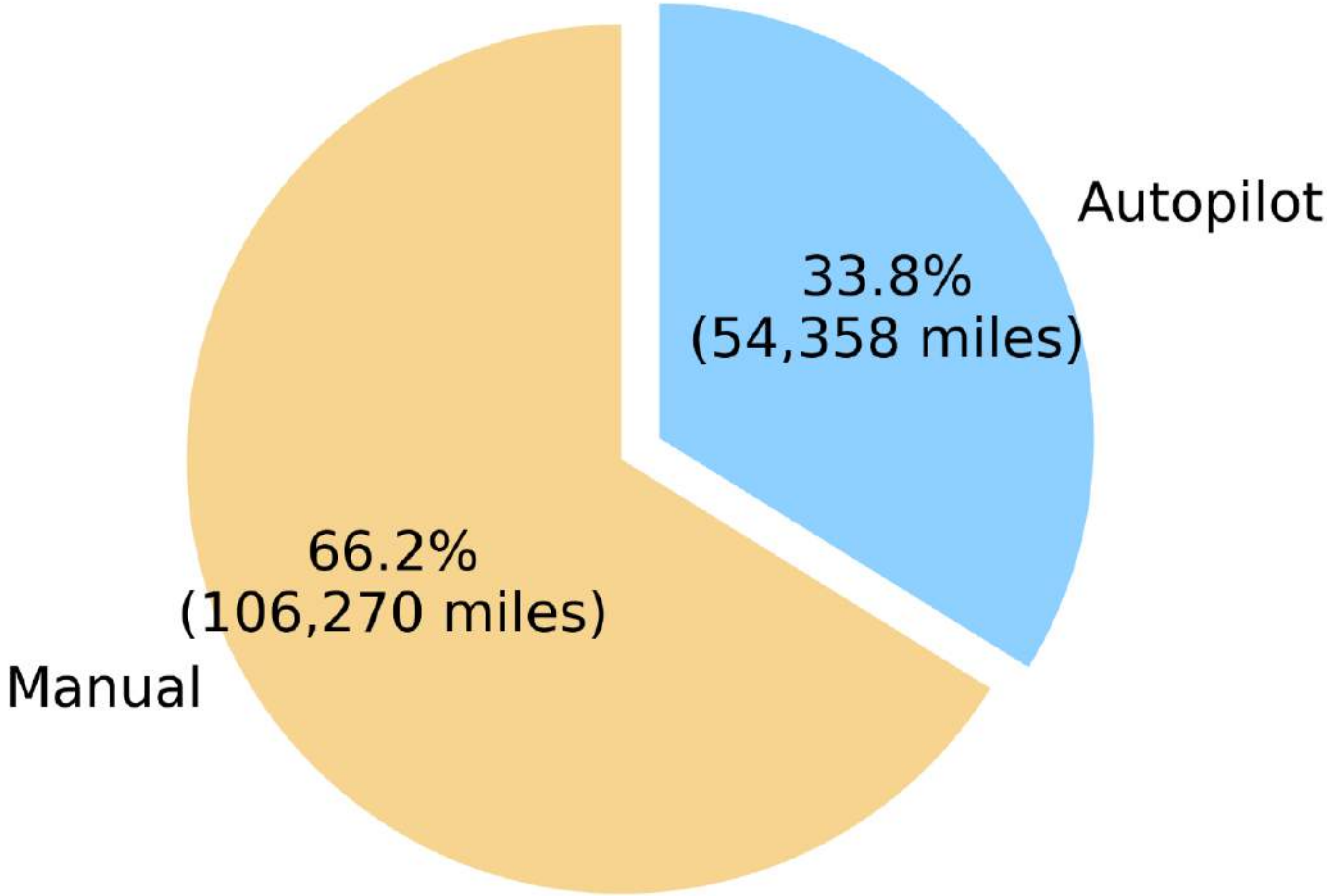
MIT-AVT Naturalistic Driving Dataset



GPS points for Autopilot control

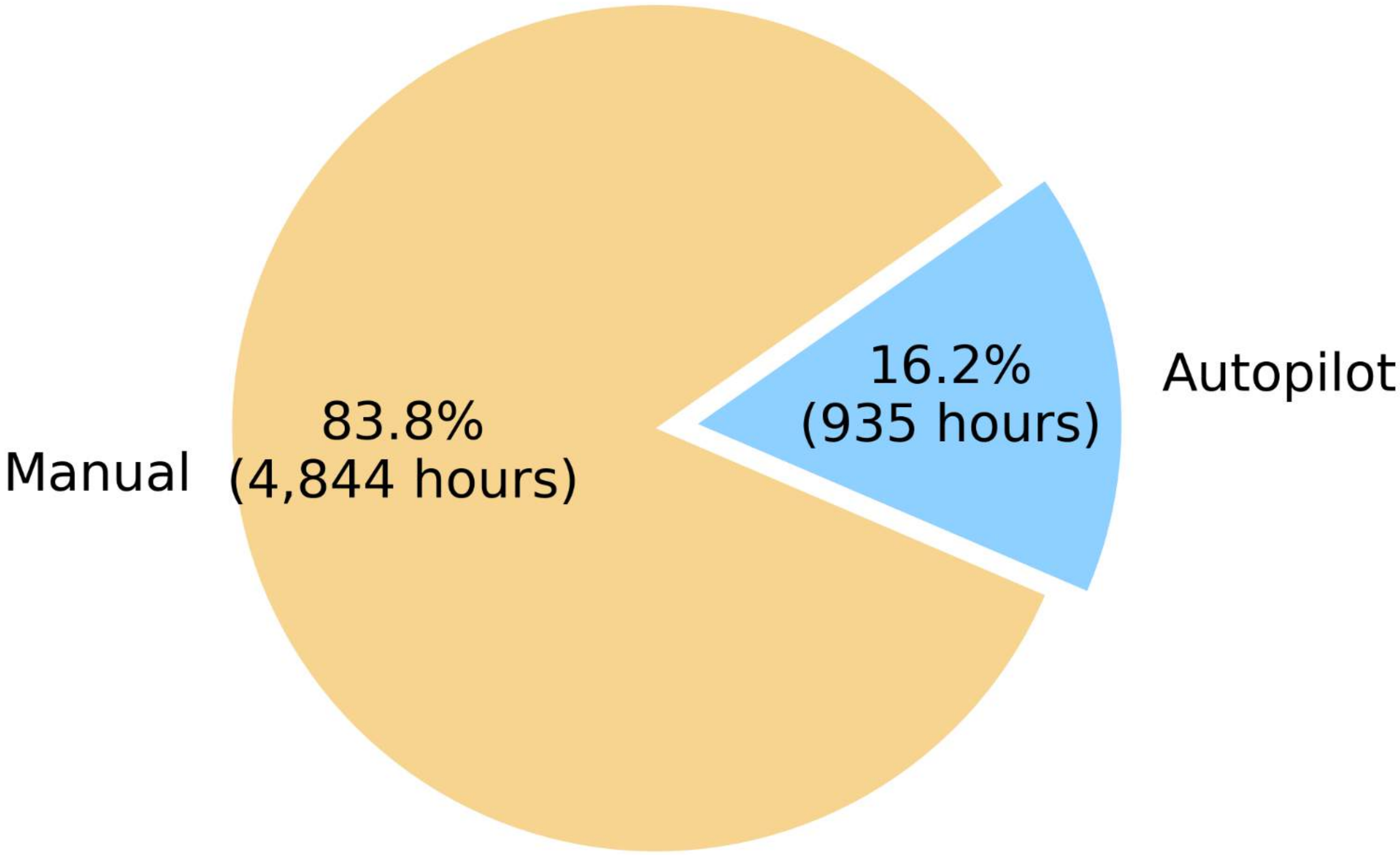
GPS points with manual control

Tesla Autopilot: Patterns of Use



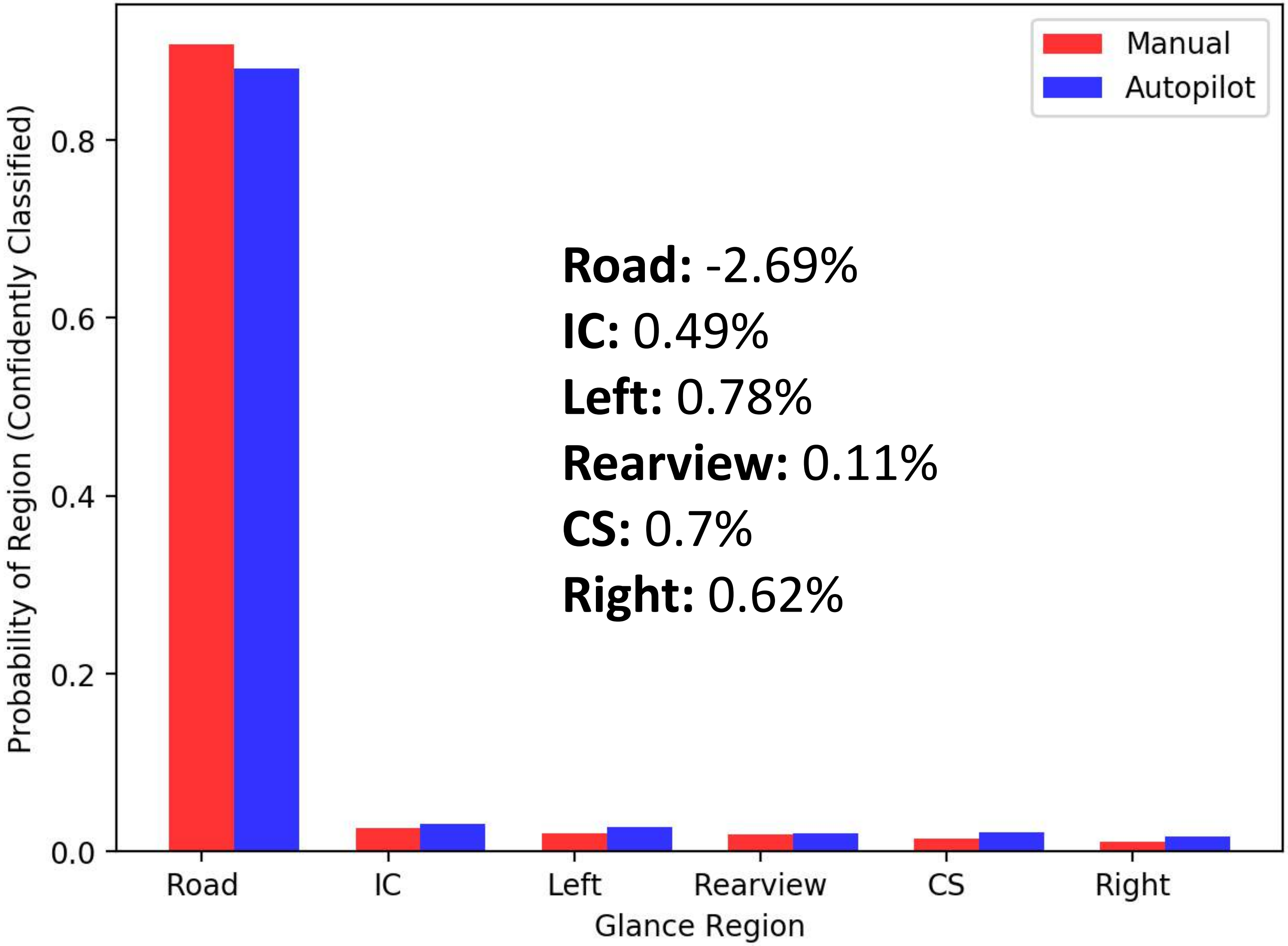
33.8% of the **miles** driven are with Autopilot engaged

Tesla Autopilot: Patterns of Use



16.2% of the **hours** driven are with Autopilot engaged

Physical Engagement: Glance Classification



Road: -2.69%
IC: 0.49%
Left: 0.78%
Rearview: 0.11%
CS: 0.7%
Right: 0.62%

Tesla Autopilot: Observed Patterns of Behavior

- **Usage:** People use autopilot a lot (% miles, % hours)
- **Road Type:** People use it on highway (using speed limit)
- **Mental Engagement:** 8,000 transfers of control from machine show that they remain vigilant to cases when Autopilot creates risk.
- **Physical Engagement:** Glance profile remains the same (% glance in manual vs autopilot by same road type)
- The "how" of successful human-robot interaction:

Use but Don't Trust.

- The "why" of successful human-robot interaction:

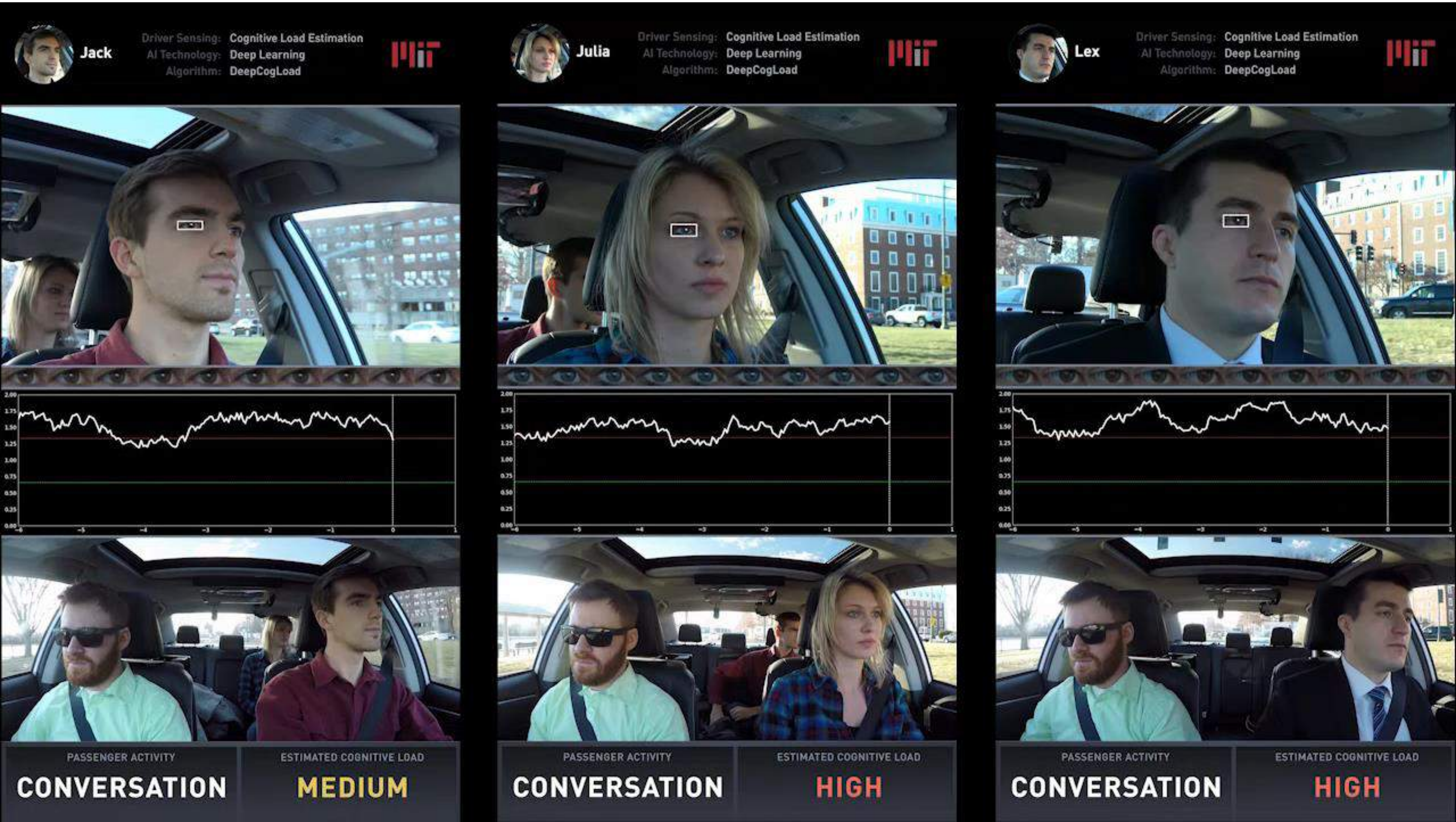
Learn Limitations by Exploring.

Self-Driving Cars: Personal Robotics View

- First wide reaching and profound integration of personal robots in society.
  - **Wide reaching:** 1 billion cars on the road.
  - **Profound:** Human gives control of his/her life directly to robot.
  - **Personal:** One-on-one relationship of communication, collaboration, understanding and trust.

A self-driving car may be more a Personal Robot and less a perfect Perception-Control system. Why:

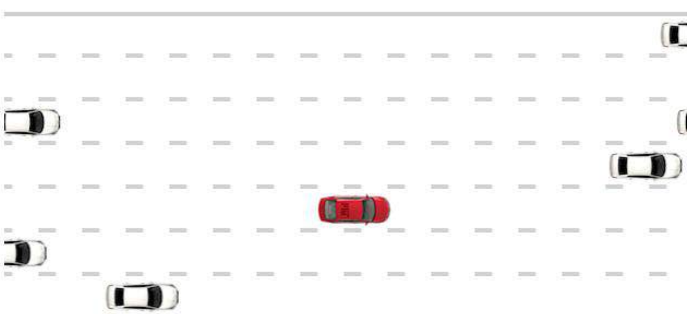
- **Flaws need humans:**
  The scene understanding problem requires much more than pixel-level labeling
- **Exist with humans:**
  Achieving both an enjoyable and safe driving experience may require "driving like a human".



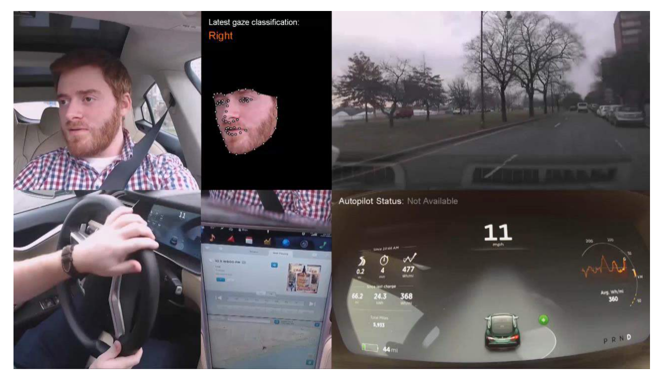
Human-Centered Artificial Intelligence Approach

90% ← No — Human Needed — Yes → 10%

Solve the perception-control problem where **possible**:



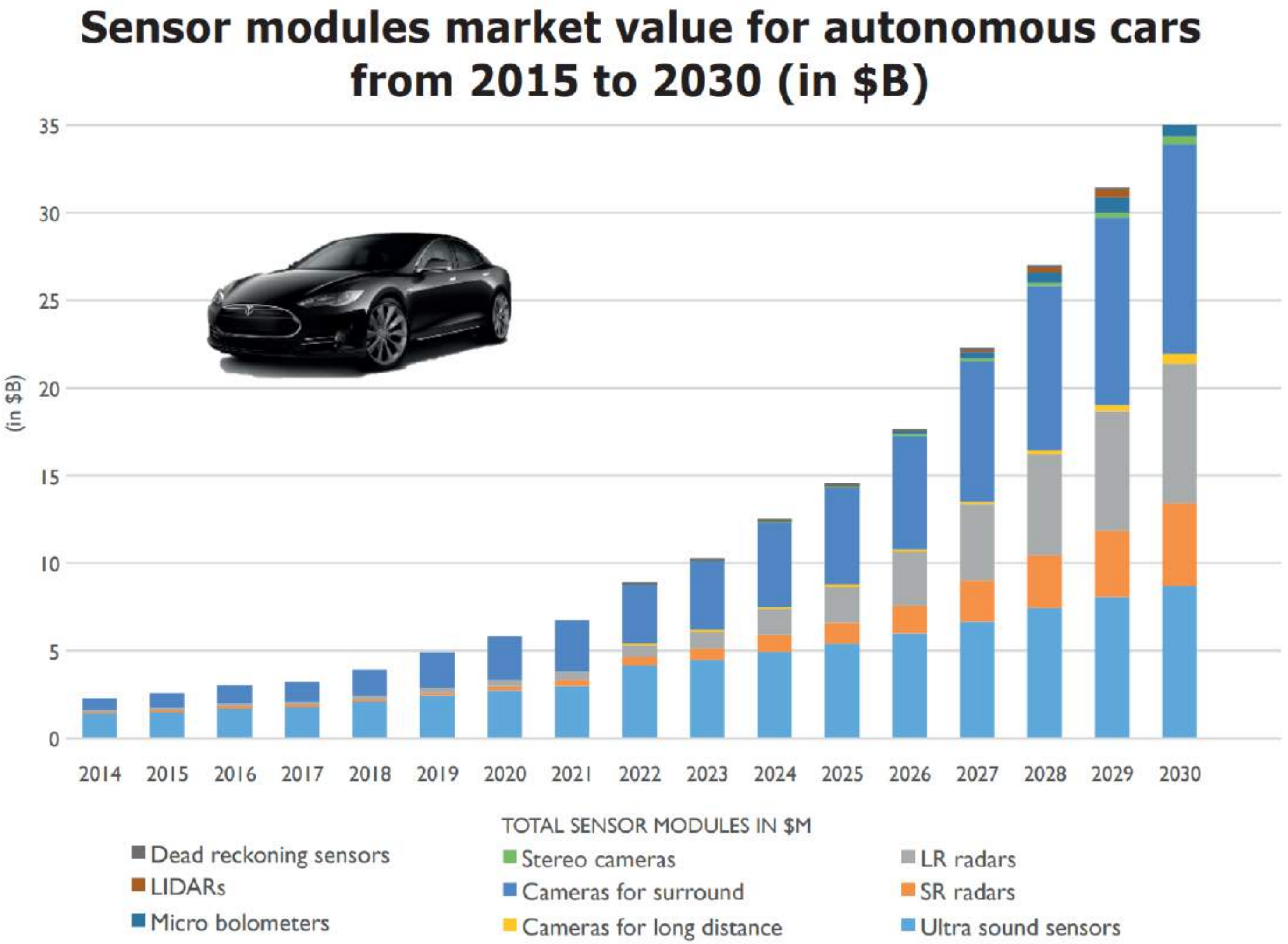
And where **not possible**: involve the human



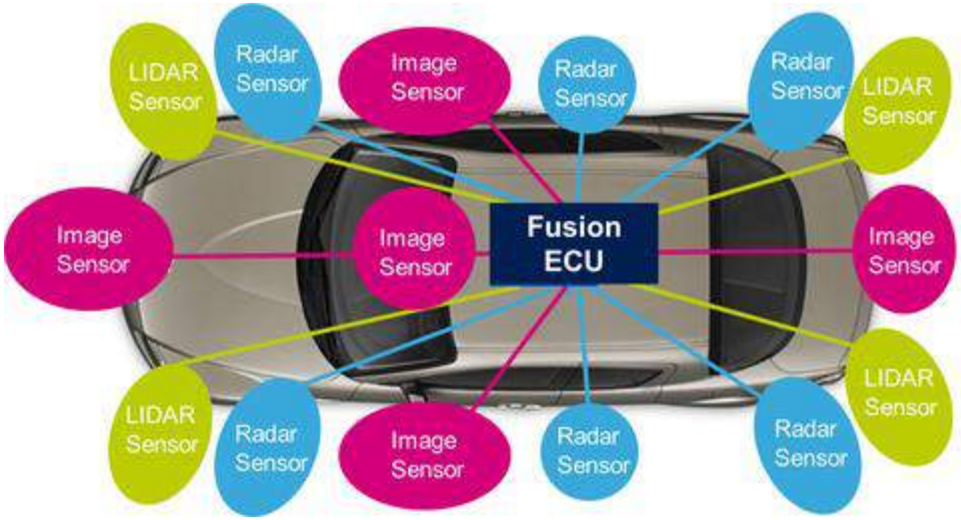
Overview

- Different approaches autonomy
- **Sensors**
- Companies doing it
- Opportunities for AI and deep learning

By 2030: Sensor Market Estimated at $36 Billion



Automotive AI Sensors

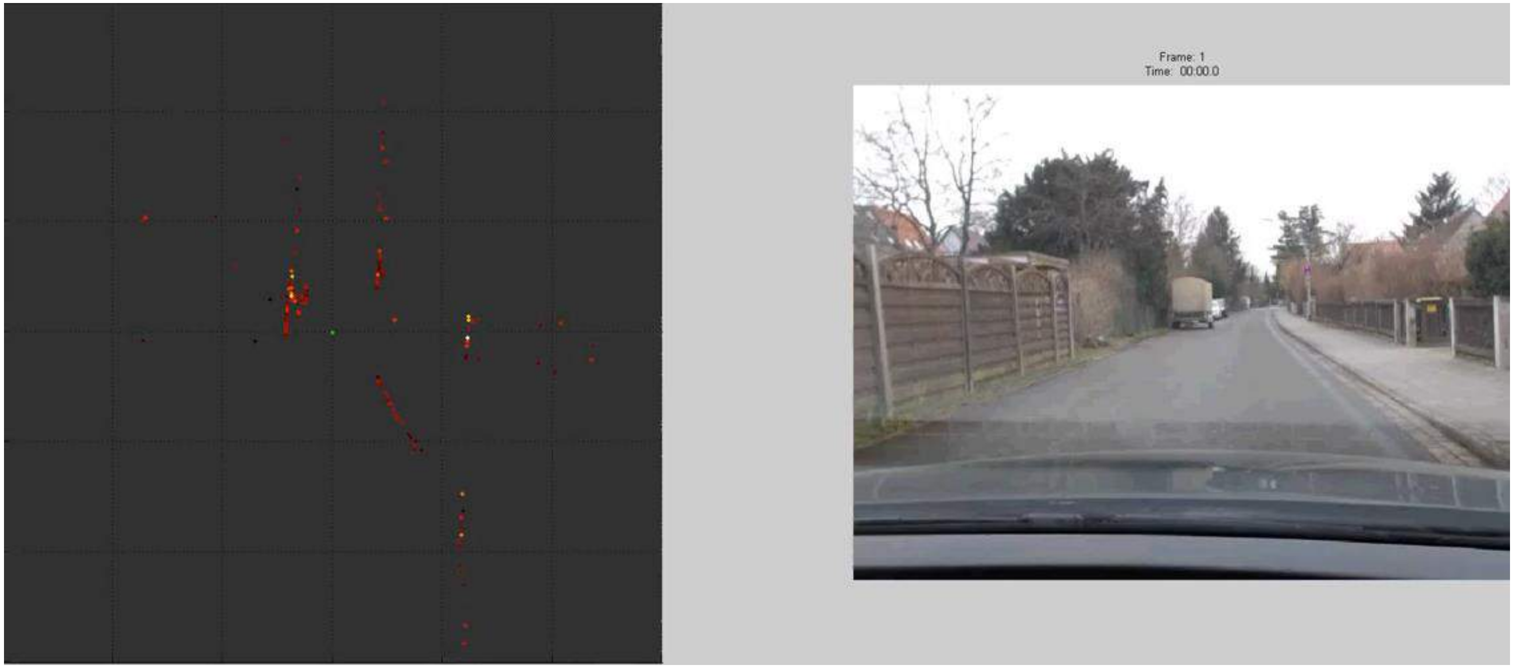


Camera

Radar

LIDAR

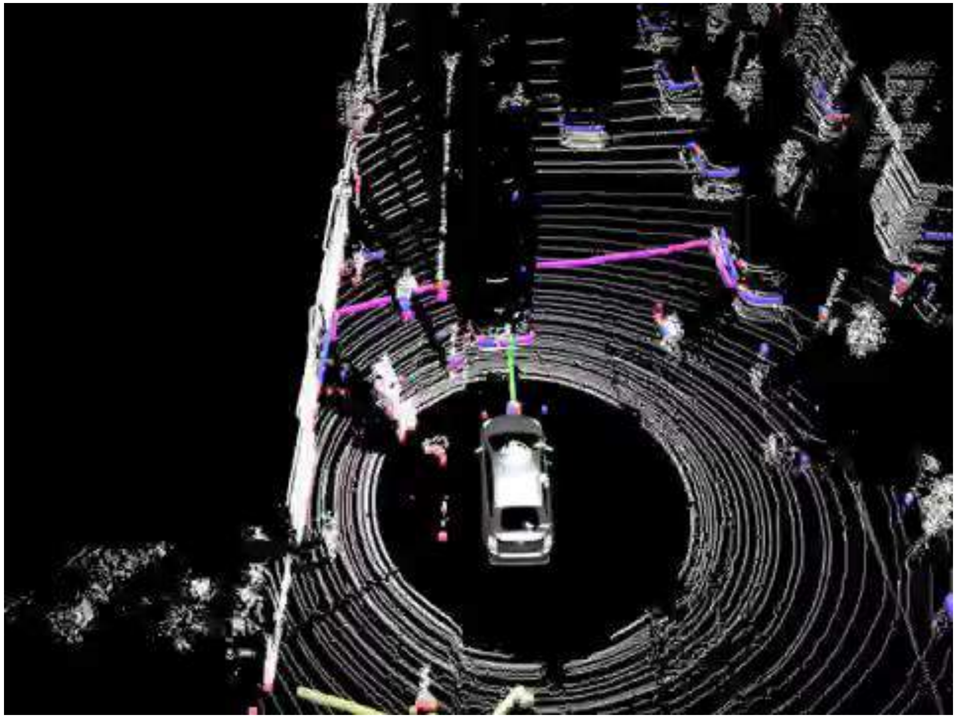
Radar



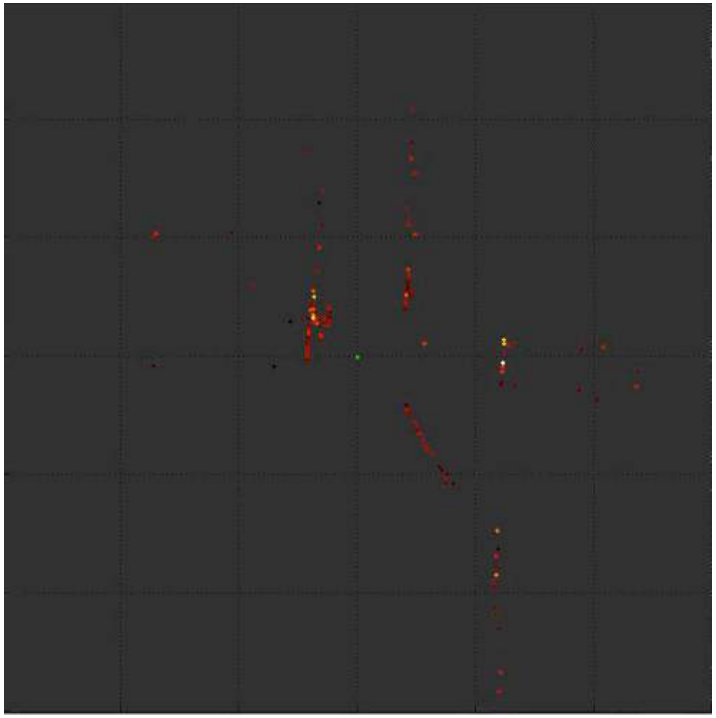
- Cheap
- Does well in extreme weather
- Low resolution
- Most used automotive sensor for object detection and tracking

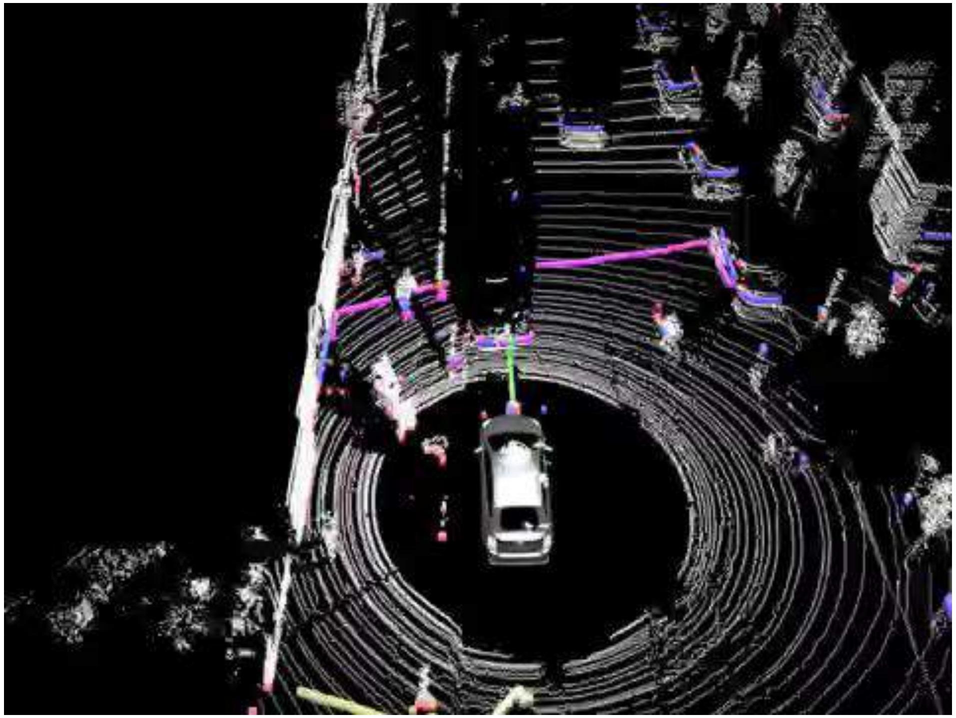
LIDAR

- Expensive
- Extremely accurate depth information
- Resolution much higher than radar
- 360 degrees of visibility



Resolution: LIDAR vs Radar

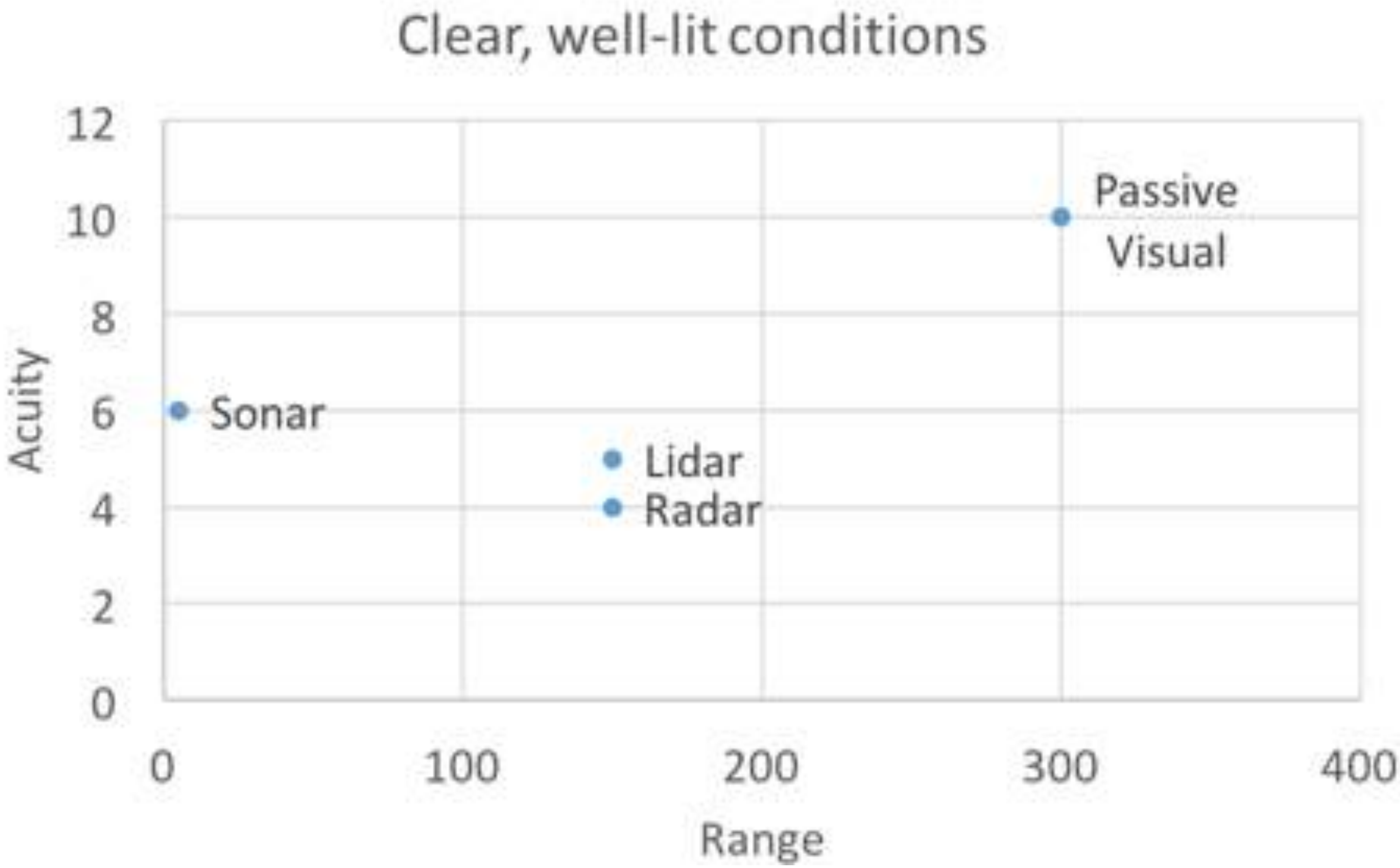


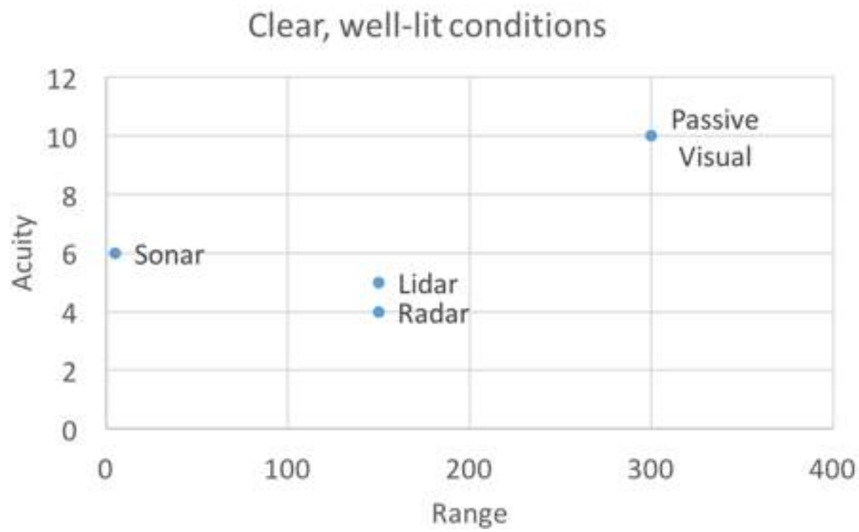


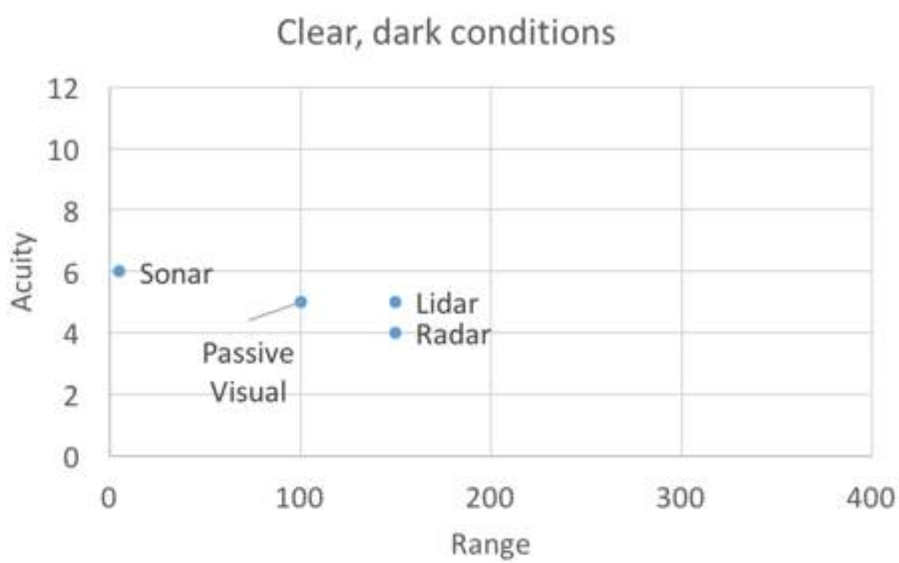
Camera

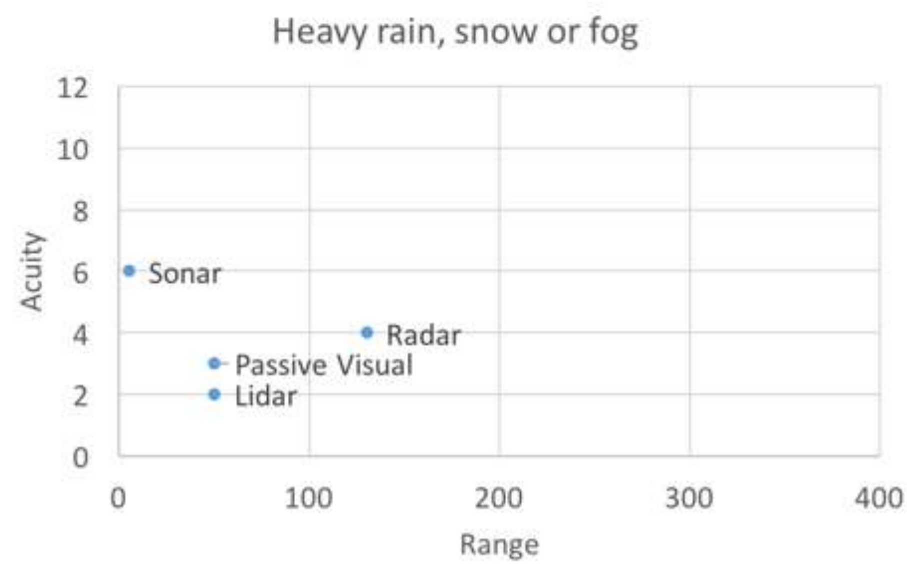
- Cheap
- Highest resolution
- Huge data = deep learning
- Human brains use similar sensor technology for driving
- Bad at depth estimation
- Not good in extreme weather

Range Comparison

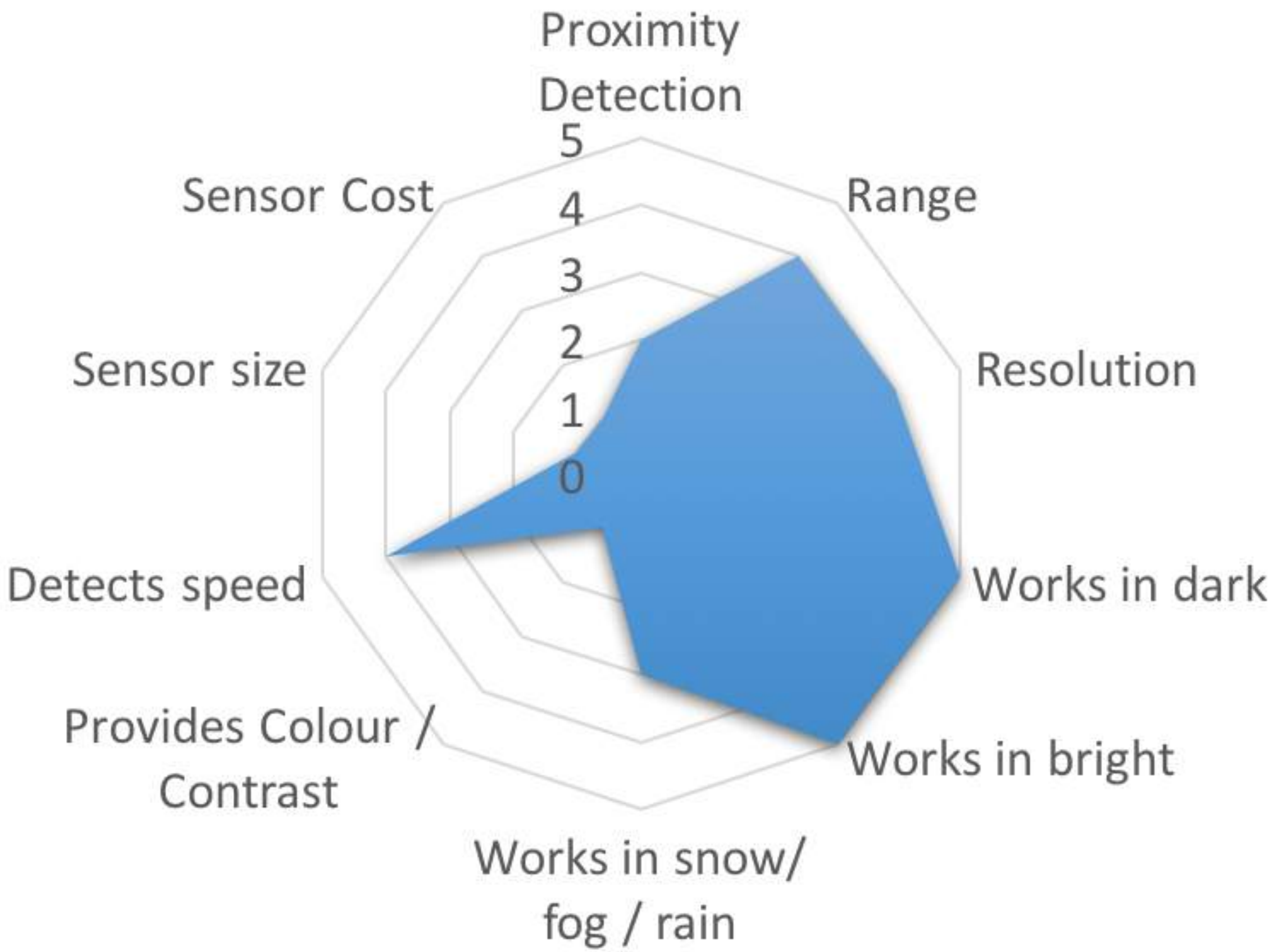



Clear, well-lit conditions
12
10
8
6
4
2
0
Acuity
Passive Visual
Sonar
Lidar
Radar
0
100
200
300
400
Range

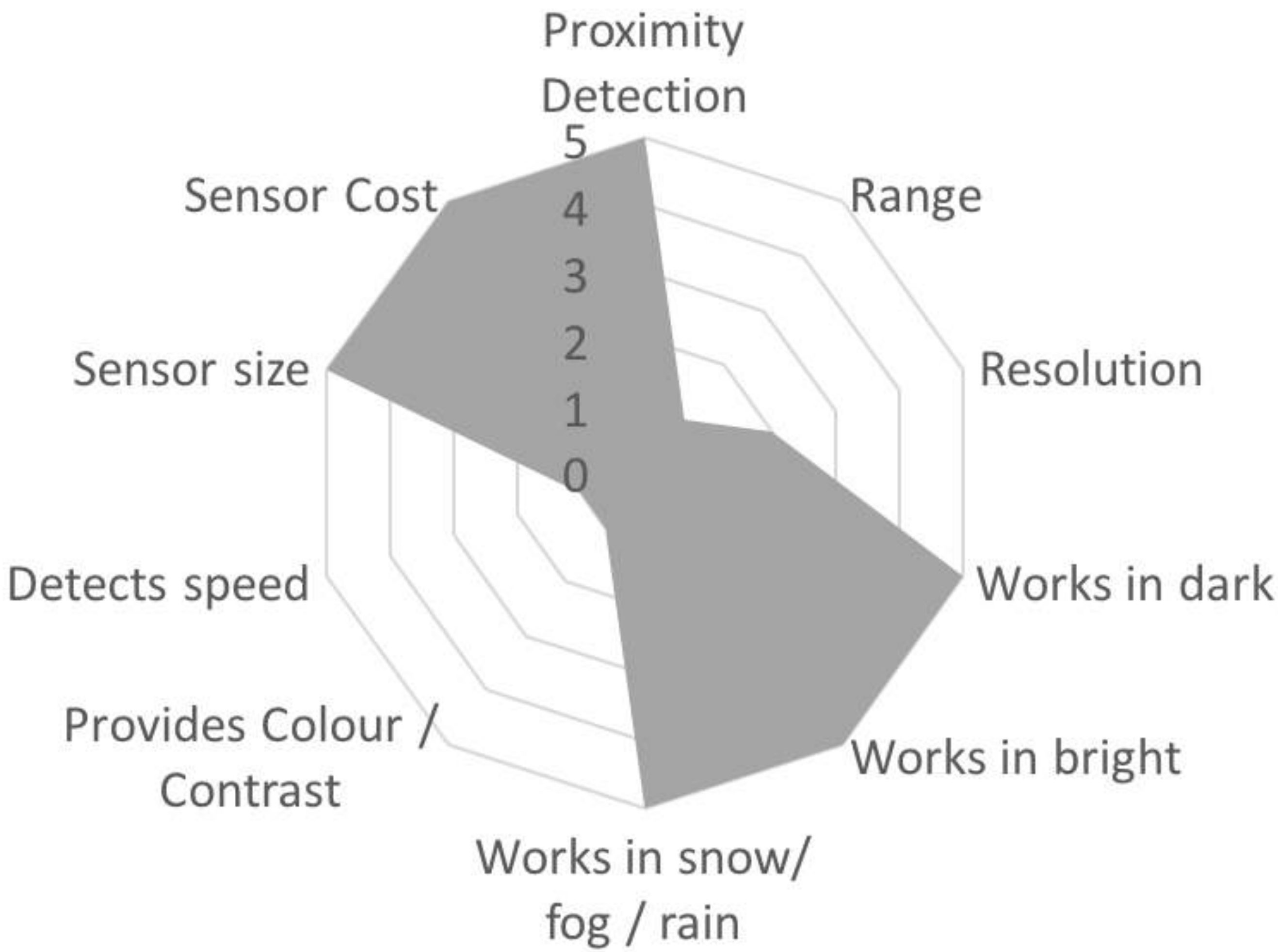

Clear, dark conditions
12
10
8
6
4
2
0
Acuity
Sonar
Lidar
Radar
Passive Visual
0
100
200
300
400
Range


Heavy rain, snow or fog
12
10
8
6
4
2
0
Acuity
Sonar
Radar
Passive Visual
Lidar
0
100
200
300
400
Range

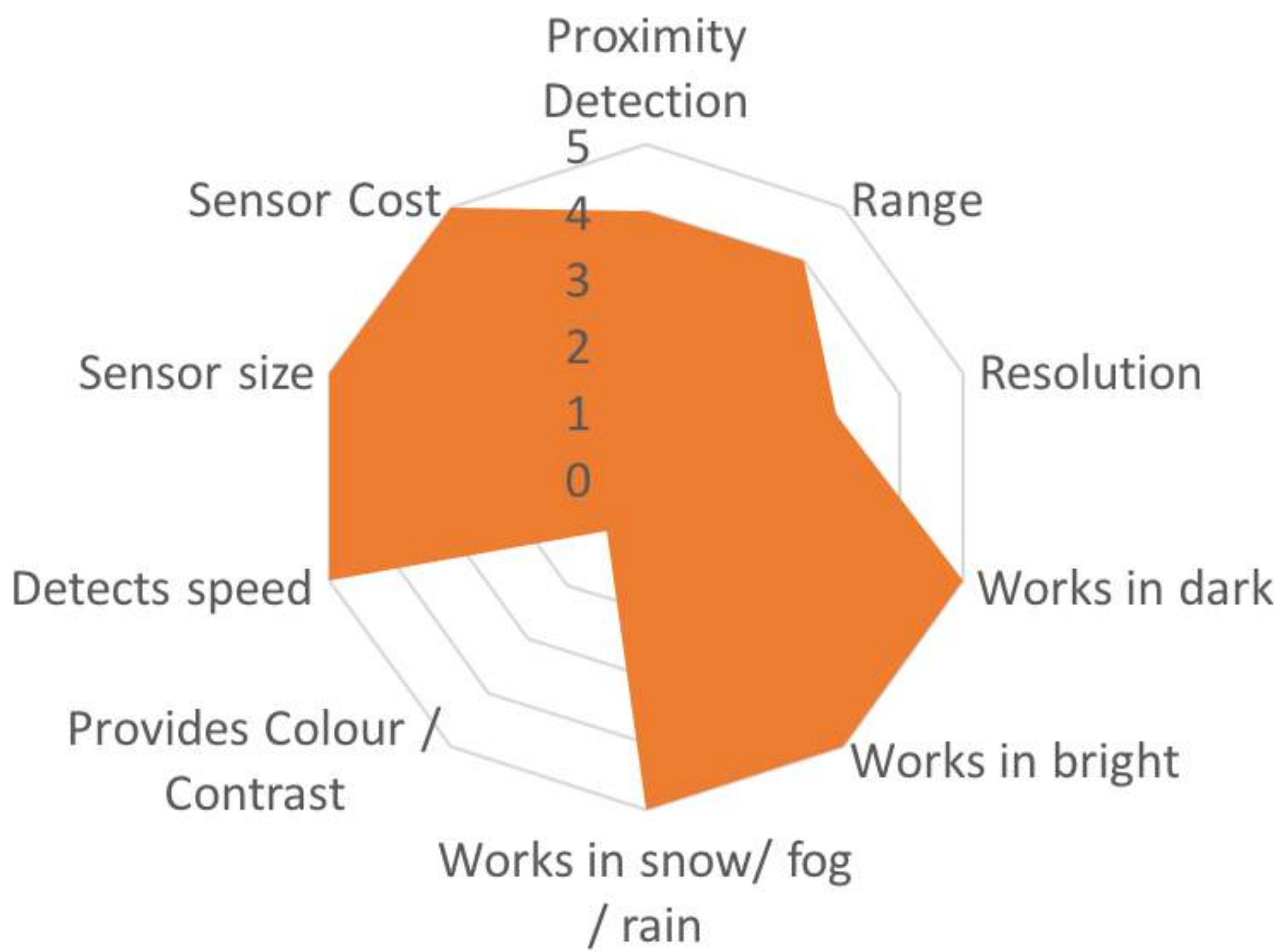
Lidar


Proximity Detection
5
4
3
2
1
0
Range
Resolution
Works in dark
Works in bright
Works in snow/ fog / rain
Provides Colour / Contrast
Detects speed
Sensor size
Sensor Cost

Ultrasonic

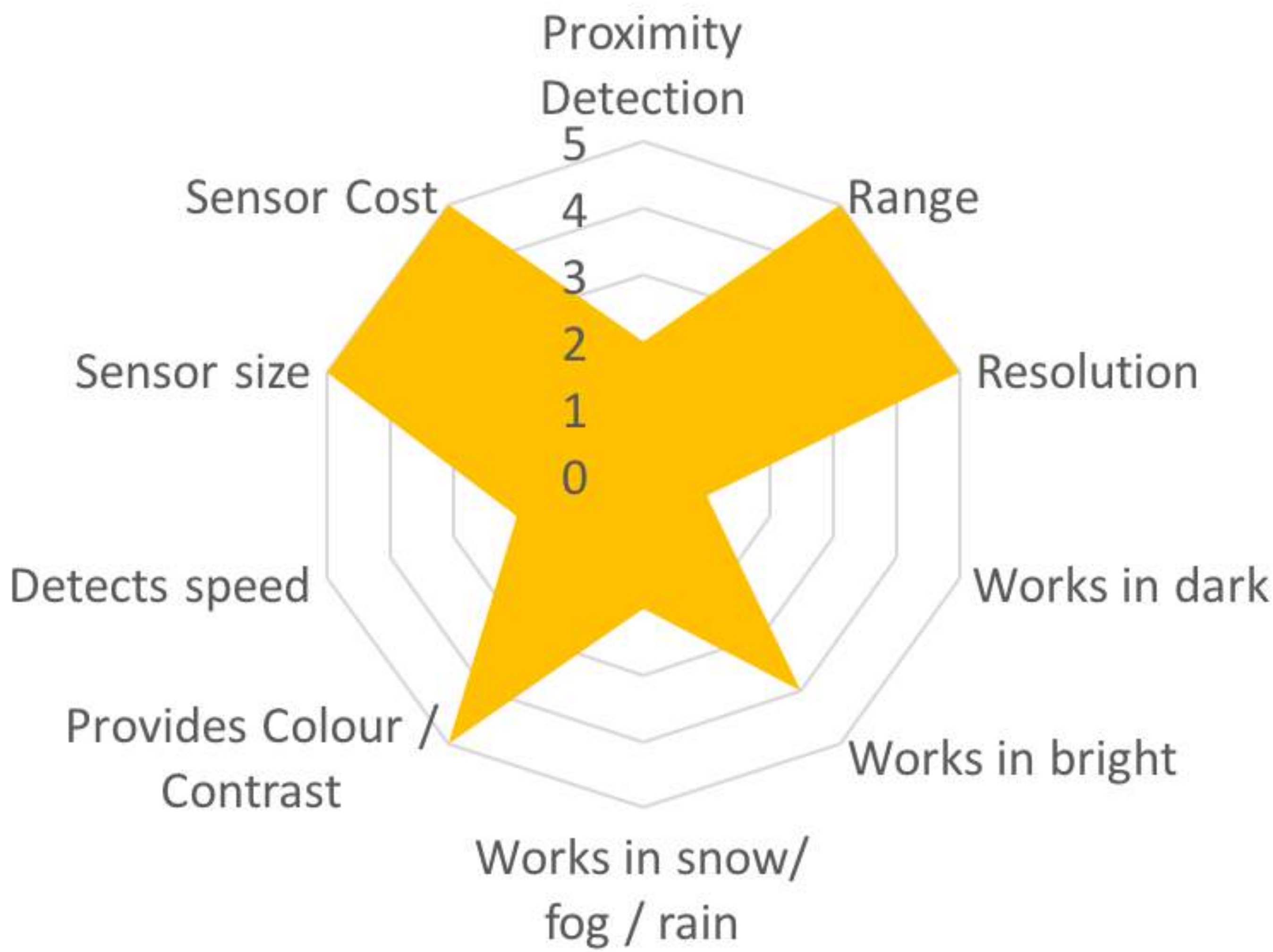

Proximity Detection
Range
Resolution
Works in dark
Works in bright
Works in snow/ fog / rain
Provides Colour / Contrast
Detects speed
Sensor size
Sensor Cost
5
4
3
2
1
0

Radar

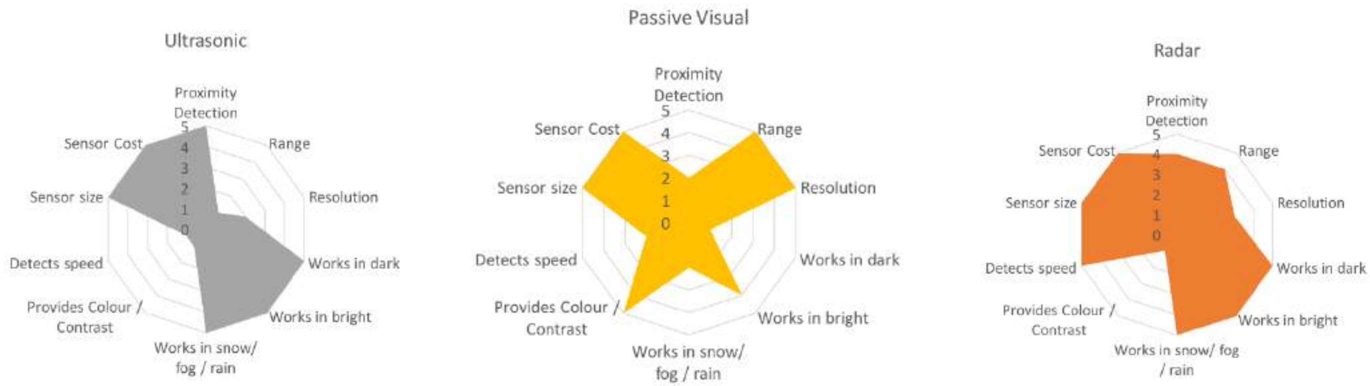


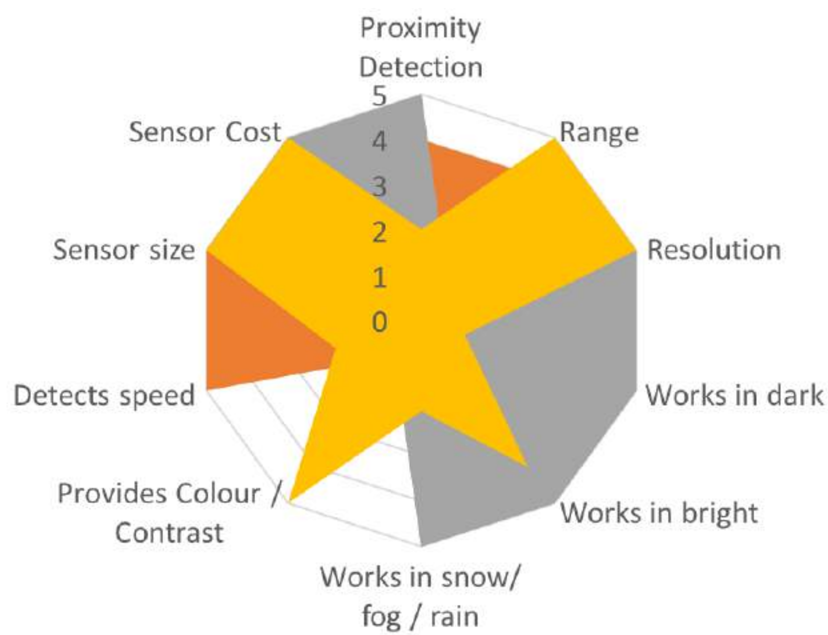
Proximity Detection
5
4
3
2
1
0
Range
Resolution
Works in dark
Works in bright
Works in snow/ fog / rain
Provides Colour / Contrast
Detects speed
Sensor size
Sensor Cost

Passive Visual



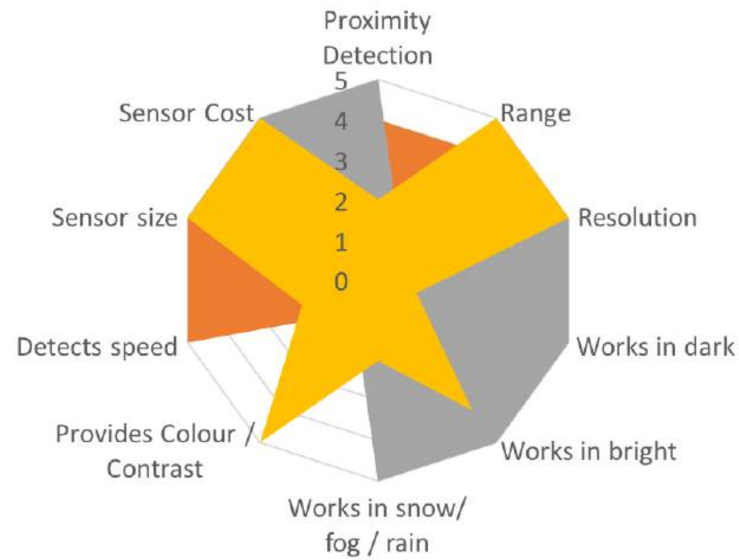
Sensor Fusion

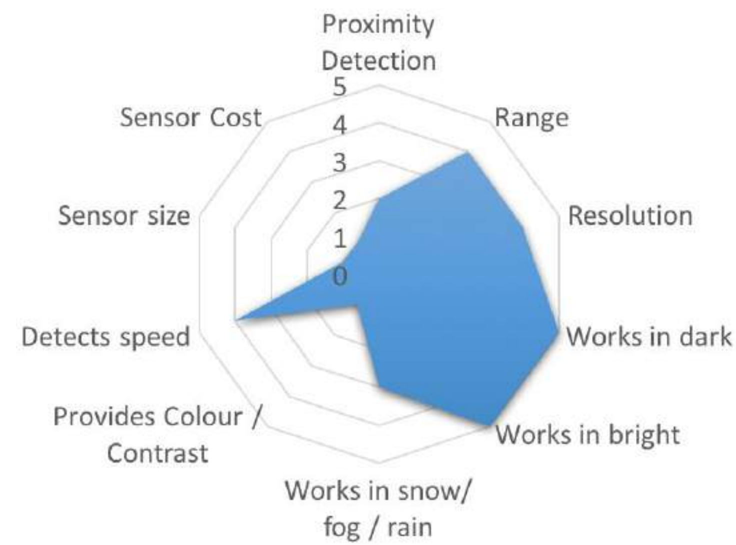




Future of Sensor Technology: Camera vs LIDAR

- **Radar and Ultrasonic:**
  - Always there to help
- **Camera:**
  - Annotated driving data grows
  - Deep learning algorithms improve
- **LIDAR:**
  - Range increases
  - Cost drops (solid-state LIDAR)





Overview

- Different approaches autonomy
- Sensors
- **Companies doing it**
- Opportunities for AI and deep learning

Waymo

Notable:

- April 2017: Exits testing**:** first rider in Phoenix
- November 2017: 4 million miles driven autonomously
- December 2017: No safety driver in Phoenix

Uber

Notable:

- December 2017: 2 million miles driven autonomously

Tesla

Notable:

- Sep 2014: Released Autopilot
- Oct 2016: Started Autopilot 2 from scratch.
- Jan 2018: ~1 billion miles driven in Autopilot
- Jan 2018: ~300,000 Autopilot equipped vehicles

Audi A8

(Released end of 2018)

- Thorsten Leonhardt, head of Automated Driving, Audio:
  “When the function is operated as intended, if the customer turns the traffic jam pilot on and uses it as intended, and the car was in control at the time of the accident, the driver goes to his insurance company and the insurance company will compensate the victims of the accident and in the aftermath they come to us and we have to pay them,” he said.

Notable Progress

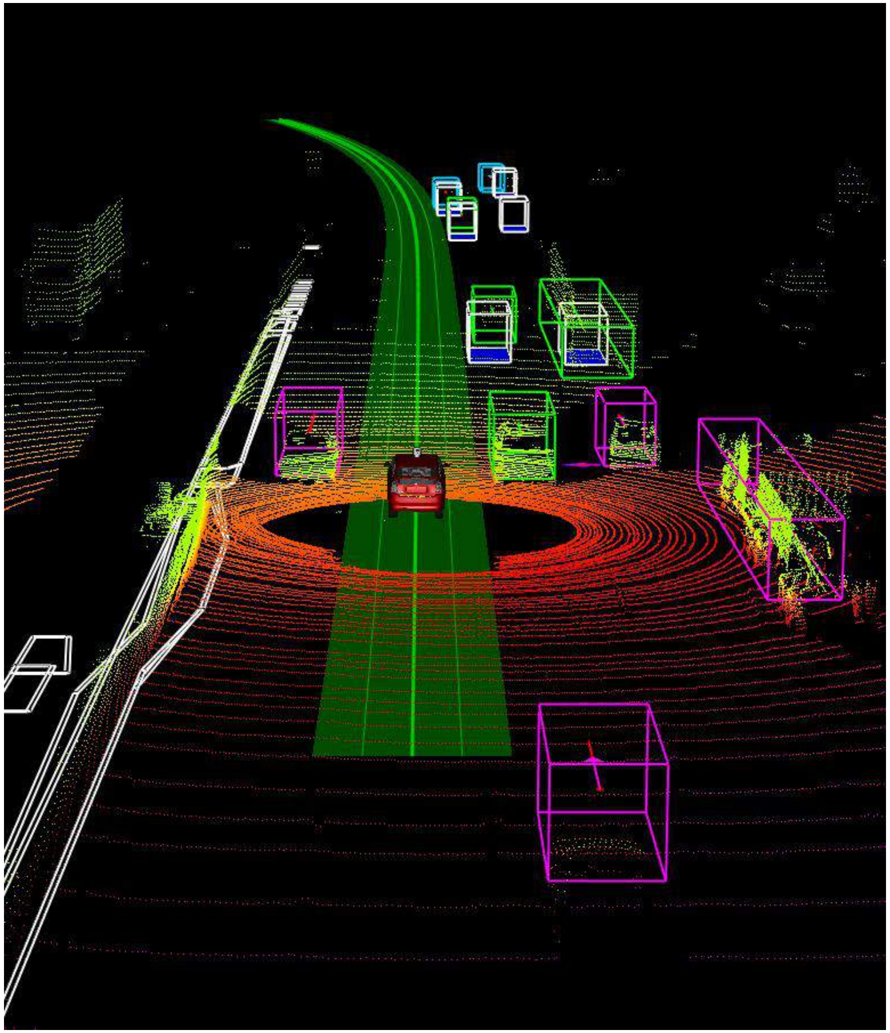
- Full autonomy (A2)
  - Waymo
  - Uber
  - GM Cruise
  - nuTonomy
  - OptimusRide
  - Zenuity
  - Voyage
  - …
- Human-centered autonomy (A1)
  - Tesla Autopilot - Model S/3/X
  - Volvo PilotAssist - S90/XC90/XC60/V90
  - Audi Traffic Jam Assist - A8
  - Mercedes-Benz Drive Pilot Assist - E-Class
  - Cadillac Super Cruise - CT6
  - Comma.ai openpilot
  - …

Overview

- Different approaches autonomy
- Sensors
- Companies doing it
- **Opportunities for AI and deep learning**

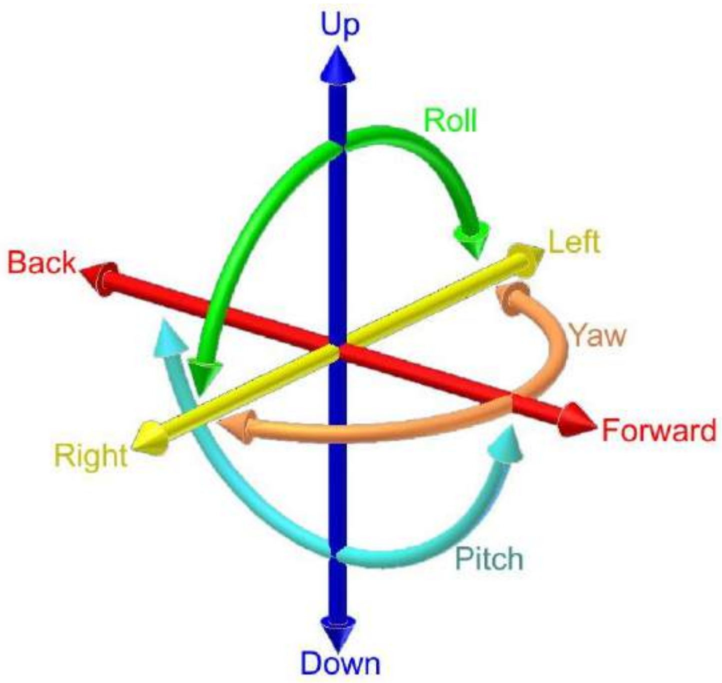
Self-Driving Car Tasks

- **Localization and Mapping:** Where am I?
- **Scene Understanding:** Where is everyone else?
- **Movement Planning:** How do I get from A to B?
- **Driver State:** What's the driver up to?



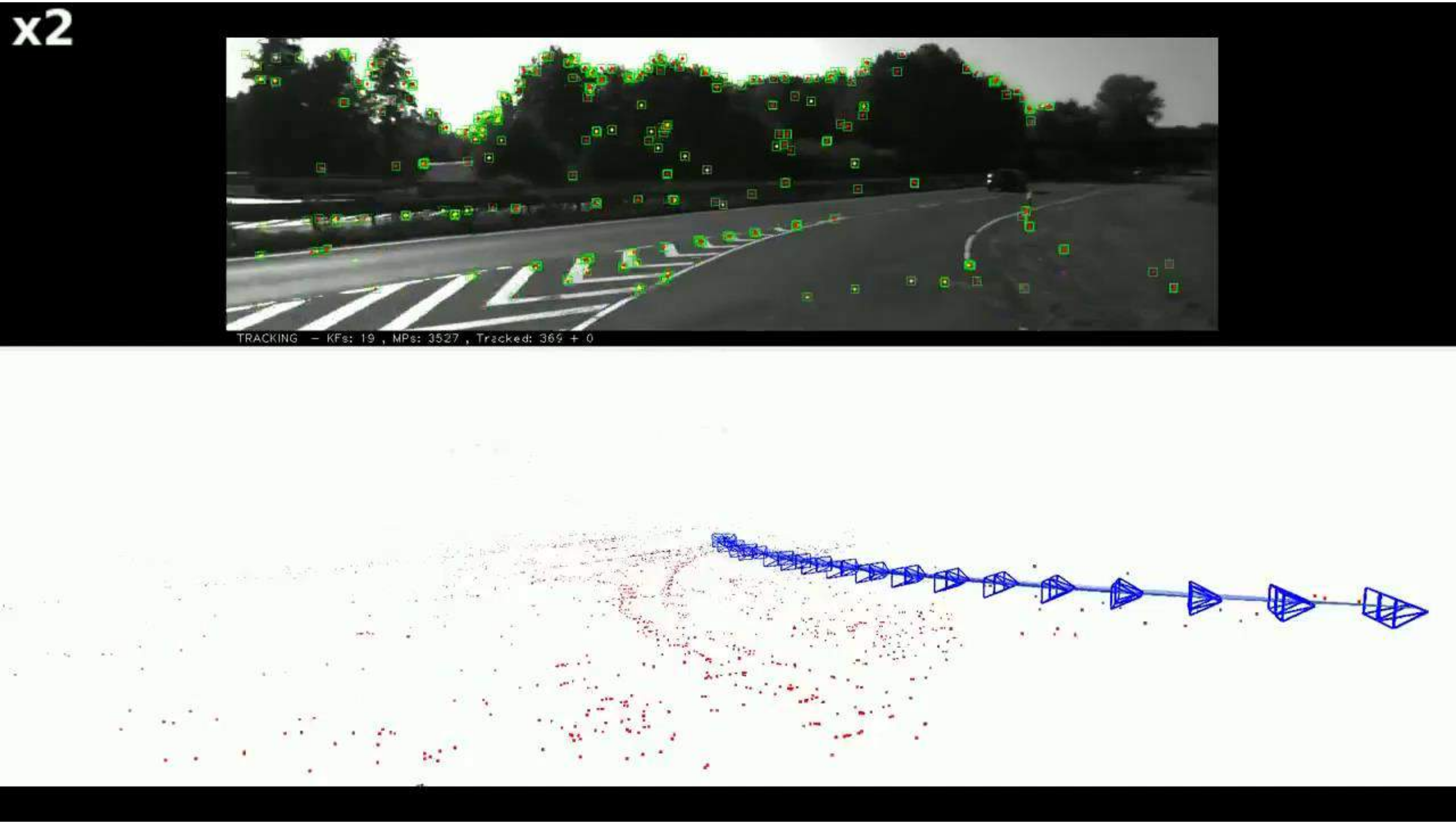
Visual Odometry

- 6-DOF: freed of movement
  - Changes in position:
    - Forward/backward: surge
    - Left/right: sway
    - Up/down: heave
  - Orientation:
    - Pitch, Yaw, Roll
- Source:
  - **Monocular:** I moved 1 unit
  - **Stereo:** I moved 1 meter
  - Mono = Stereo for far away objects
    - PS: For tiny robots everything is “far away” relative to inter-camera distance



SLAM: Simultaneous Localization and Mapping

What works: SIFT and optical flow



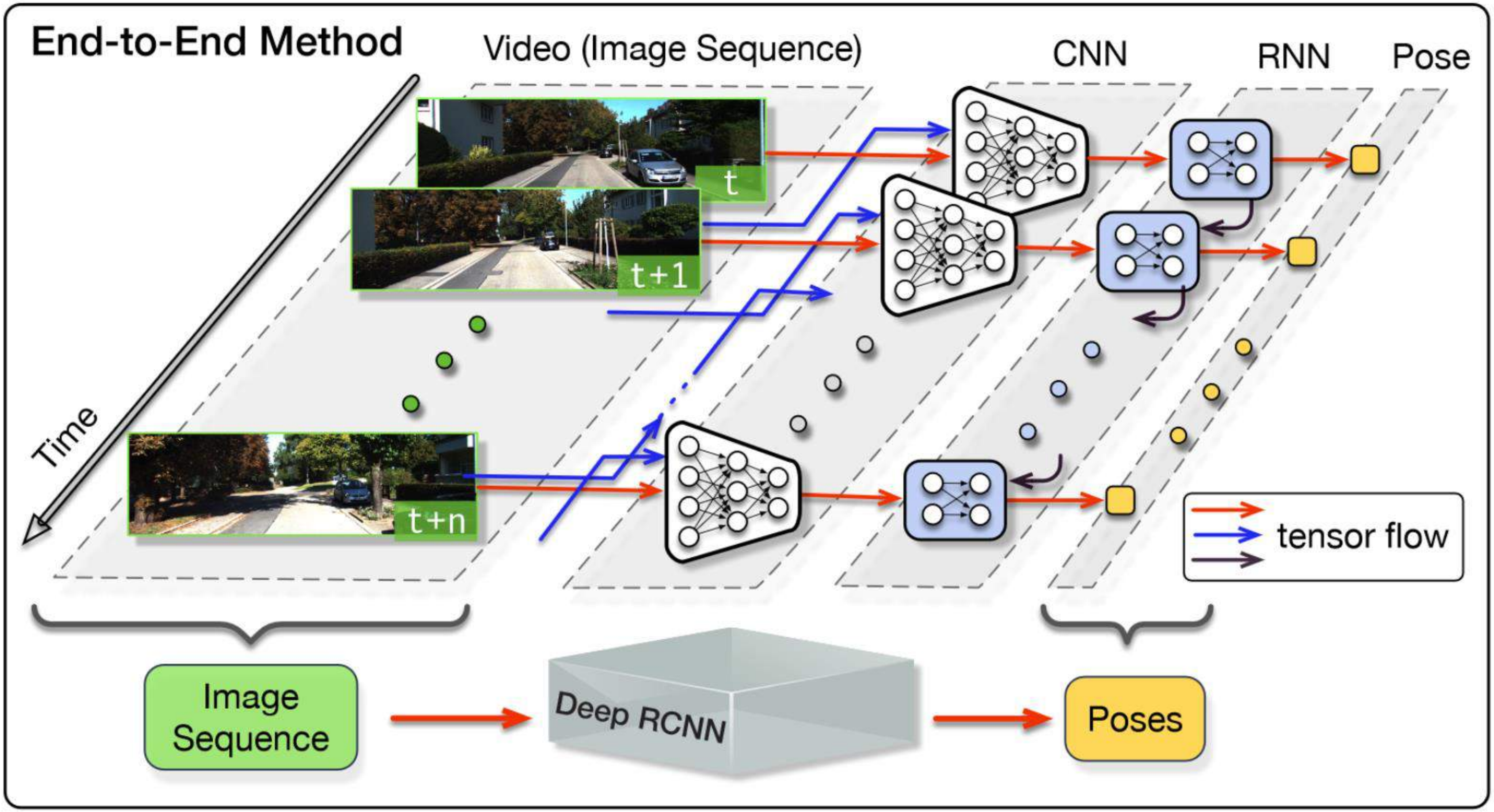
References: [98, 99]

Visual Odometry in Parts

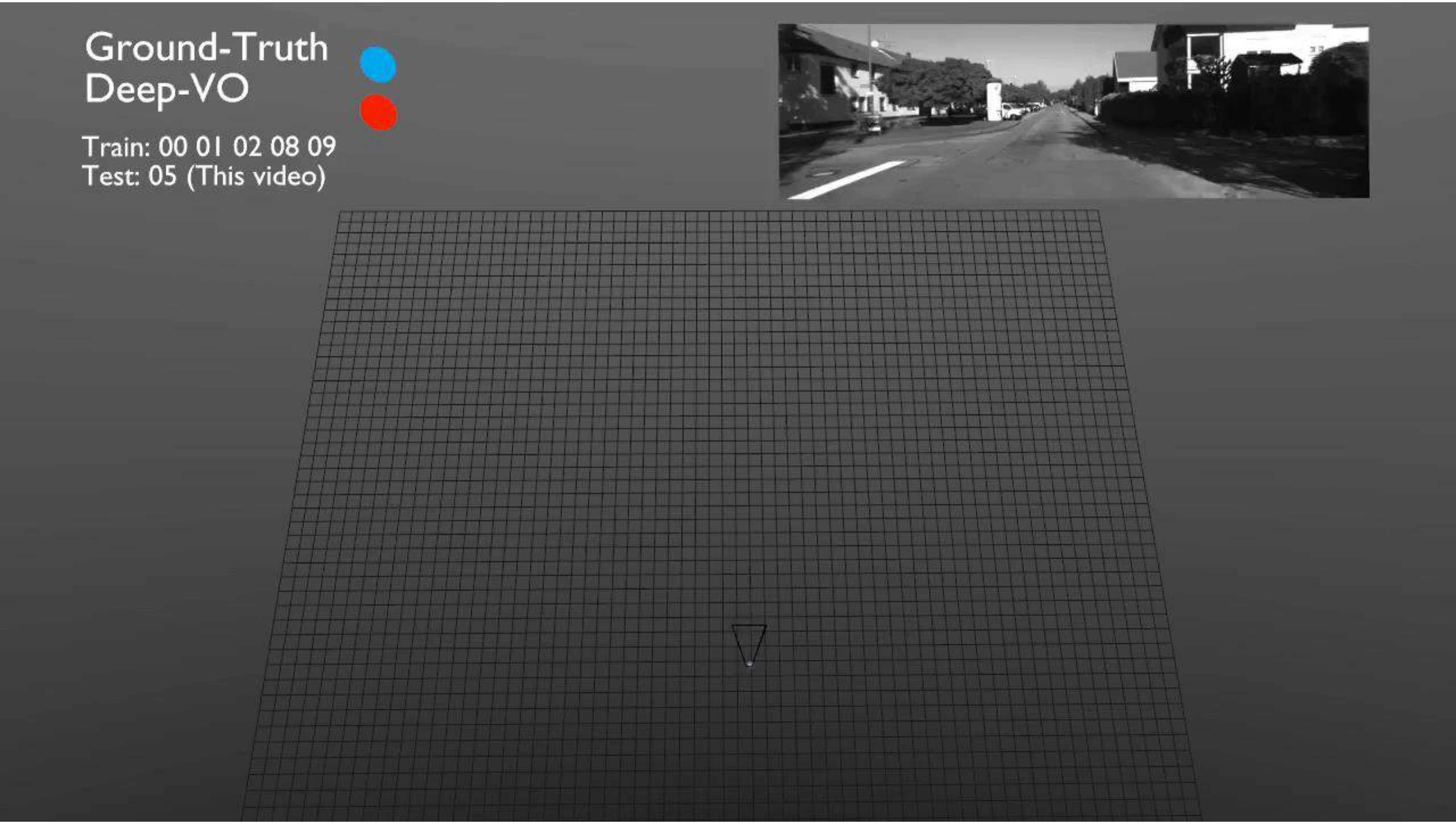
- (Stereo) Undistortion, Rectification
- (Stereo) Disparity Map Computation
- Feature Detection (e.g., SIFT, FAST)
- Feature Tracking (e.g., KLT: Kanade-Lucas-Tomasi)
- Trajectory Estimation
  - Use rigid parts of the scene (requires outlier/inlier detection)
  - For mono, need more info* like camera orientation and height of off the ground

* Kitt, Bernd Manfred, et al. "Monocular visual odometry using a planar road model to solve scale ambiguity." (2011).

DeepVO: Deep Learning Based Visual Odometry

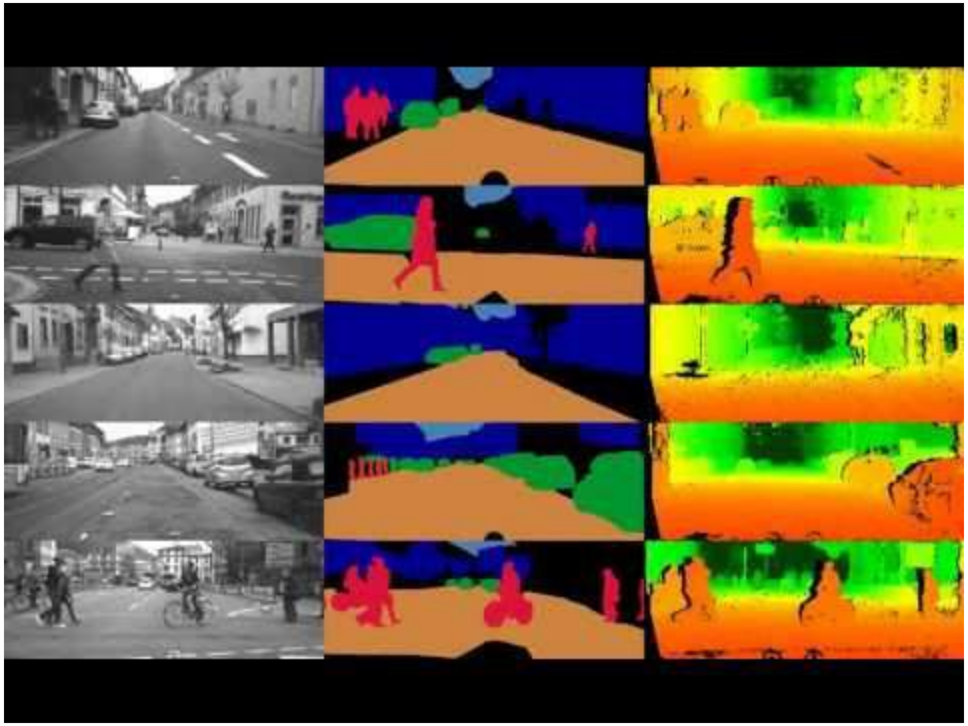


DeepVO: Deep Learning Based Visual Odometry



Self-Driving Car Tasks

- **Localization and Mapping:**
  Where am I?
- **Scene Understanding:**
  Where is everyone else?
- **Movement Planning:**
  How do I get from A to B?
- **Driver State:**
  What's the driver up to?



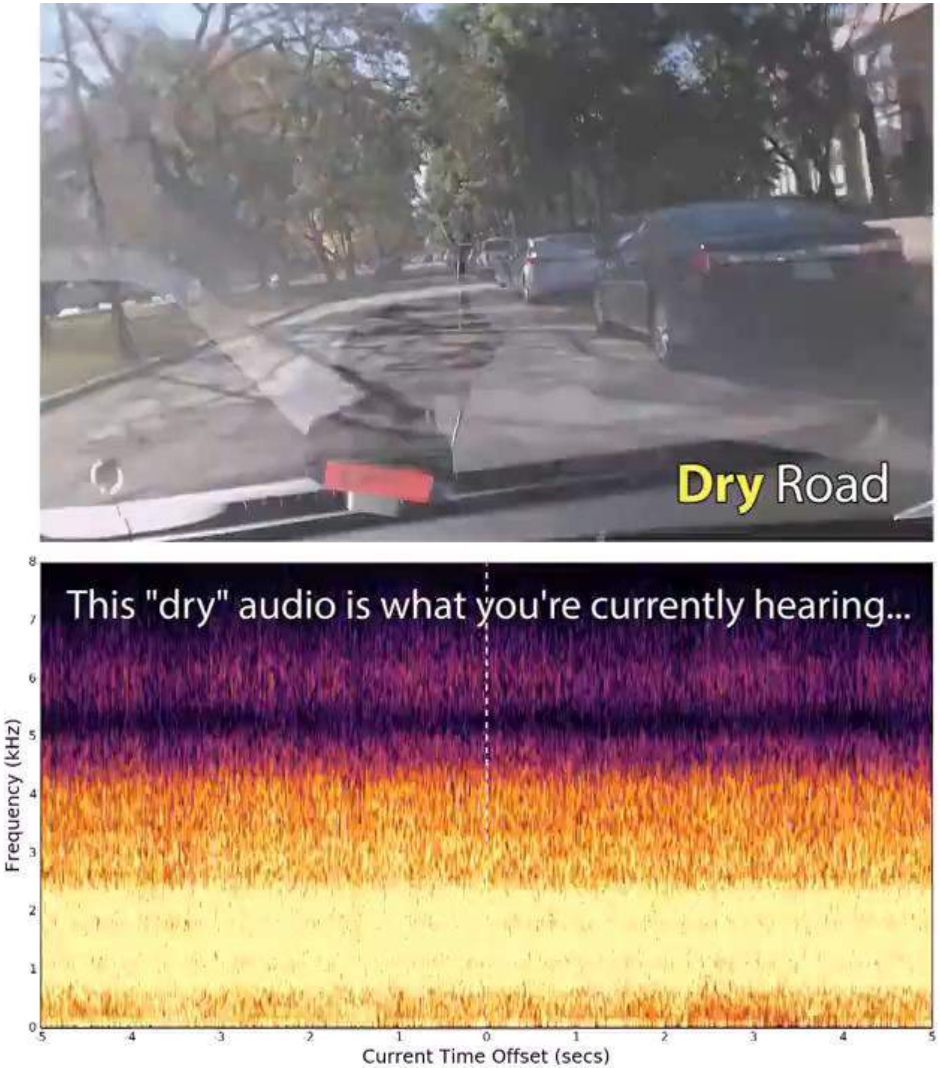
Object Detection

- Past approaches: cascades classifiers (Haar-like features)
- Where deep learning can help:
  recognition, classification, detection

Driving Scene Segmentation

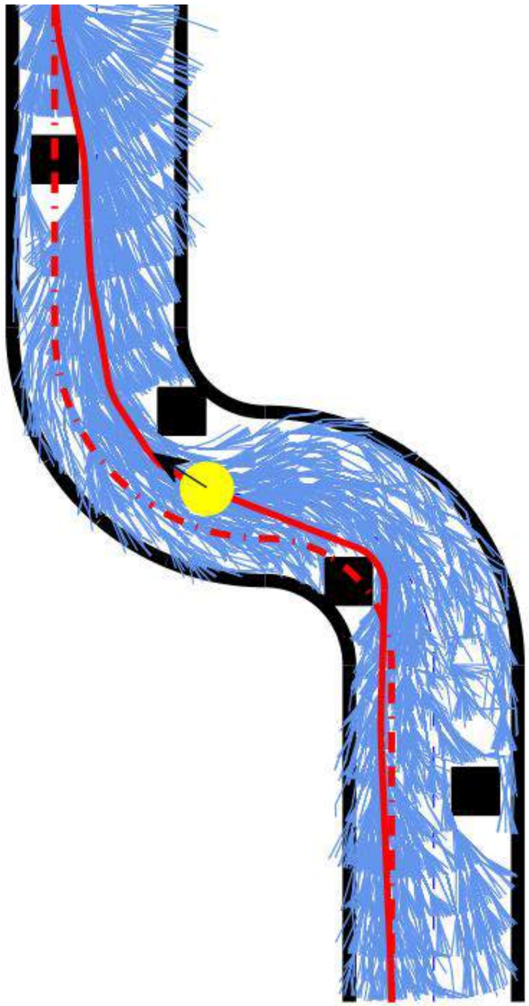
Road Texture and Condition from Audio

(with Recurrent Neural Networks)

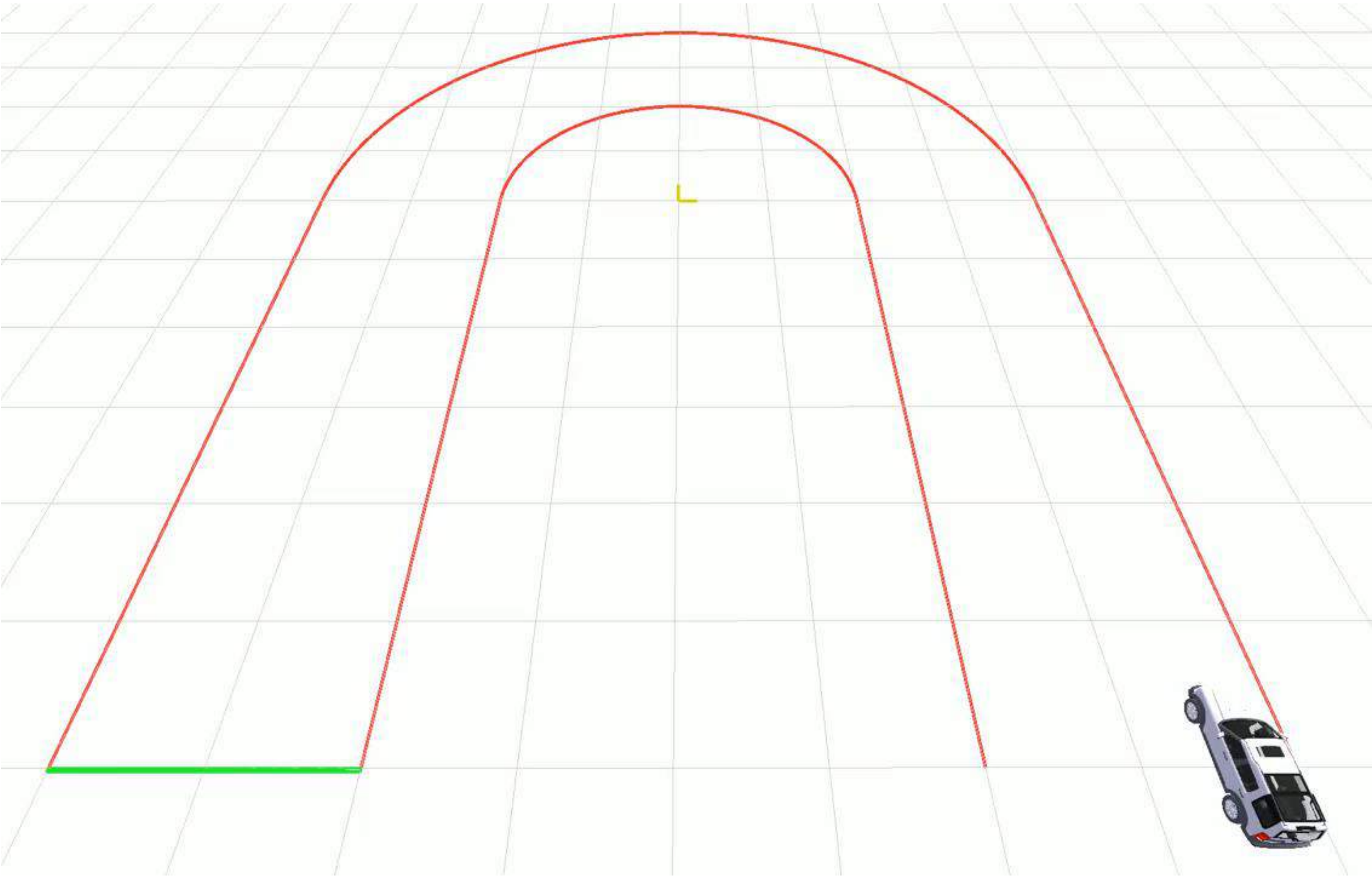


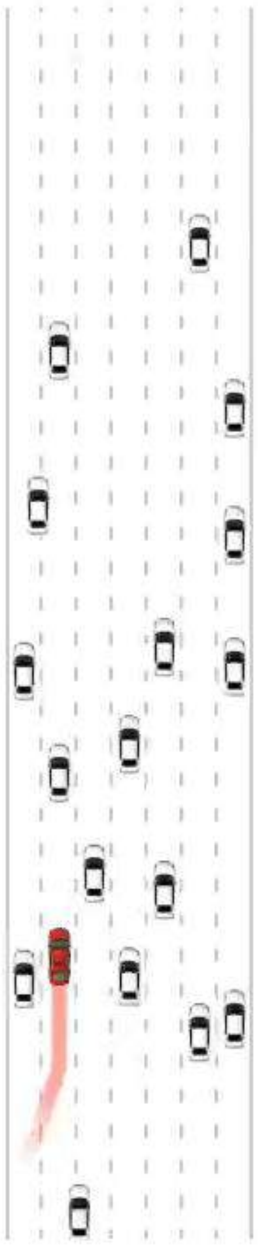
Self-Driving Car Tasks

- **Localization and Mapping:**
  Where am I?
- **Scene Understanding:**
  Where is everyone else?
- **Movement Planning:**
  How do I get from A to B?
- **Driver State:**
  What's the driver up to?



- **Previous approaches:** optimization-based control
- **Deep reinforcement learning:** give the ability to deal with under-actuated control, uncertainty, motion blur, lack of sensor calibration or prior map information.

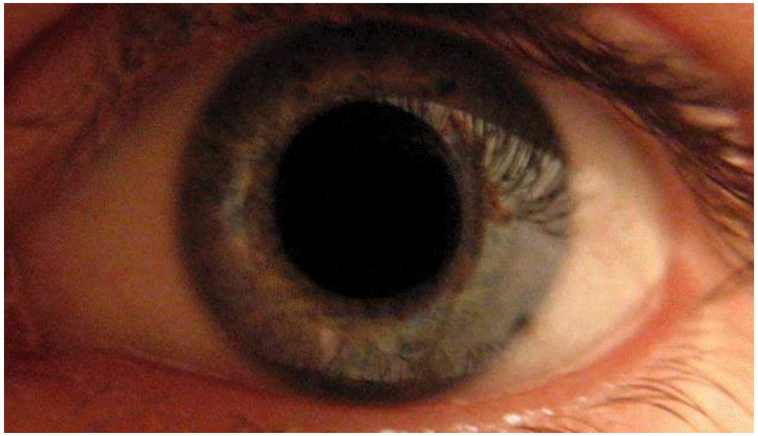


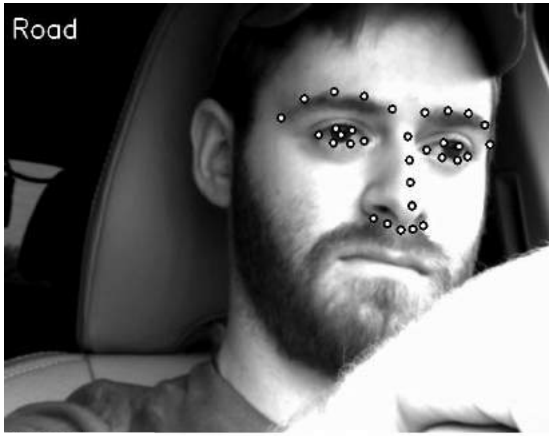


- **Previous approaches:** optimization-based control
- **Deep reinforcement learning:** give the ability to deal with under-actuated control, uncertainty, motion blur, lack of sensor calibration or prior map information.

Self-Driving Car Tasks

- **Localization:**
  Where am I?
- **Object detection:**
  Where is everyone else?
- **Movement planning:**
  How do I get from A to B?
- **Driver state:**
  What's the driver up to?



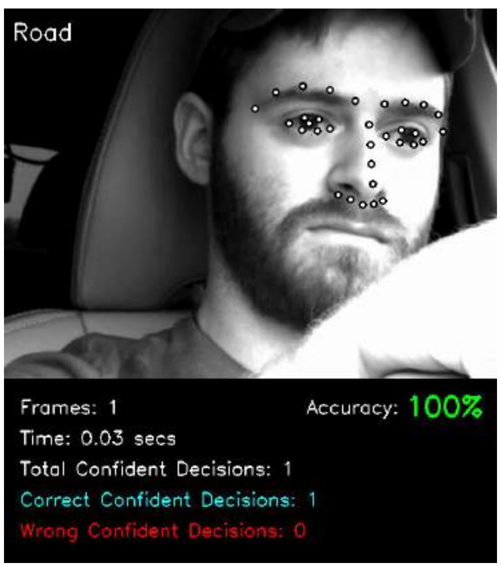


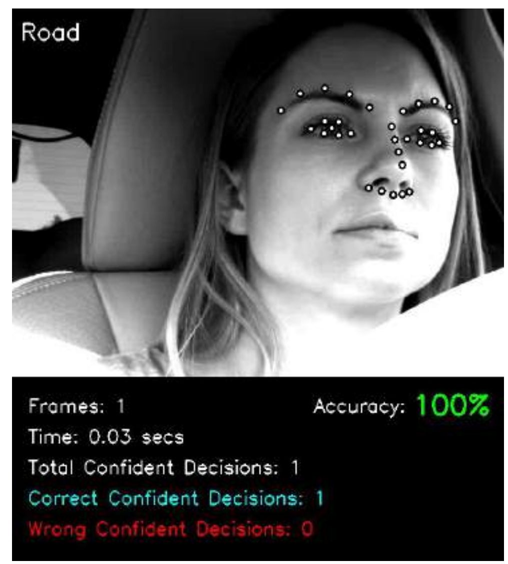
Drive State Detection:
A Multi-Resolutional View

Increasing level of detection resolution and difficulty →

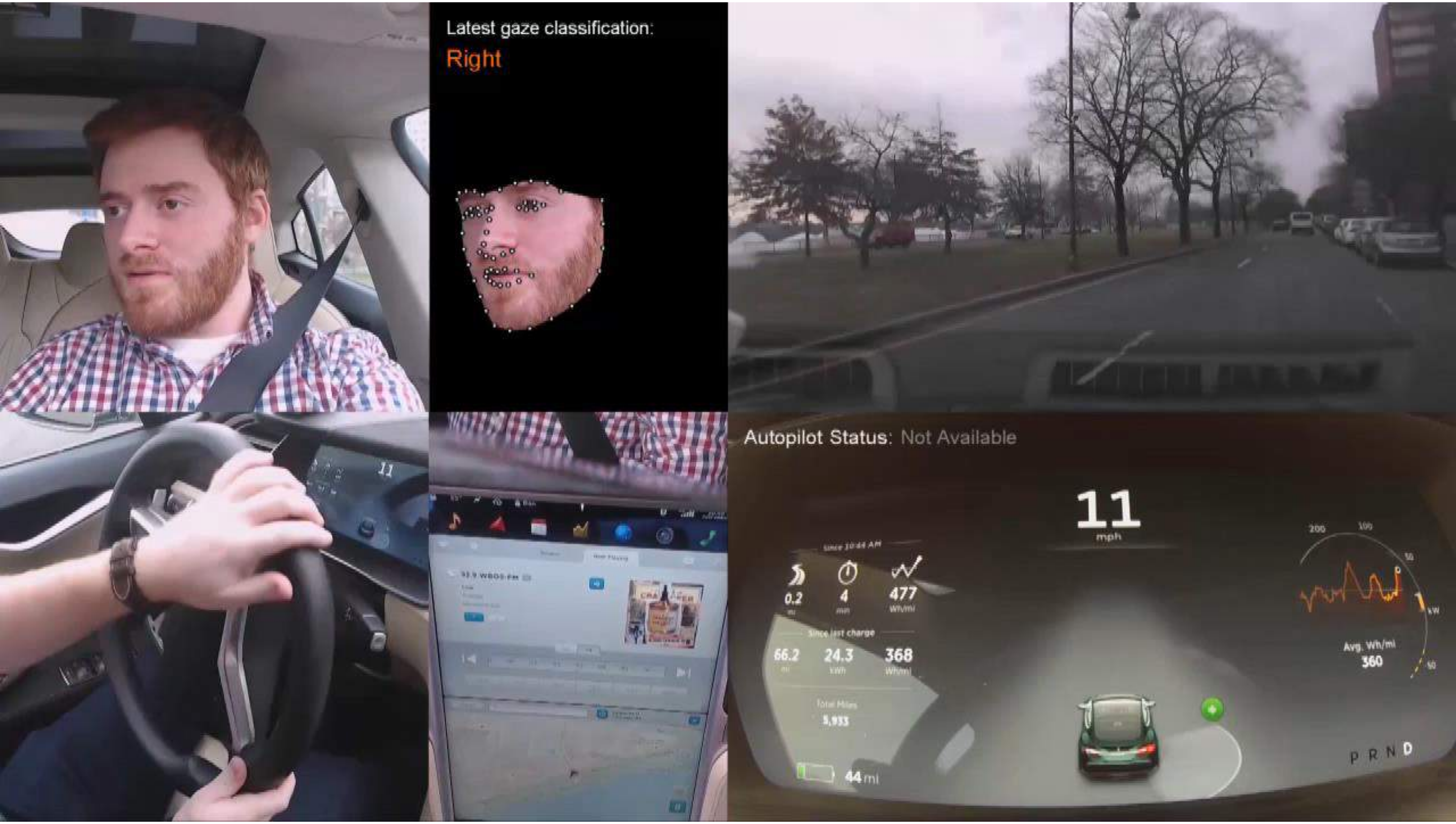
Body Pose | Head Pose | Blink Rate | Blink Duration | Eye Pose | Blink Dynamics | Pupil Diameter | Micro Saccades

Gaze Classification | Drowsiness | Micro Glances | Cognitive Load

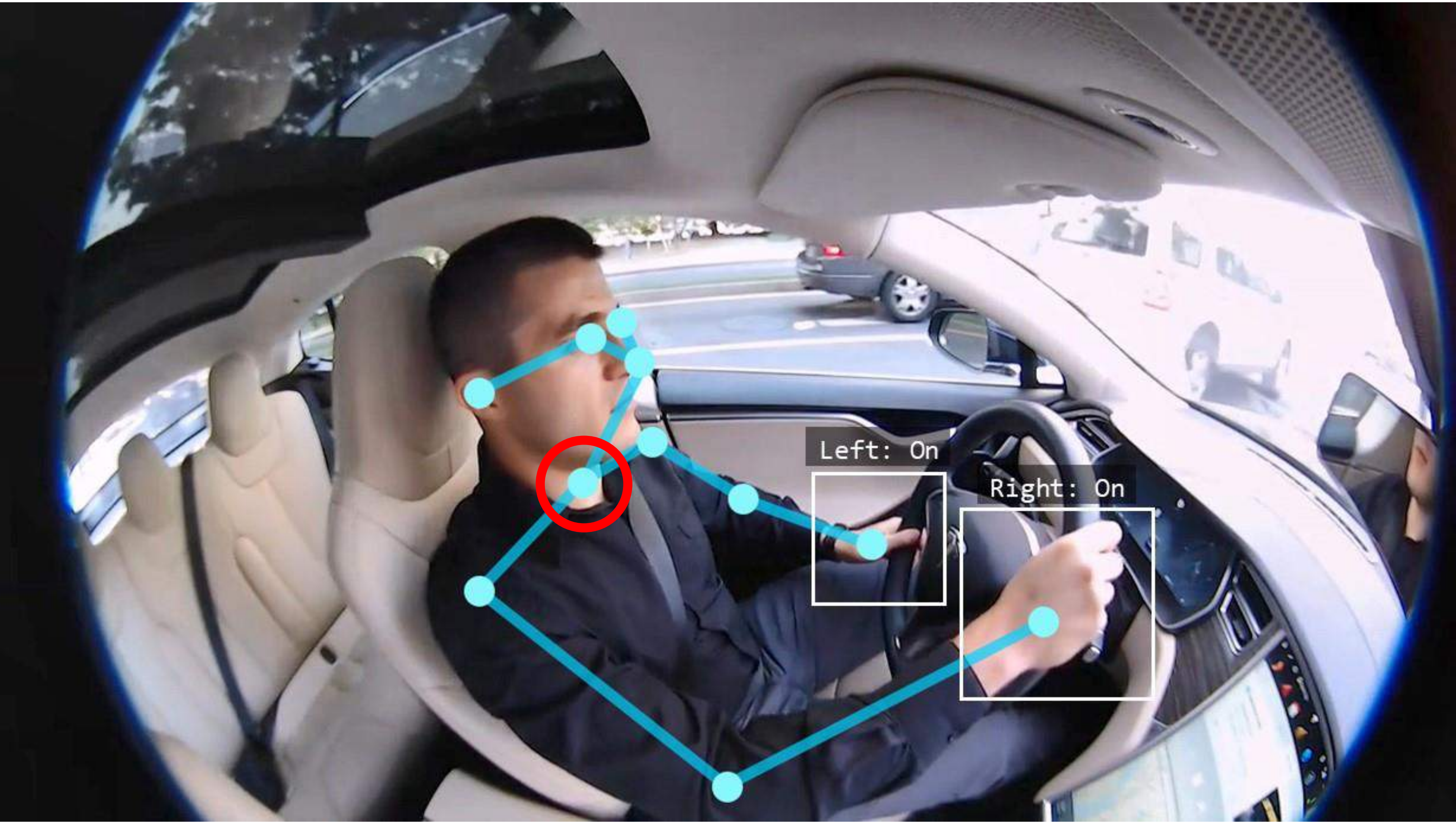




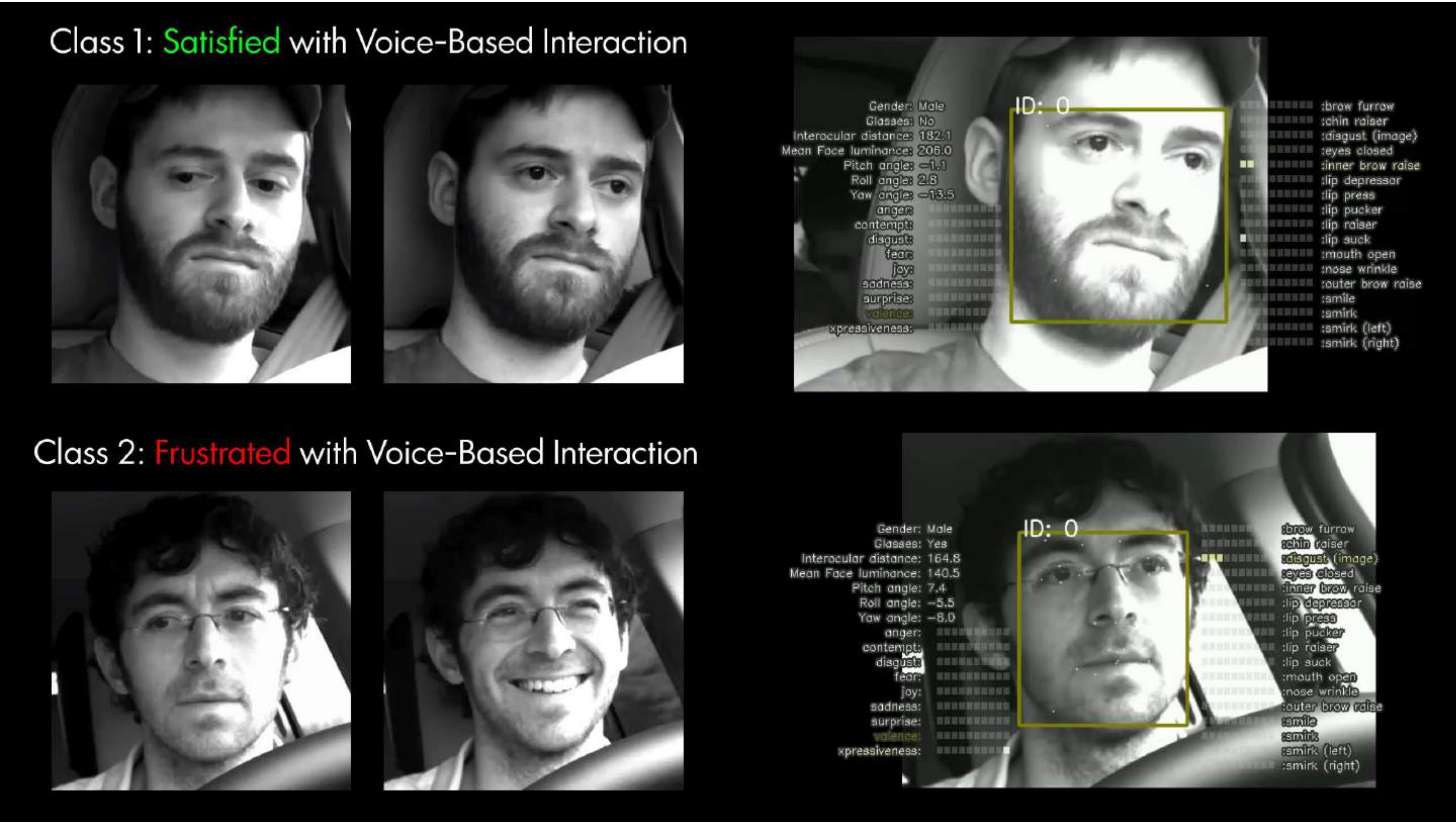
Driver Glance Region Classification



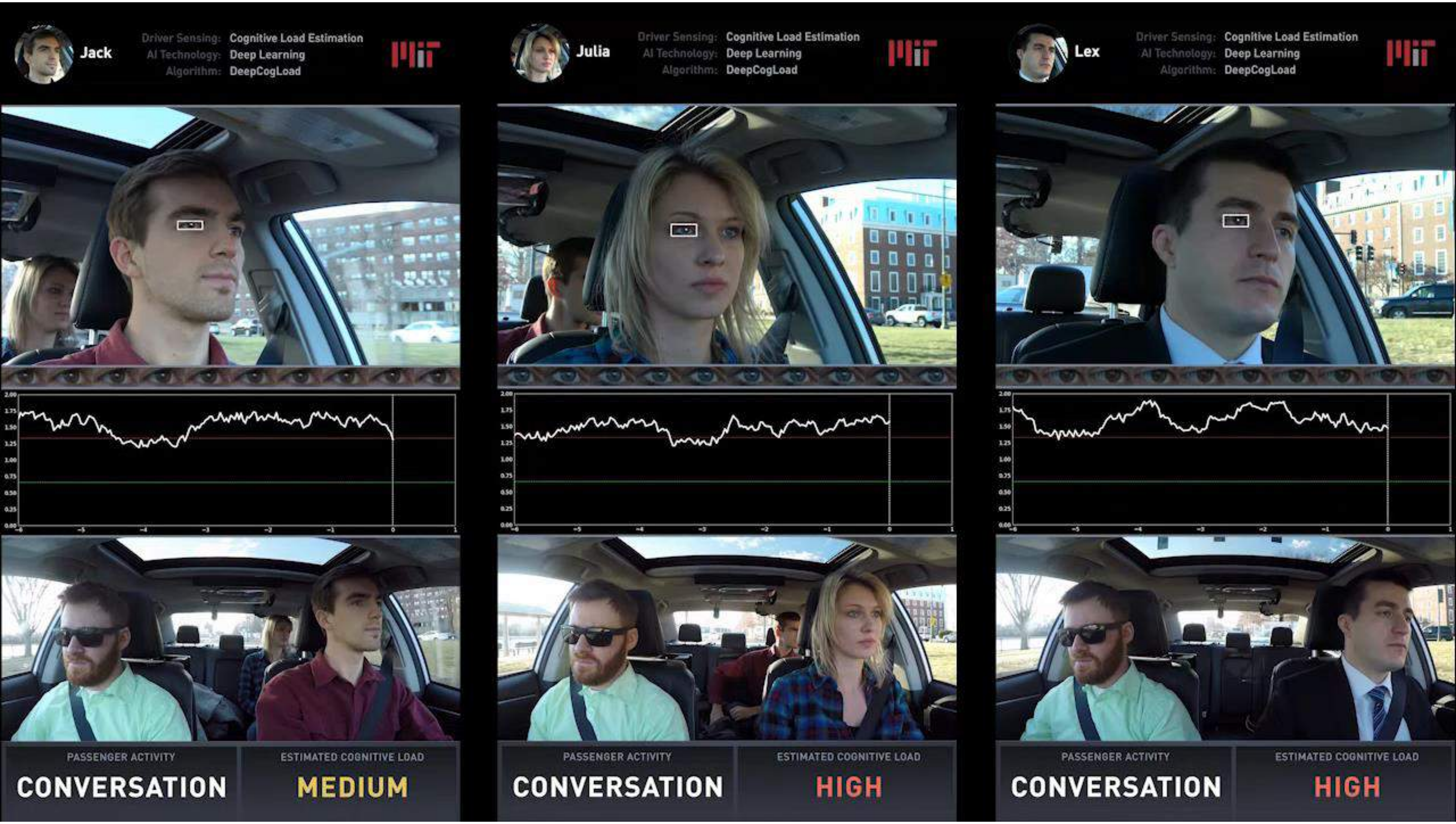
Driver Body Pose Estimation



Driver Emotion



Cognitive Load

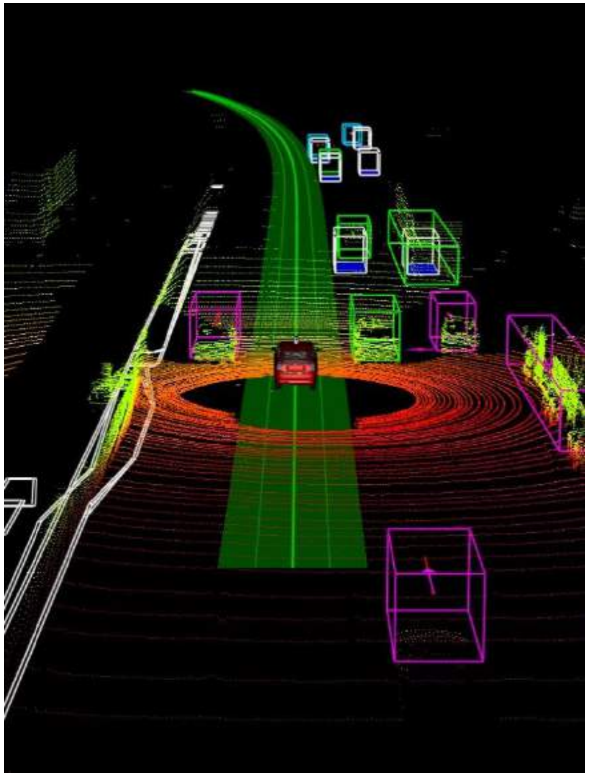


Two Paths to an Autonomous Future

A1:
Human-Centered Autonomy

- **Localization and Mapping:**
  Where am I?
- **Scene Understanding:**
  Where/who/what/why of everyone else?
- **Movement Planning:**
  How do I get from A to B?
- **Human-Robot Interaction:**
  What is the physical and mental state of the driver?
- **Communicate:**
  How to I convey intent to the driver and to the world?

Blue Text: Easier
Red Text: Harder



A2:
Full Autonomy

- **Localization and Mapping:**
  Where am I?
- **Scene Understanding:**
  Where/who/what/why of everyone else?
- **Movement Planning:**
  How do I get from A to B?
- **Human-Robot Interaction:**
  What is the physical and mental state of the driver?
- **Communicate:**
  How to I convey intent to the driver and to the world?

Full Autonomy (A2) Requires Solving This

Full Autonomy (A2) Requires Solving This

Full Autonomy (A2) Requires Solving This

Full Autonomy (A2) Requires Solving This

Full Autonomy (A2) Requires a Good Reward Function

(that balances driving safety and enjoyment)

Thank You

Next lecture: ***Deep Reinforcement Learning***