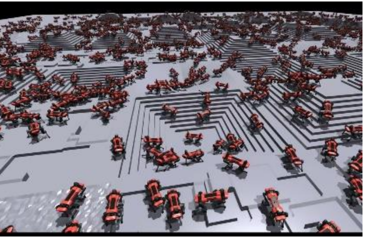
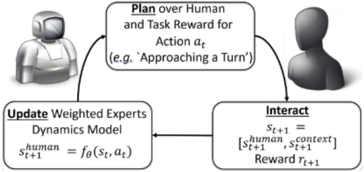
EC518: Robot Learning



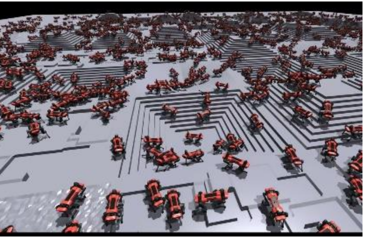
Eshed Ohn-Bar



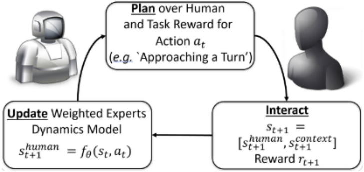
Sep. 11, 2023

EC518: Robot Learning

(and Vision for Navigation)



Eshed Ohn-Bar



Sep. 11, 2023

Last Time

Robot Learning

- Last Time:

- Interactive intelligence, system can ***act and move***

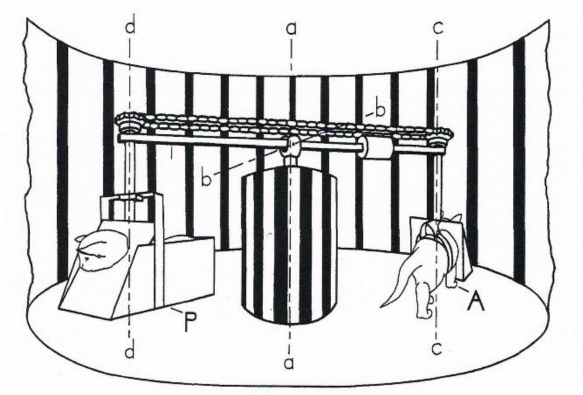
- ***How*** to design this mapping function?

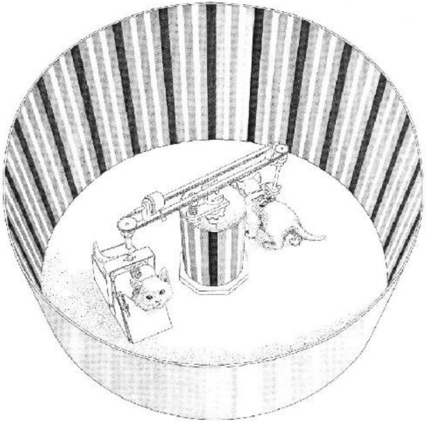
Critique of Pure Vision

Vision, like other sensory functions, has its evolutionary rationale rooted in improved motor control. Although organisms can of course see when motionless or paralyzed, the visual system of the brain has the organization, computational profile, and architecture it has in order to facilitate the organism's thriving at the four Fs: feeding fleeing, fighting, and reproduction. By contrast, a pure visionary would say that the visual system creates a fully elaborated model of the world in the brain, and that the visual system can be studied and modeled without worrying too much about the nonvisual influences on vision.
- Churchland, Ramachandran, and Sejnowski (1994)

https://papers.cnl.salk.edu/PDFs/A%20Critique%20of%20Pure%20Vision%201994-2933.pdf

On the Importance of Interactive Intelligence





Held and Hein: Movement-produced stimulation in the development of visually guided behavior. Journal of Comparative and Physiological Psychology, 1963.

On the Importance of Interactive Intelligence

The ***<u>Gaze Heuristic</u>*** – how to catch a flying ball?

CALCULATE TRAJECTORY:

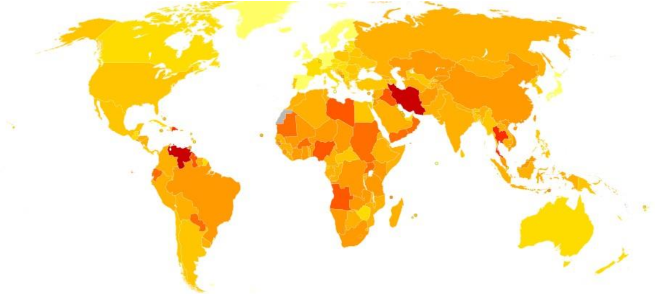
$$z(x) = x\left(\tan\alpha_0 + \frac{mg}{\beta v_0 \cos\alpha_0}\right) + \frac{m^2 g}{\beta^2}\ln\left(1 - \frac{\beta}{m}\frac{x}{v_0 \cos\alpha_0}\right)$$

GAZE HEURISTIC:

1. Fix your gaze on the ball,
2. start running, and
3. adjust your running speed so that the angle of gaze remains constant.

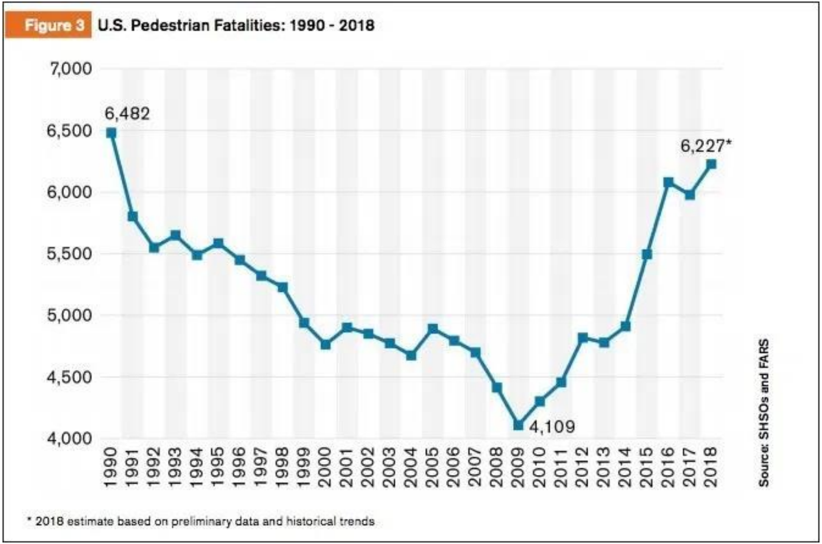
Gigerenzer, "Homo Heuristicus: Why Biased Minds Make Better Inferences“, 2009

Specific Application – Autonomous Driving
Road Fatalities in 2021



- ►World: 1,350,000
- ►Main factors: speeding, intoxination, distraction

U.S. Pedestrian Fatalities Reach Highest Level in 40 Years (in 2022, NHTSA)



Additionally, the traffic fatalities in the following categories showed relatively large increases in 2021, as compared to 2020:

- Fatalities in multi-vehicle crashes up 16%
- Fatalities on urban roads up 16%
- Fatalities among drivers 65 and older up 14%
- Pedestrian fatalities up 13%
- Fatalities in crashes involving at least one large truck up 13%
- Daytime fatalities up 11%
- Motorcyclist fatalities up 9%
- Bicyclist fatalities up 5%
- Fatalities in speeding-related crashes up 5%
- Fatalities in police-reported, alcohol-involvement crashes up 5%

https://www.nhtsa.gov/press-releases/early-estimate-2021-traffic-fatalities

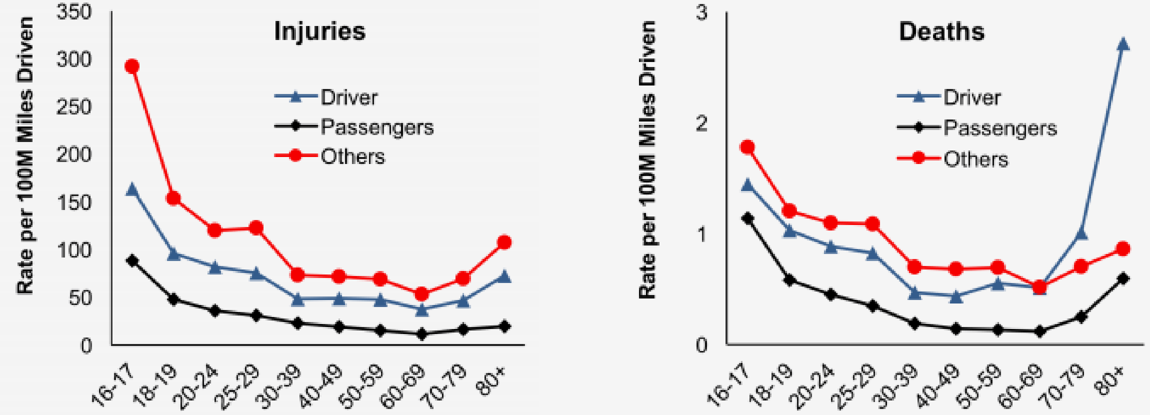


Figure 2. Injuries (left) and deaths (right) in crashes involving a driver of age shown per 100 million miles driven by drivers of that age, by role of person injured or killed, United States, 2014-2015.

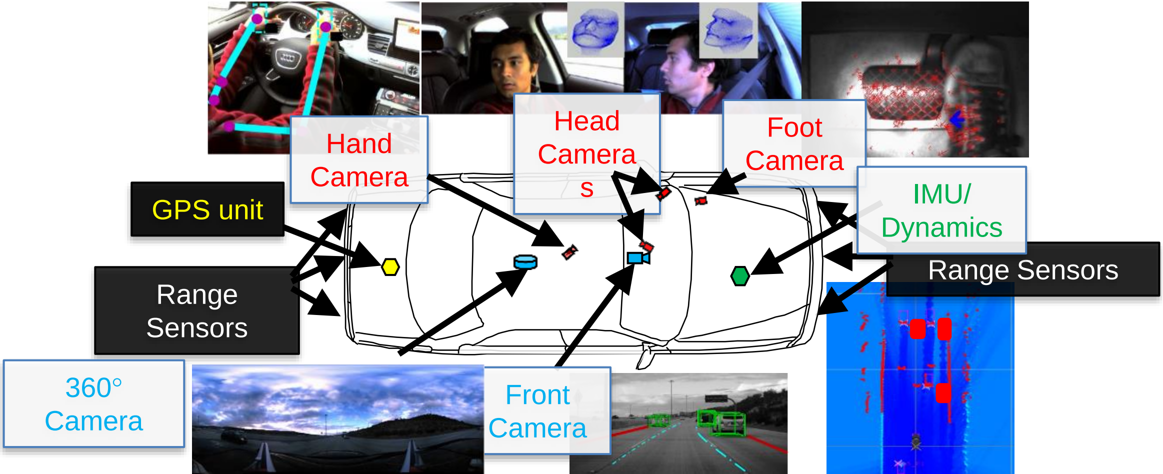
National Highway Traffic Safety Administration

Benefits of Autonomous Driving

- Lower risk of accidents
- Provide mobility for elderly and people with diabilities
- Decrease pollution
- Reduce number of cars (95% of the time a car is parked)

Building an Intelligent System

- **Sense** and **perceive** the world
- **Plan**, predict the world
- **Interact** with environment, humans



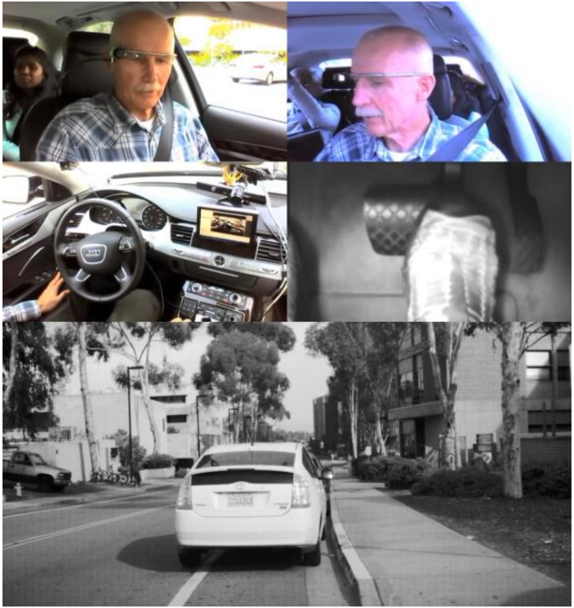
Can You Guess What Is Going to **Happen Next**?

a) Merge Right

b) Merge Left

c) Forward

d) Other



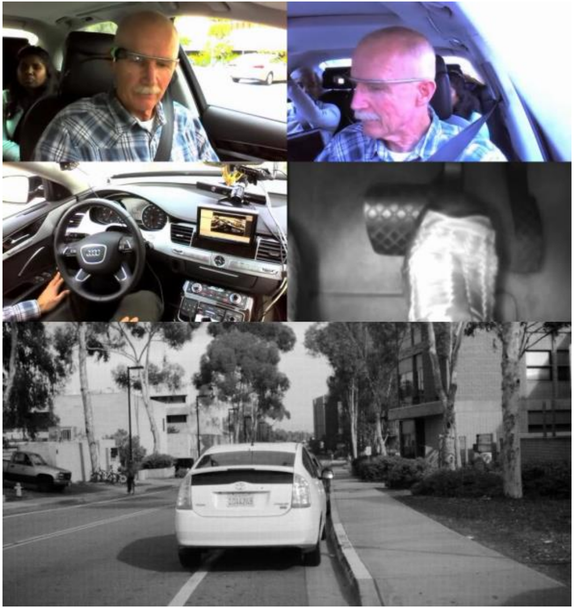
Can You Guess What Is Going to **Happen Next**?

a) Merge Right

b) Merge Left

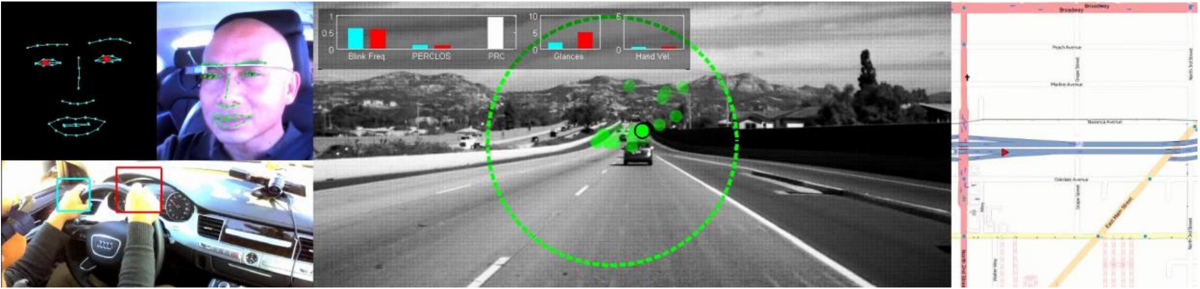
c) Forward

d) Other



Can You Guess What Is Going to **Happen Next**?

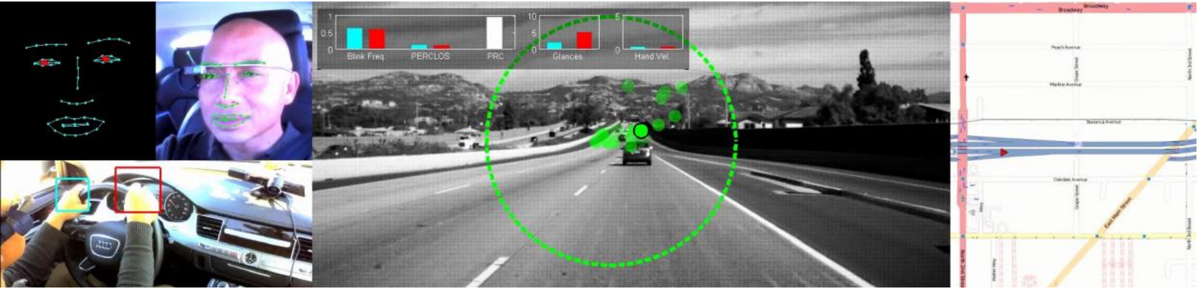
a) Forward   b) Right Change   c) Left Change   d) Other



Ohn-Bar, et al., Predicting Driver Maneuvers, Computer Vision and Image Understanding, 2015

Can You Guess What Is Going to **Happen Next**?

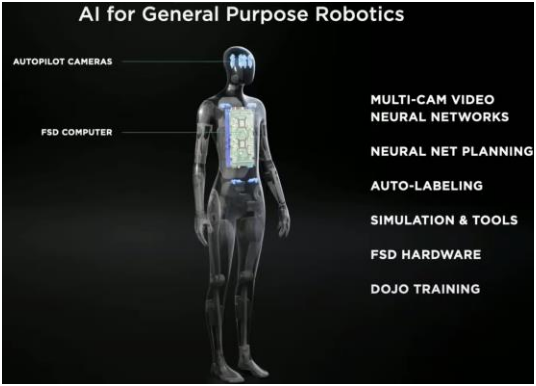
a) Forward   b) Right Change   c) **Left Change**   d) Other



Ohn-Bar, et al., Predicting Driver Maneuvers, Computer Vision and Image Understanding, 2015

Bush
1500
FRANKLIN
1500
FRANKLIN
Franklin
ACURA
coffee
5:28

TeslaBot


AI for General Purpose Robotics
AUTOPILOT CAMERAS
FSD COMPUTER
MULTI-CAM VIDEO
NEURAL NETWORKS
NEURAL NET PLANNING
AUTO-LABELING
SIMULATION & TOOLS
FSD HARDWARE
DOJO TRAINING

How Did We Get Here?

Objective for This Lecture

- Historical Perspective
- Frameworks for Sensorimotor Algorithms

900 AD

1886 – First Automobile

Benz Patent-Motorwagen Nummer 1

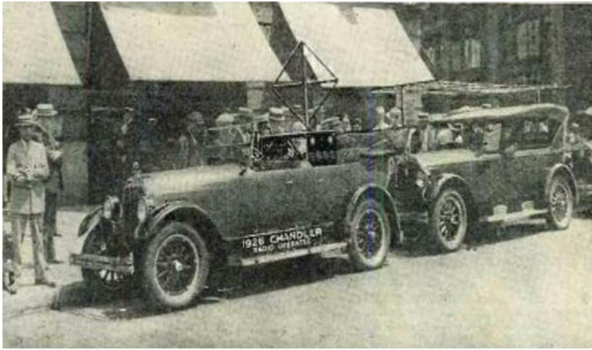
1886: Benz Patent-Motorwagen Nummer 1

- Benz 954 cc single-cylinder four-stroke engine (500 watts)
- Weight: 100 kg (engine), 265 kg (total)
- Maximal speed: 16 km/h
- Consumption: 10 liter / 100 km (!)

1886: Benz Patent-Motorwagen Nummer 1

► First long distance trip (106 km / 66 miles) by ***Bertha Benz*** in 1888 with Motorwagen Nummer 3 (without knowledge of her husband) fostered commercial interest

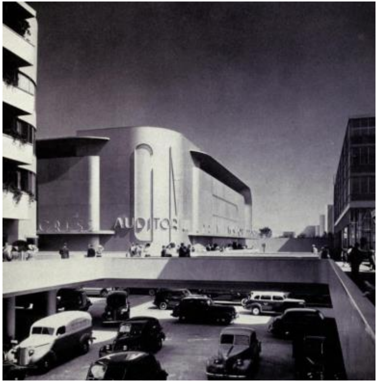
1925: Phantom Auto – "American Wonder" (Houdina Radio Control)



Houdina's driverless car, called the American Wonder, traveled along Broadway in New York City—trailed by an operator in another vehicle—and down Fifth Avenue through heavy traffic. It turned corners, sped up, slowed down and honked its horn. Unfortunately, the demonstration ended when the American Wonder crashed into another vehicle filled with photographers documenting the event.

(Discovery Magazine)

1939: Futurama – New York World's Fair

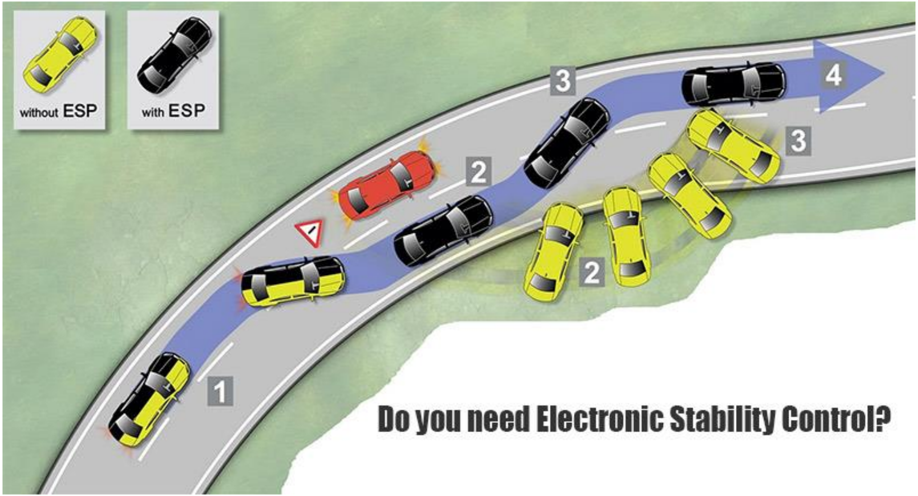




- ►Exhibit at the New York World's Fair in 1939 sponsored by **General Motors**
- ►Designed by Norman Bel Geddes' -his vision of the world 20 years later (1960)
- ►Radio-controlled electric cars, electromagnetic field via circuits in roadway
- ►#1 exhibition, very well received (great depression), prototypes by RCA & GM

1956: General Motors Firebird II


1,450
°F
INLET TEMP
GAS TURBINE

1983-Now: Driver Assistance Systems



ESP – Electronic Stability Program

1983-Now: Driver Assistance Systems

ESC Prevents

- 40% of single vehicle crashes
- 43% of all crash fatalities
- 56% of single vehicle fatalities
- 80% of rollover fatalities

Insurance Institute of Highway Safety - 2006

1986-1994: The invention of the self-driving car

- Developed by Ernst Dickmanns, Longitudinal & lateral guidance with lateral acceleration feedback
- 1678 km autonomous ride Munich to Odense, 95% autonomy (up to 158 km)
- Autonomous driving speed record: 180 km/h (lane keeping)

Dickmanns and Mysliwetz: Recursive 3-D Road and Relative Ego-State Recognition. PAMI, 1992.

1986-1995: Navlab

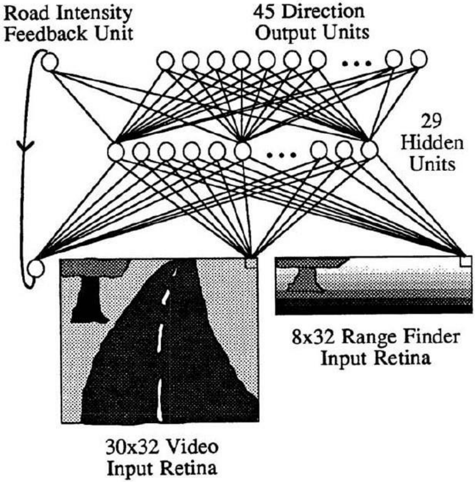
Jochem, Pomerleau, Kumar and Armstrong: PANS: A Portable Navigation Platform. IV, 1995.

1986-1995: Navlab

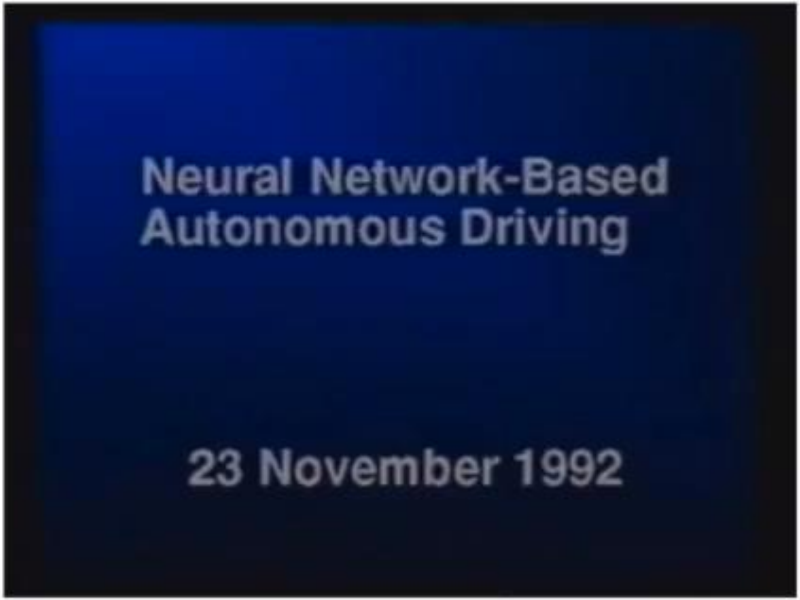
Using computer vision, the **Navlab 5** steered 98.2% of the distance from Washington, D.C. to San Diego CA. (2797 miles out of 2849)

1988: ALVINN An Autonomous Land Vehicle in a Neural Network

- ► Forward-looking, vision based driving
- ► Fully connected neural network maps road images to vehicle turn radius
- ► Directions discretized (45 bins)
- ► Trained on simulated road images
- ► Tested on unlined paths, lined city streets and interstate highways
- ► 90 consecutive miles at up to 70 mph



Pomerleau: ALVINN: An Autonomous Land Vehicle in a Neural Network. NeurIPS, 1988.

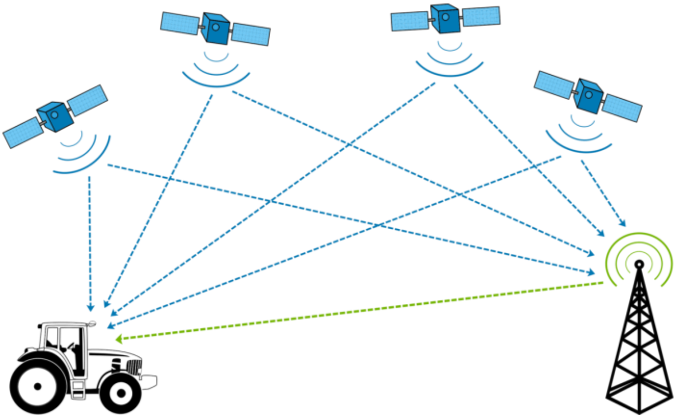


[Courtesy of Dean Pomerleau]

1995: Invention of Adaptive Cruise Control (ACC)

- 1992: Lidar-based distance control by Mitsubishi (throttle control & downshift)
- 1997: Laser adaptive cruise control by Toyota (throttle control & downshift)
- 1999: Distronic radar-assisted ACC by Mercedes-Benz (S-Class), level 1 autonomy

2000: "**Main Breakthrough #1**": GPS, IMUs & Maps



- NAVSTAR GPS available with 1meter accuracy, IMUs improve up to 5 cm
- Navigation systems and road maps available
- Accurate self-localization and ego-motion estimation algorithms

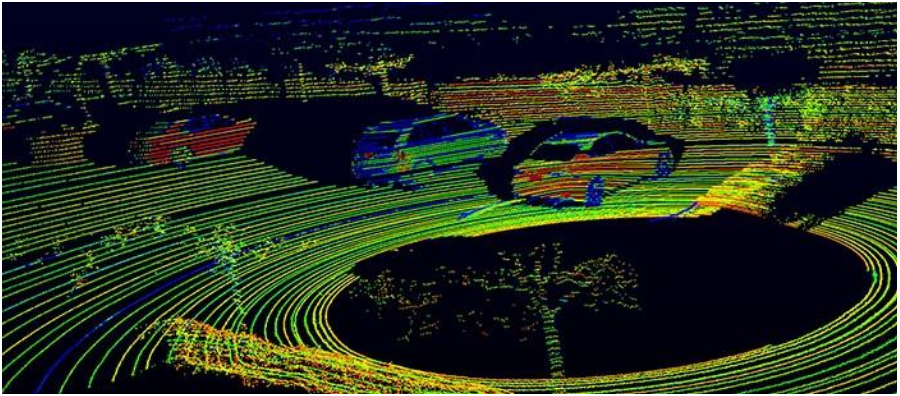
2004: Darpa Grand Challenge 1

- 1st competition in the Mojave Desert along a 240 km route, $1 mio prize money
- No traffic, dirt roads, driven by GPS (2935 points, up to 4 per curve).
- None of the robot vehicles finished the route. CMU traveled the farthest distance, completing 11.78 km of the course **before hitting a rock**.

2005: Darpa Grand Challenge 2

- 2nd competition in the Mojave Desert along a 212 km route, $2 mio prize money
- Five teams finished (Stanford team 1st in 6:54 h, CMU team 2nd in 7:05 h)

2006: **Main Breakthrough #2**: Lidars & High-res Sensors



- High-resolution Lidar
- Camera systems with increasing resolution
- Accurate 3D reconstruction, 3D detection & 3D localization

2007: Darpa Urban Challenge

- 3nd competition at George Air Force Base, 96 km route, urban driving, \$2 mio
- Rules: obey traffic law, negotiate, avoid obstacles, merge into traffic
- 11US teams received \$1 mio funding for their research
- Winners: CMU 1st (4:10), Stanford's Stanley 2nd (4:29).


START

2009: Google starts working on Self-Driving Car

- ► Led by Sebastian Thrun, former director of Stanford AI lab and Stanley team
- ► Others: Chris Urmson, Dmitri Dolgov, Mike Montemerlo, Anthony Levandowski
- ► Renamed "Waymo" in 2016

2012: **Breakthrough #3**: Benchmarks and Methods

Geiger, Lenz and Urtasun: Are we ready for Autonomous Driving? The KITTI Vision Benchmark Suite. CVPR, 2012.

Still Ongoing: New Benchmarks and Methods

Dosovitskiy et. al.: CARLA: An open urban driving simulator. CoRL, 2018.

Still Today: Third Technological Revolution: New Benchmarks and Methods

Dosovitskiy et. al.: CARLA: An open urban driving simulator. CoRL, 2018.

2013: Mercedes Benz S500 Intelligent Drive

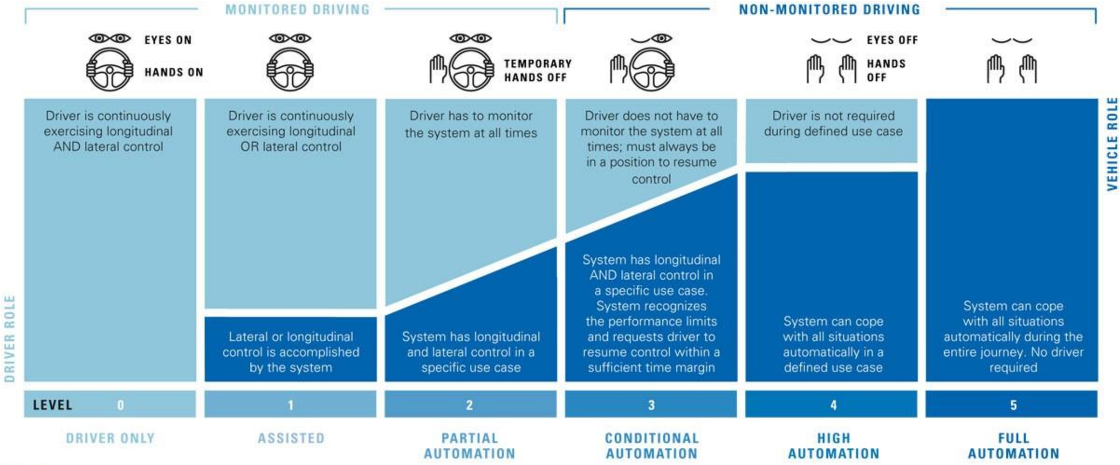
- Autonomous ride on historic Bertha Benz route by Daimler R&D and KIT/FZI
- Novelty: close to production stereo cameras / radar (but also HD maps)

2014: Mercedes S Class

Advanced ADAS (Level 2 Autonomy):

- ► Autonomous steering, lane keeping, acceleration/braking, collision avoidance, driver fatigue monitoring in city traffic and highway speeds up to 200 km/h

2014: Society of Automotive Engineers: SAE Levels of Autonomy

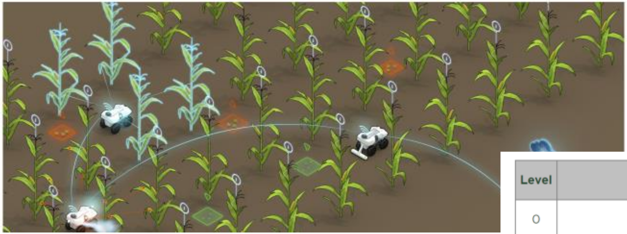


Mike Lemanski

Levels of Autonomy for Field Robots

A framework for decision makers, engineers, and users working on deploying autonomous robots.

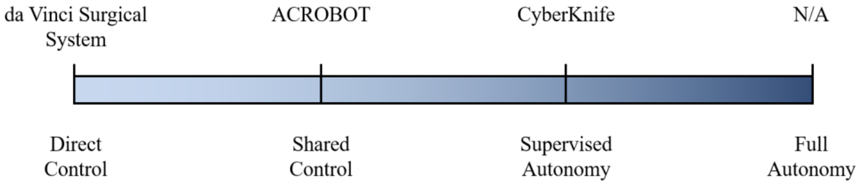
Introduction: Stakeholders in a large variety of industries are considering autonomous robots. However, we are still some time away from achieving reliable and robust long-term autonomy in the real world. Fortunately, even at the current levels of autonomy, robots can be deployed to help with a variety of tasks and deliver significant benefit to end-users across industries. We present a framework that will enable engineers, users, and decision makers to systematically evaluate the autonomy of real-world robotics systems they are considering and decide how they can best benefit from this rapidly improving technology.

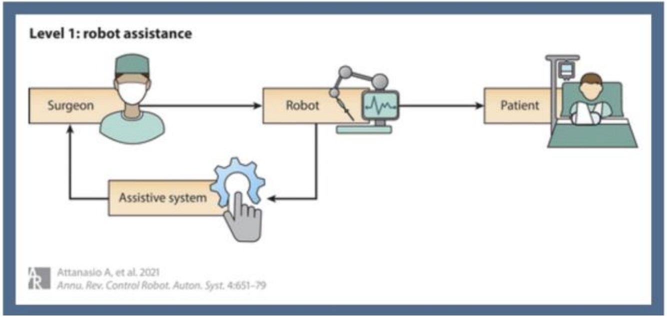


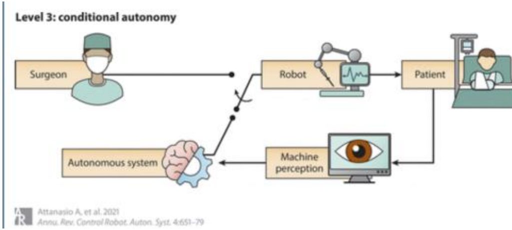
Level	Description	Time between Interventions
0	Full manual teleoperation	n/a
1	Robot within line of sight (hands off)	5 minutes
2	Operator on site or nearby (eyes off)	1 hour
3	One operator oversees many robots (mind off)	8 hours
4	Supervisor not on site (monitoring off)	3 days
5	Robots adapt and improve execution (development off)	extended operation

Table 2. Levels of Decision Making Automation (Sheridan & Verplank, 1978)

Level of Automation	**Description**
1.	The computer offers no assistance; the human must make all decisions and actions
2.	The computer offers no assistance; the human must make all decisions and actions
3.	The computer offers a complete set of decision/action alternatives, or
4.	Narrows the selection down to a few, or
5.	Suggests one alternative
6.	Executes that suggestion if the human operator approves, or
7.	Allows the human a restricted time to veto before automatic execution, or
8.	Executes automatically, then necessarily informs the human, and
9.	Informs the human only if asked, or
10.	Informs the human only if it, the computer, decides to


da Vinci Surgical System
ACROBOT
CyberKnife
N/A
Direct Control
Shared Control
Supervised Autonomy
Full Autonomy


Level 1: robot assistance
Surgeon
Robot
Patient
Assistive system
Attanasio A, et al. 2021
Annu. Rev. Control Robot. Auton. Syst. 4:651–79


Level 3: conditional autonomy
Surgeon
Robot
Patient
Autonomous system
Machine perception
Attanasio A, et al. 2021
Annu. Rev. Control Robot. Auton. Syst. 4:651–79

2015: Tesla Model S Autopilot

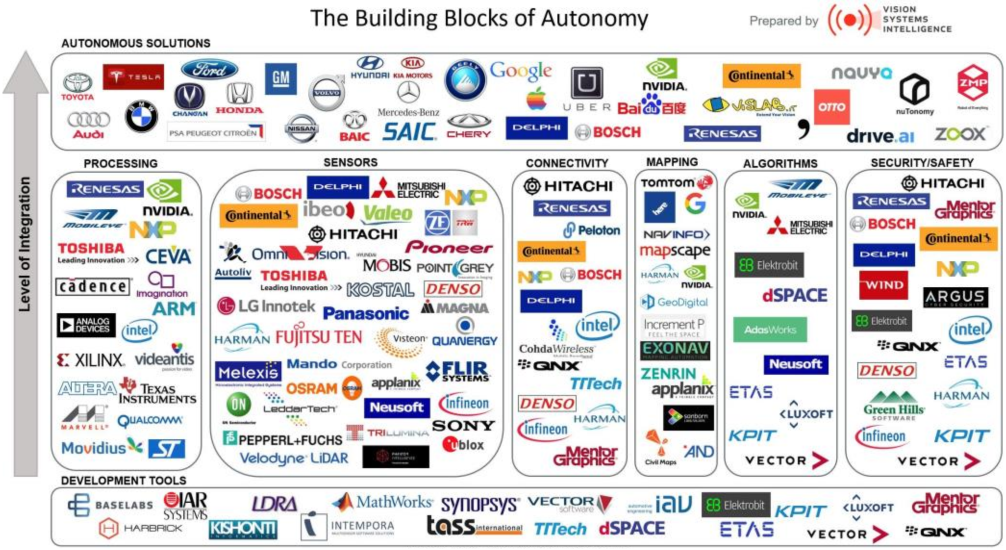
Tesla Autopilot (Level 2 Autonomy):

- ► Lane keeping for limited-access highways (hands off time: 30-120 seconds)
- ► Doesn’t read traffic signals, traffic signs or detect pedestrians/cyclists

2016: Tesla Model S Autopilot: Fatal Accident 1

2018: Tesla Model X Autopilot: Fatal Accident 2

Self-Driving Industry



Business Models

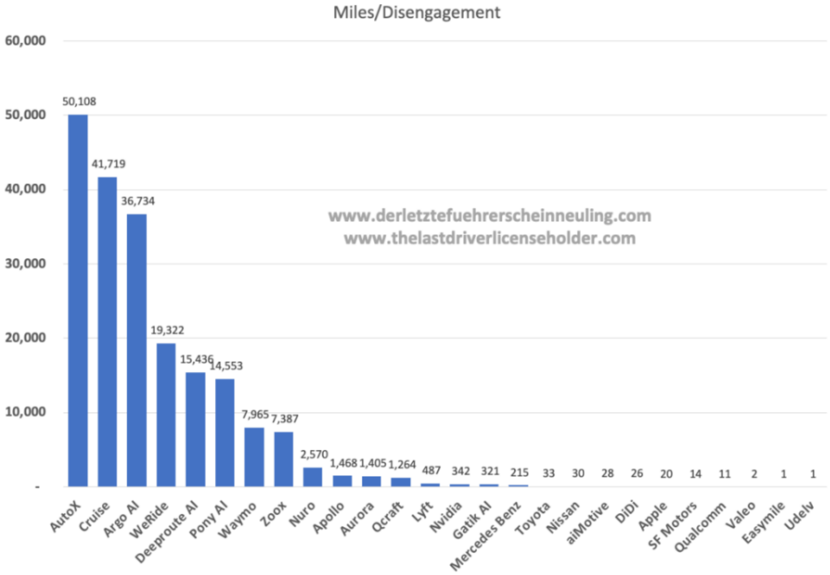
Autonomous or nothing (Google, Apple, Uber)

- Very risky, only few companies can do this
- Long term goals

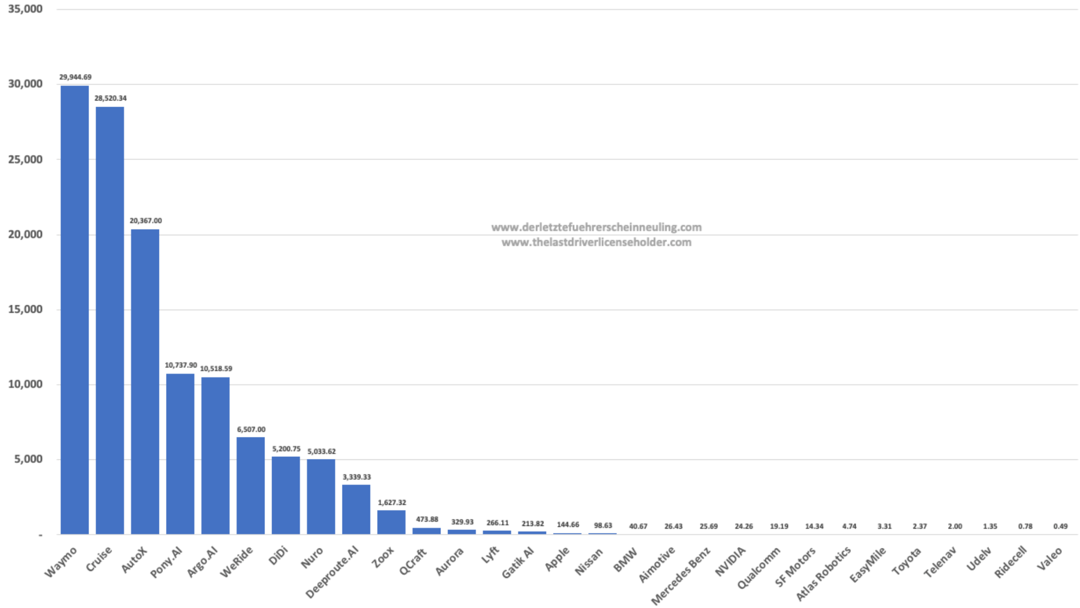
Introduce technology little by little (all car companies)

- Car industry is very conservative
- ADAS as intermediate goal
- Sharp transition: how to maintain the driver engaged?

Miles per disengagement (California Dept. of Motor Veh., 2021)



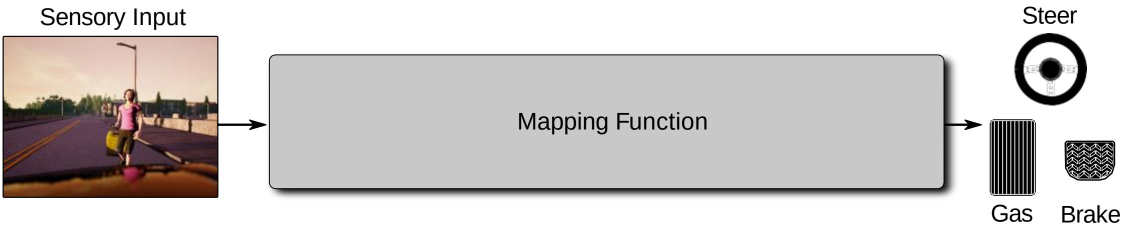
Miles per disengagement (California Dept. of Motor Veh., 2020)



Summary

- ► Self-driving has a ***long history***
- ► Highway lane-keeping of today was developed 30 years ago
- ► Increased robustness ⇒ introduction of ***level 3 for highways in 2019***
- ► Increased interest after DARPA challenge and new benchmarks (e.g., KITTI)
- ► Many claims about full self-driving (e.g., Elon Musk), but level 4/5 stays hard
- ► Waymo seems ahead of competition in fully self-driving, also largest investments
- ► But several setbacks (Uber, Tesla accidents)
- ► Existing systems require laser scanners and HD maps
- ► Key technological challenges remain – autonomous driving is a research problem.

Approaches for Sensorimotor Navigation

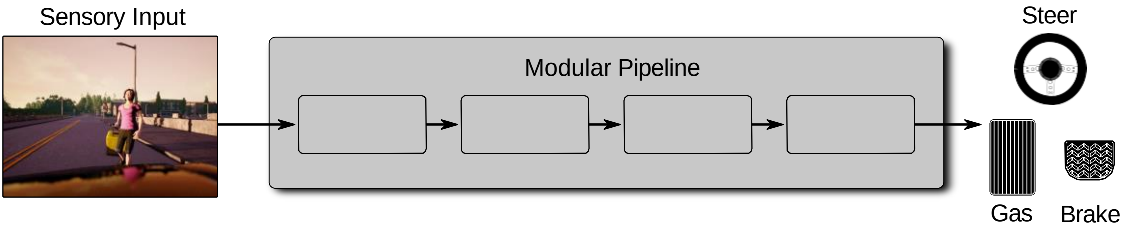


Major Paradigms:

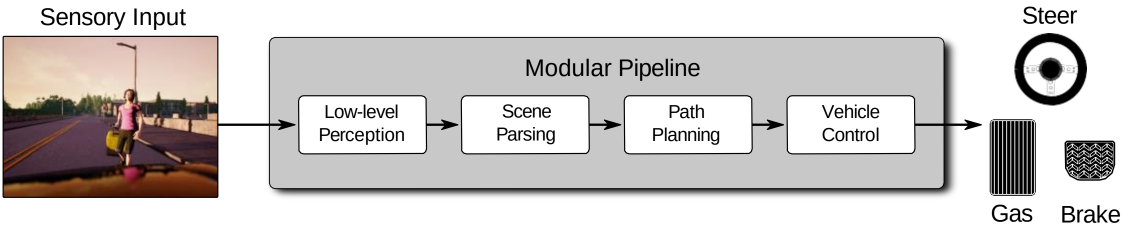
- Modular Pipelines
- End-to-End Learning (Deep Imitation Learning, Reinforcement Learning)
- Direct Perception

Modular Pipeline

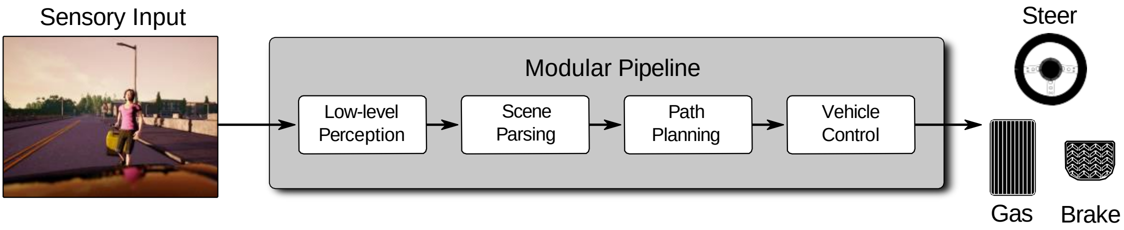
Modular Pipeline



Modular Pipeline



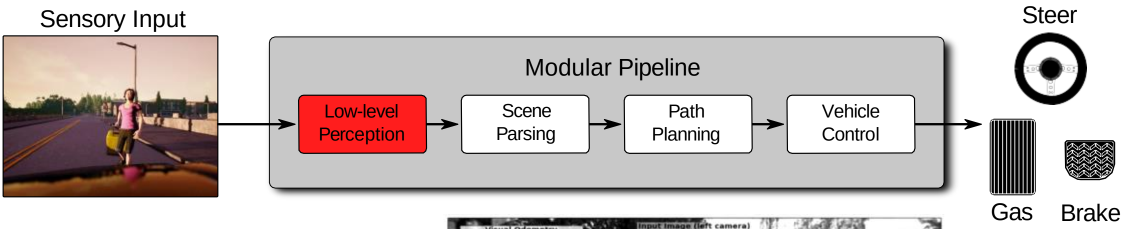
Modular Pipeline

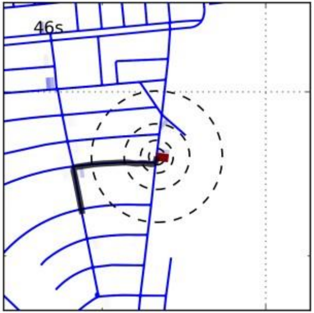


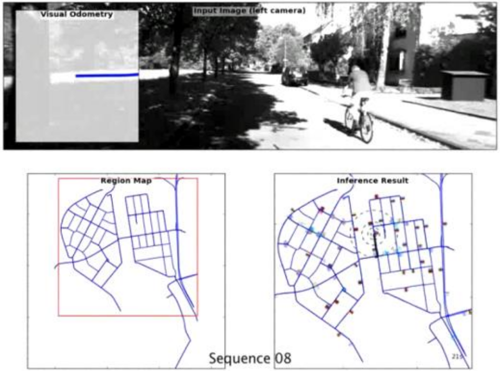
Examples:

- [Montemerlo et al., JFR 2008]
- [Urmson et al., JFR 2008]
- Waymo, Uber, Tesla, Zoox, …

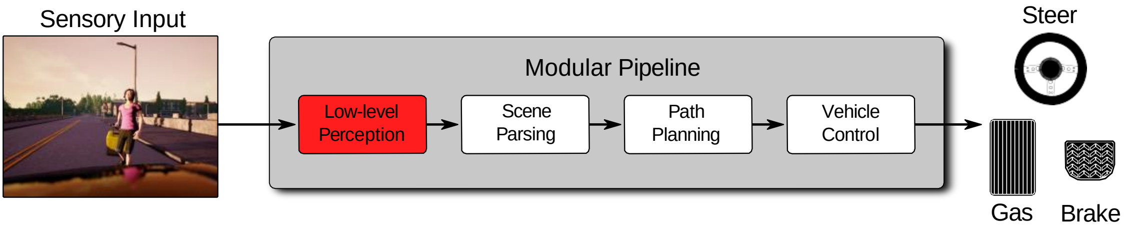
Modular Pipeline



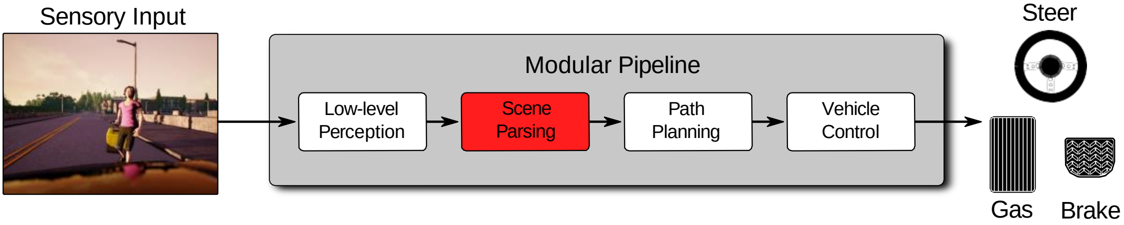


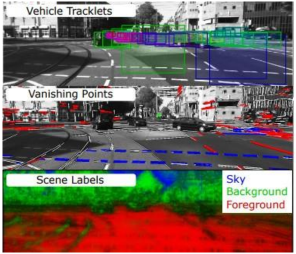


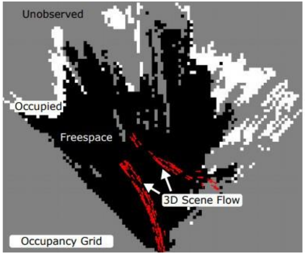
Modular Pipeline

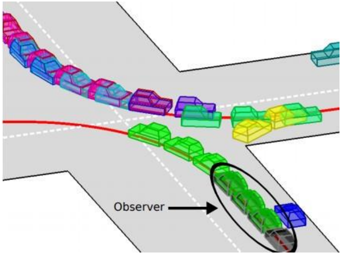


Modular Pipeline

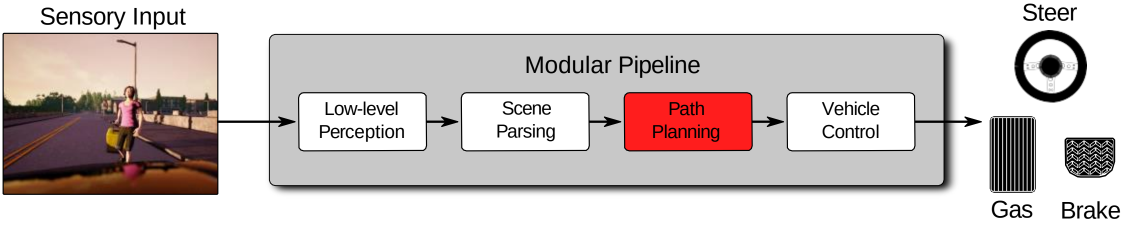


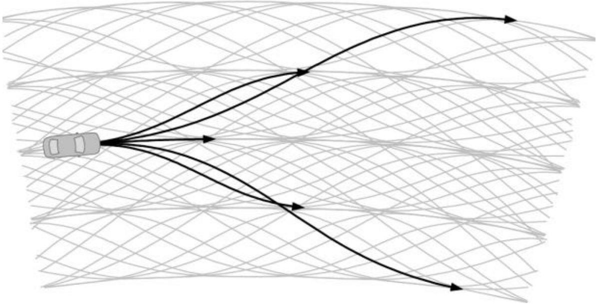




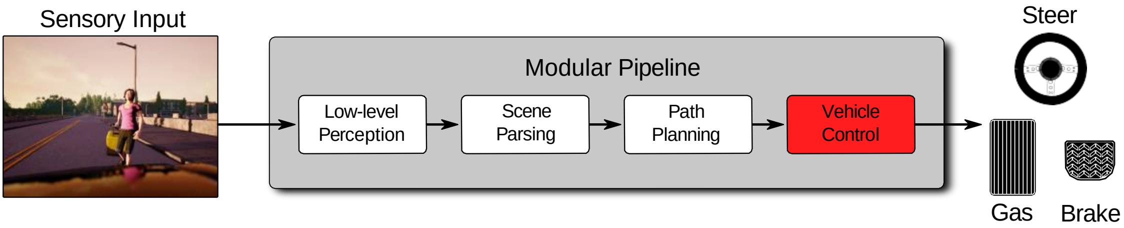


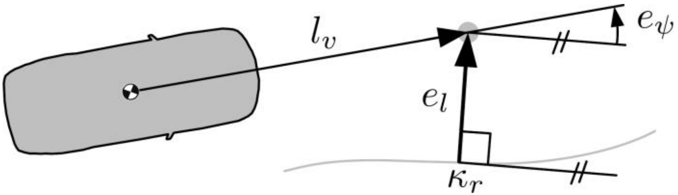
Modular Pipeline



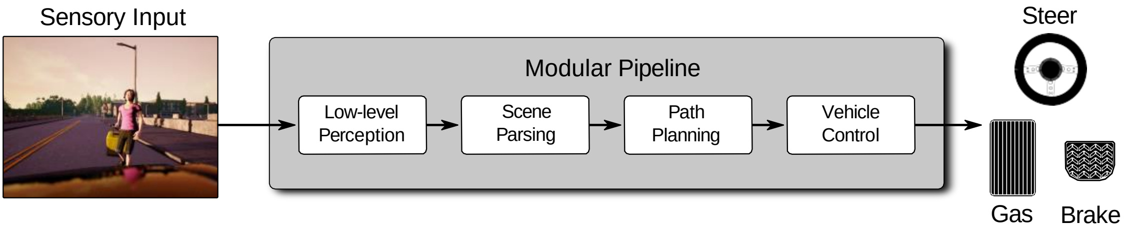


Modular Pipeline





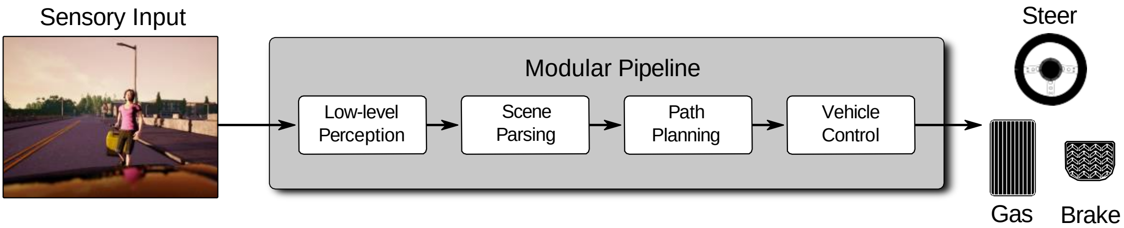
Modular Pipeline



Pros:

Cons:

Modular Pipeline

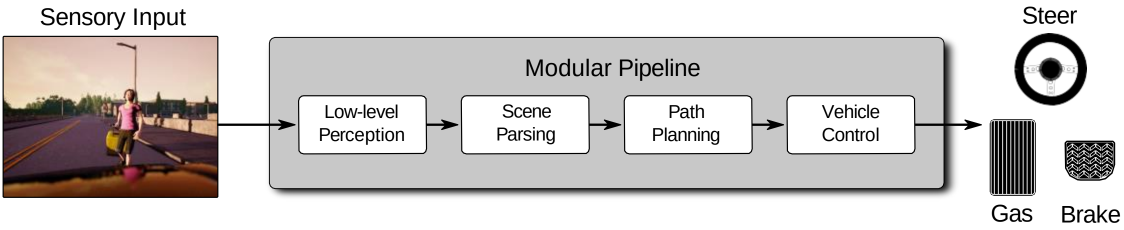


Pros:

- Small components, easy to develop in parallel

Cons:

Modular Pipeline

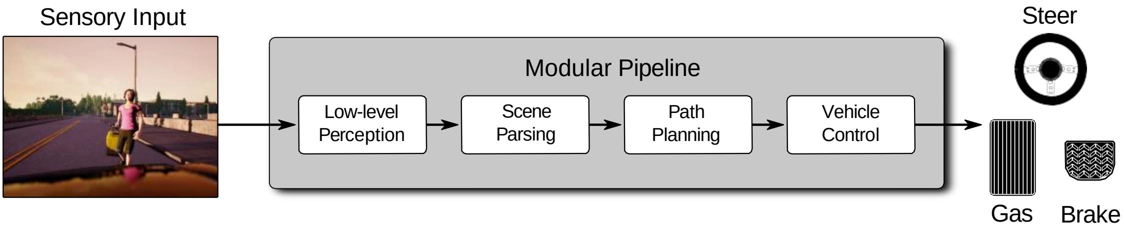


Pros:

- Small components, easy to develop in parallel
- Interpretability

Cons:

Modular Pipeline

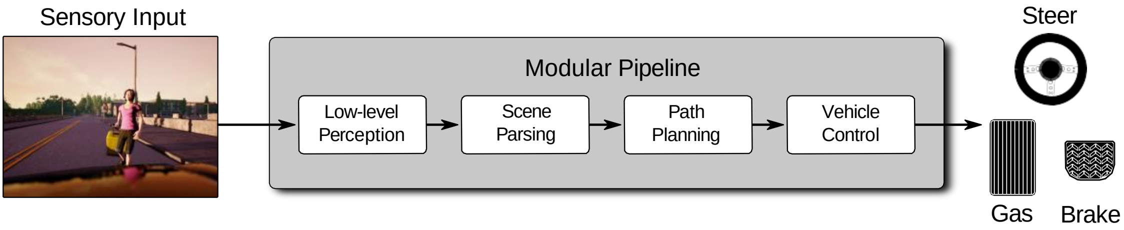


Pros:

- Small components, easy to develop in parallel
- Interpretability
- Engineered bias/prior, inductive structure

Cons:

Modular Pipeline

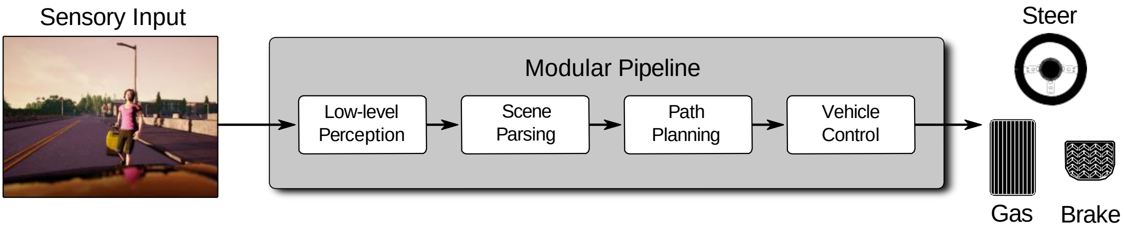


Pros:

- Small components, easy to develop in parallel
- Interpretability
- Engineered bias/prior, inductive structure

Cons:

Modular Pipeline



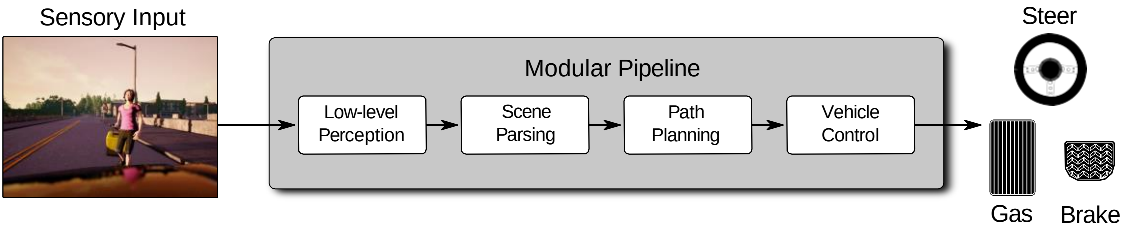
Pros:

- Small components, easy to develop in parallel
- Interpretability
- Engineered bias/prior, inductive structure

Cons:

- Piece-wise training (not jointly)

Modular Pipeline



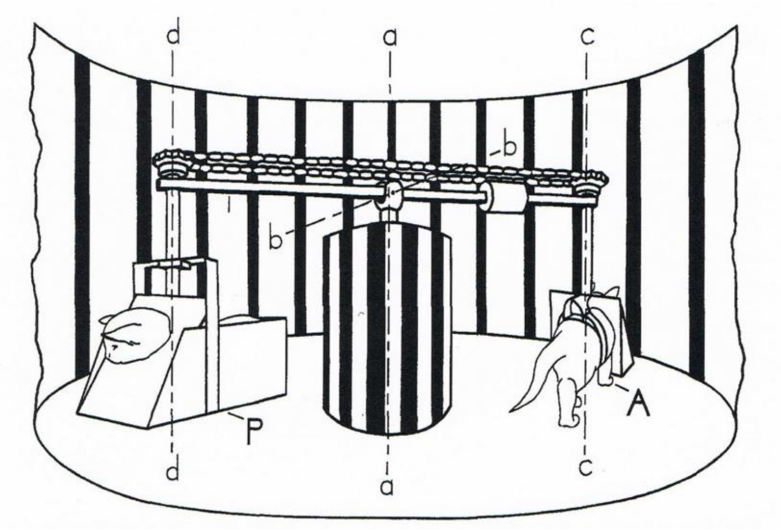
Pros:

- Small components, easy to develop in parallel
- Interpretability
- Engineered bias/prior, inductive structure

Cons:

- Piece-wise training (not jointly)
- Requires tedious and costly annotations

Experiment by Held and Hein: ***Coupling Perception and Action***



Held and Hein: Movement-produced stimulation in the development of visually guided behavior. Journal of Comparative and Physiological Psychology, 1963.

Modular Pipeline

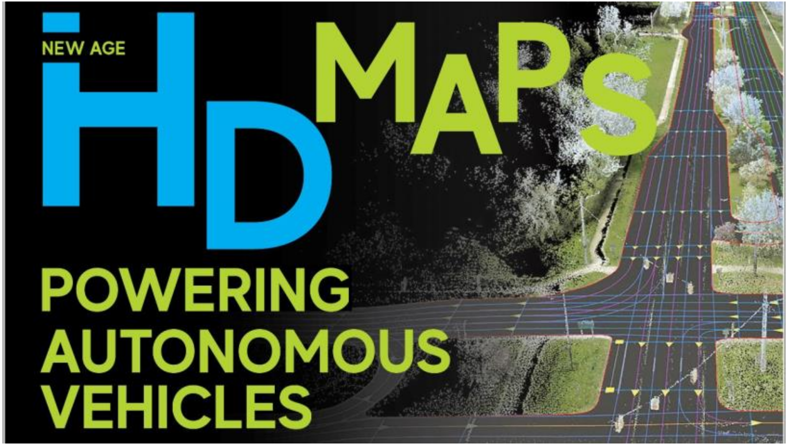
- Piece-wise training dificult: not all objects are equally important!

Ohn-Bar and Trivedi: Are All Objects Equal? Deep Spatio-Temporal Importance Prediction in Driving Videos. PR, 2017.

Modular Pipeline

Cityscapes Dataset

Modular Pipeline

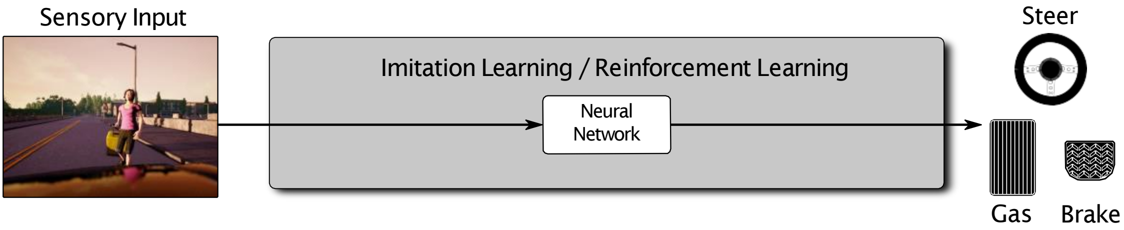


https://www.geospatialworld.net/article/hd-maps-autonomous-vehicles/

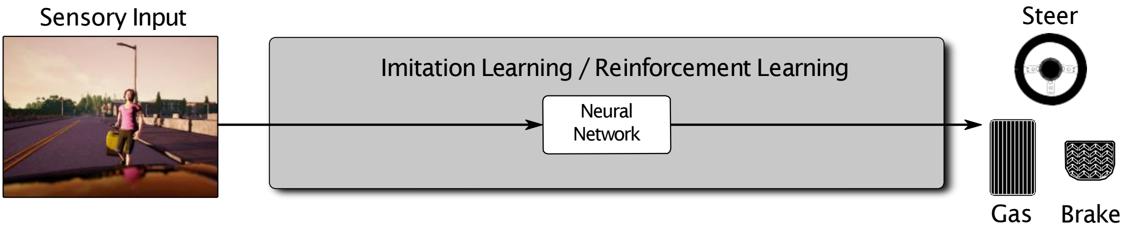
- HD Maps are expensive to create (data collection & annotation effort)

HD maps: Centimeter precision lanes, markings, traffic lights/signs, human annotated

Deep End-to-End Learning



Deep End-to-End Learning for Navigation



Examples:

[Pomerleau, NIPS 1989]

[Bojarski, Arxiv 2016]

[Codevilla et al., ICRA 2018]

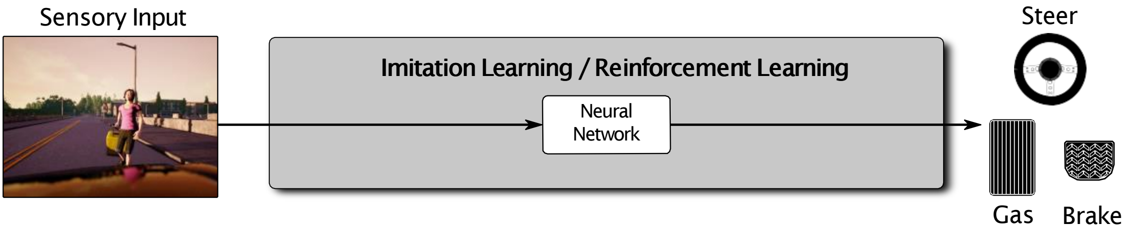

End-to-end Reinforcement learning
CARLA
Speed x3

Robot Motor Skill Coordination with EM-based Reinforcement Learning

Petar Kormushev, Sylvain Calinon, and Darwin G. Caldwell

Italian Institute of Technology

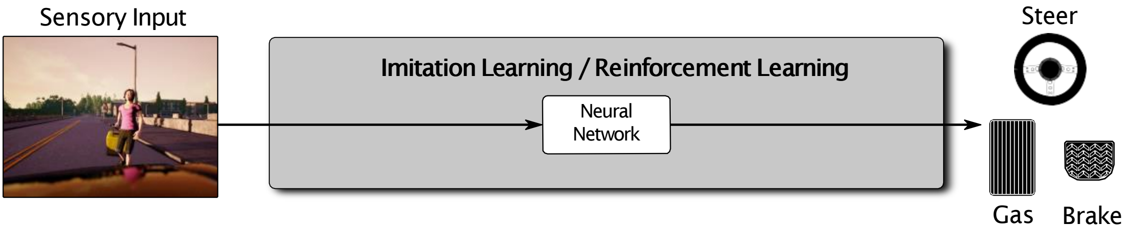
Deep End-to-End Learning for Navigation



Pros:

Cons:

Deep End-to-End Learning for Navigation

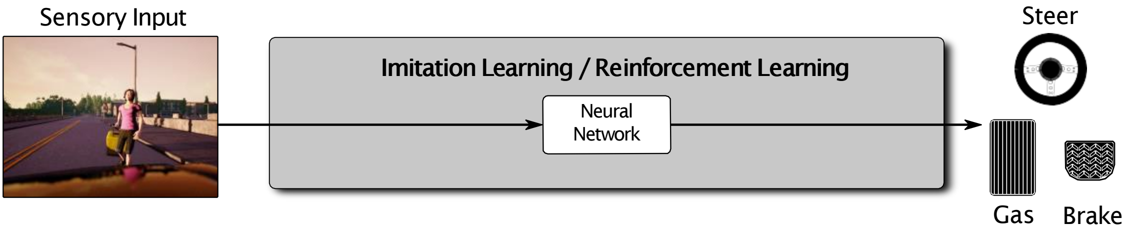


Pros:

- Task-driven training/optimization
- Cheap annotations

Cons:

Deep End-to-End Learning for Navigation



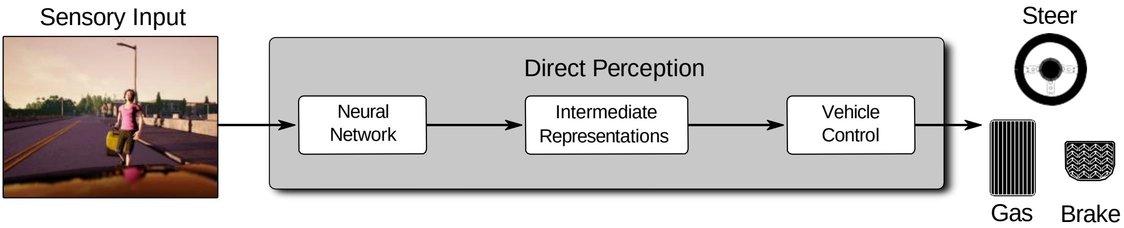
Pros:

- Task-driven training/optimization
- Cheap annotations

Cons:

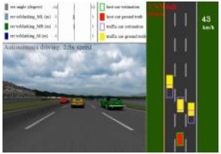
- Overfitting / Generalization
- Interpretability?

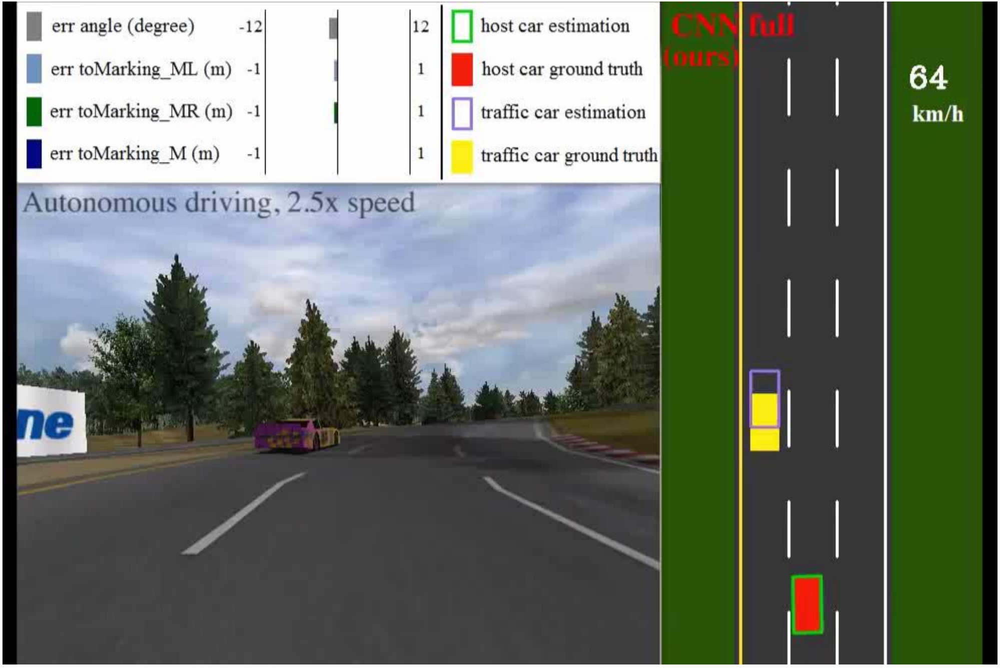
Direct Perception for Navigation



Examples:

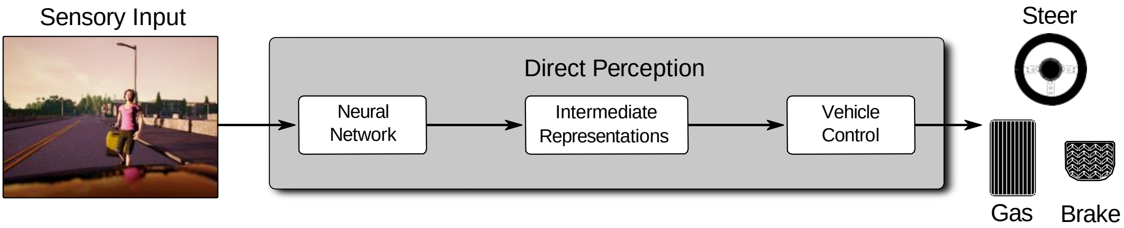
- [Chen et al., ICCV 2015]
- [Sauer et al., CoRL 2018]
- Affordances [J. J. Gibson]




err angle (degree)
-12
12
err toMarking_ML (m)
-1
1
err toMarking_MR (m)
-1
1
err toMarking_M (m)
-1
1
host car estimation
host car ground truth
traffic car estimation
traffic car ground truth
Autonomous driving, 2.5x speed
CNN full
(ours)
64
km/h


FPS: 10
Speed: 0 km/h
Gear: N
Speed Limit: 30 km/h
Traffic Light: Green
Location: (24.410, 202, 3.879)
Orientation: (1.00, -0.00, 0.00)
Acceleration: (-0.00, 0.00, 0.00)
Collision (Cars): 0
Collision (Pedestrian): 0
Collision (Other): 0
Intersection (Lane): 0%
Intersection (OffRoad): 0%

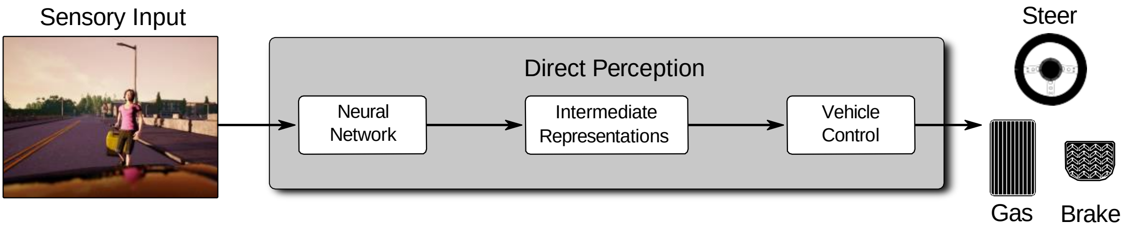
Direct Perception for Navigation



Pros:

Cons:

Direct Perception for Navigation

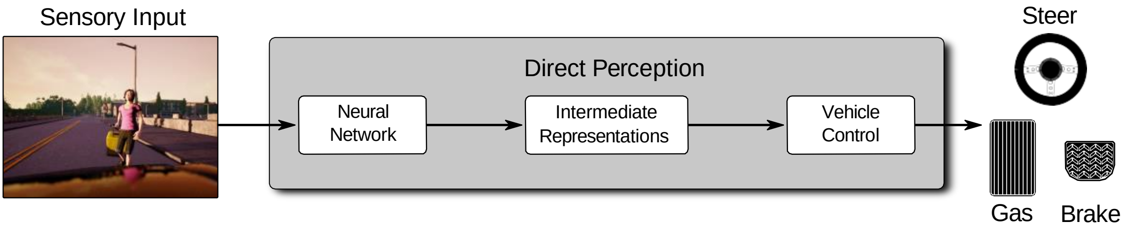


Pros:

- Compact Representation
- Interpretability

Cons:

Direct Perception for Navigation



Pros:

- Compact Representation
- Interpretability

Cons:

- What if representation recognition fails?
- **How to choose representations?**

Further Readings

- Jochem, Pomerleau, Kumar and Armstrong: PANS: A Portable Navigation Platform. IV, 1995.
- Pomerleau: ALVINN: An Autonomous Land Vehicle in a Neural Network. NIPS, 1988.
- Chen, Jochem and Pomerleau: AURORA: A Vision-Based Roadway Departure Warning System. IROS, 1995.
- Dickmanns and Mysliwetz: Recursive 3-D Road and Relative Ego-State Recognition. PAMI, 1992.
- Ernst Dickmanns, "The development of machine vision for road vehicles in the last decade." IEEE Intelligent Vehicles Symposium, 2002
- Geiger, Lenz and Urtasun: Are we ready for Autonomous Driving? The KITTI Vision Benchmark Suite. CVPR, 2012.
- Dosovitskiy, Ros, Codevilla, Lopez, Kolton CARLA: An Open Urban Driving Simulator, CoRL, 2017