

# L'IA appliquée à la conduite le deep learning dans les véhicules autonomes

Stéphane Canu

<https://chaire-raimo.github.io/>  
[scanu@insa-rouen.fr](mailto:scanu@insa-rouen.fr)

61ème édition du congrès annuel du Club EEA



Vendredi 24 juin 2022

# Road map

- 1 Today's vehicles with AI
- 2 Embedded AI
- 3 Data to train the AI
- 4 The future of Artificial Intelligence

ANR RAIMo

Home News Team Publications Contact

Research Chair  
RAIMo

A road toward safe artificial intelligence in mobility

**PROJECT**

*Recent progress in machine learning in general and especially in deep learning make it possible to include this technology in more and more autonomous vehicles. However, before this possible future becomes reality and our roads are made safer with algorithms replacing human drivers, it is necessary to know how to prove the quality of the decisions made.*

This project aims at strengthening local research dynamics about safety issues associated with the use of artificial intelligence in mobility. To achieve this goal, the project will endeavor to formalize the problem, to propose algorithms to solve it and to demonstrate its feasibility

## Autonomous vehicles: when?

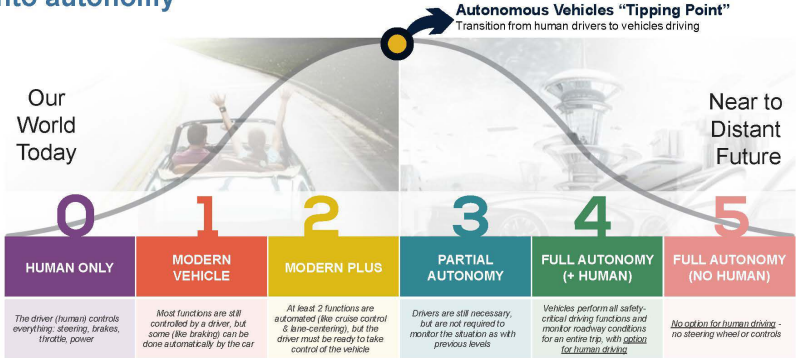
Tesla : prévues en 2014, 2015, 2016, 2018, 2019, 2020, 2021 et 2022, les voitures autonomes sont maintenant promises pour 2023



Un jour, Elon Musk aura raison.

Elon Musk removing his hands from the wheel with Autopilot engaged during an interview (Bloomberg, 2014).  
[numerama.com/vroom/972975-tesla-prevues-en-2014-2015-2016-2018-2019-2020-2021-et-2022-les-voitures-autonomes-sont-maintenant-promises-pour-2023.html](https://numerama.com/vroom/972975-tesla-prevues-en-2014-2015-2016-2018-2019-2020-2021-et-2022-les-voitures-autonomes-sont-maintenant-promises-pour-2023.html)

# ➤ Level 3 Partial Autonomy adoption is when the market "tips" into autonomy



<http://www.techrepublic.com/article/autonomous-driving-levels-0-to-5-understanding-the-differences/>

A COX AUTOMOTIVE BRAND

DRIVER ROLE	MONITORED DRIVING			NON-MONITORED DRIVING	
	Eyes on Hands on	Eyes on Hands off	Temporary Hands off	Eyes off Hands off	Eyes off Hands off
LEVEL 0	Driver is continuously exercising longitudinal AND lateral control	Driver is continuously exercising longitudinal OR lateral control	Driver has to monitor the system at all times	Driver does not have to monitor the system at all times, must always be in a position to resume control	Driver is not required during defined use case
LEVEL 1	Lateral or longitudinal control is recognized by the system	System has longitudinal and lateral control in a specific use case	System has longitudinal AND lateral control for a specific use case. System recognizes the performance limits and requests driver to resume control within a sufficient time margin	System can cope with all situations automatically during the entire journey. No driver required	System can cope with all situations automatically during the entire journey. No driver required
LEVEL 2	ASSISTED	PARTIAL AUTOMATION	CONDITIONAL AUTOMATION	HIGH AUTOMATION	FULL AUTOMATION

# Level 2/3 = ADAS Ratings

## Consumer Reports' for major Advanced Driver Assistance Systems (2020)

SYSTEM NAME	SCORE	CAPAB. & PERF.	KEEPING DRIVER ENGAGED	EASE OF USE	CLEAR WHEN SAFE TO USE	UNRE-SPONSIVE DRIVER
<b>Comma Two</b> Open Pilot	78	8	9	8	6	8
<b>Cadillac</b> Super Cruise	69	8	7	3	8	9
<b>Tesla</b> Autopilot	57	9	3	7	2	6
<b>Ford/Lincoln</b> Co-Pilot 360	52	8	4	3	4	5
<b>Audi</b> Driver Assistance Plus	48	8	3	3	2	6
<b>Mercedes-Benz</b> Driver Assistance	46	6	4	4	2	5
<b>Subaru</b> Eyesight	46	7	4	3	4	5
<b>Hyundai</b> Smart Sense, <b>Kia</b> Drive Wise	46	5	4	5	4	4
<b>BMW</b> Active Driving Assistance Pro	44	7	3	3	2	6
<b>Porsche</b> Active Safe	41	4	3	6	2	5
<b>Volvo</b> Pilot Assist	41	6	3	3	2	5
<b>Toyota/Lexus</b> Safety Sense 2.0	40	5	4	2	4	5
<b>Honda/Acura</b> Sensing	40	6	4	2	4	4
<b>Nissan/Infiniti</b> ProPILOT Assist	40	5	3	3	4	7
<b>Volkswagen</b> Driver Assistance	39	4	3	6	2	5
<b>Land Rover</b> Driver Assist	38	4	3	6	2	4
<b>Buick/Chevy</b> Driver Confidence	36	3	3	5	2	6
<b>Mazda</b> i-ACTIVSENSE	27	3	2	5	2	1
















Some players:

- OpenPilot (open source 50 k)
- Super Cruise (Cadillac 110 k)
- AutoPilot (Tesla, 2M)
- Mobil Eye (54 M)

# Autonomous vehicle performance ranking

## The Self-Driving Car Companies Going The Distance

Number of autonomous test miles and miles per disengagement (Dec 2019-Nov 2020)\*

			Miles	Miles per disengagement	
Waymo (Alphabet)			628,839		29,945
Cruise (GM)			770,049		28,520
AutoX			40,734		20,367
Pony.AI			225,496		10,738
Argo.AI (Ford, VW)			21,037		10,519
WeRide			13,014		6,507
DiDi Chuxing			10,401		5,201
Nuro			55,370		5,034



\* Cases where a car's software detects a failure or a driver perceived a failure, resulting in control being seized by the driver.

Source: DMV California, via The Last Driver License Holder



Some player:

- Waymo (Google)
- Cruise (GM)
- Apollo (Baidu)
- ...

Related initiatives:

- La stratégie nationale de développement de la mobilité routière automatisée
- L3 Pilot (European project)
- ...

# Two kind of AI systems for cars

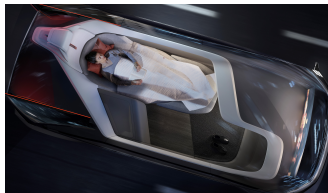
## 1 Driver assistance      Driver is responsible

- ▶ Level 2/3 autonomy
- ▶ Specific intelligence
- ▶ it works: how many seconds for take-over?



## 2 Full Autonomous driving      Car is responsible

- ▶ Level 4/5 autonomy
- ▶ Generic Intelligence
- ▶ Experience level: it doesn't scale yet!



Lex Fridman long term vision

When will we have more than 10,000 Full Autonomous cars?

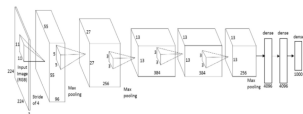
# Road map

1 Today's vehicles with AI

2 Embedded AI

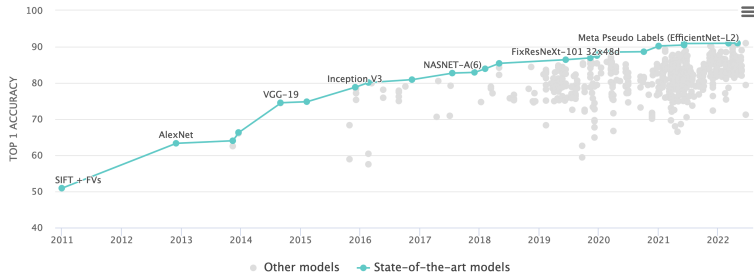
3 Data to train the AI

4 The future of Artificial Intelligence for driving





# ImageNet results: from 50% to 91%



2012 Alex Net

2014 VGG

2015 GoogLeNet / Inception

2016 Residual Network

2018 NAS Network

2020 EfficientNet (Transformers)

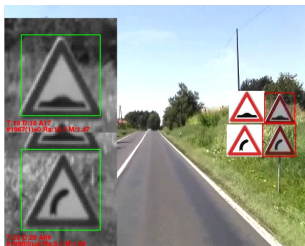
2022 CoCa (Foundation model = image+text)

# Detection, tracking and recognition of traffic signs (2011-13)

**Recognition** German Traffic Sign  
Recognition Benchmark  
(GTSRB) data set, containing  
51839 labelled images of  
real-world traffic signs.



**Detection** The German Traffic Sign  
Detection Benchmark is a  
single-image detection  
assessment 900 images (600 for  
training and 300 for test)



and the winner is

→ Deep learning gives very good results on both tasks

# Open Pilot: 2200 \$



**openpilot**

openpilot is open source software built to improve upon the existing driver assistance in most new cars on the road today. Tesla Autopilot like functionality for your Toyota, Honda, and more.

See openpilot in action >

**Works with the push of a button.**

openpilot is simple to use. It enables your car to steer, accelerate,

## openpilot is the Android



openpilot

Join GitHub today

open source driving agent

1.0M downloads

1.2k stars

1.2k forks

1.2k issues

1.2k pull requests

1.2k commits

1.2k contributors

1.2k followers

1.2k repositories

1.2k organizations

1.2k teams

1.2k projects

1.2k packages

1.2k discussions

1.2k code snippets

1.2k gists

1.2k wikis

1.2k pages

1.2k assets

1.2k avatars

1.2k emojis

1.2k stickers

1.2k themes

1.2k fonts

1.2k icons

1.2k images

1.2k videos

1.2k audio

1.2k documents

1.2k spreadsheets

1.2k presentations

1.2k code files

1.2k data files

1.2k text files

1.2k images

1.2k videos

1.2k audio

1.2k documents

1.2k spreadsheets

1.2k presentations

1.2k code files

1.2k data files

1.2k text files

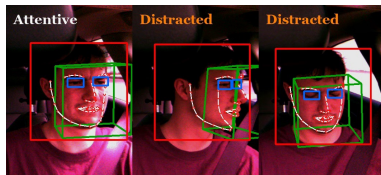
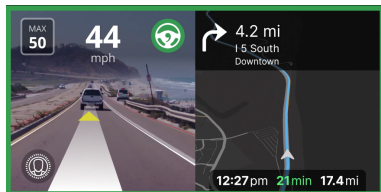
<https://github.com/commaai/openpilot>

THIS IS ALPHA QUALITY SOFTWARE FOR RESEARCH PURPOSES ONLY. THIS IS NOT A PRODUCT. YOU ARE RESPONSIBLE FOR COMPLYING WITH LOCAL LAWS AND REGULATIONS.

# Openpilot AI features

## Two AI

- Diving agent
  - ▶ Automated lane-centering
  - ▶ Adaptive cruise control
  - ▶ OpenStreetMap inside
  - ▶ Assisted lane change
- Driver monitoring system (DMS)
  - ▶ Safety concerns



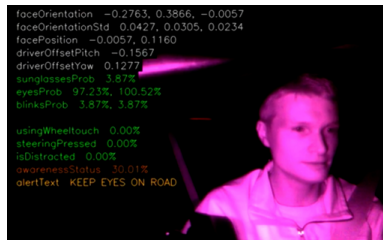
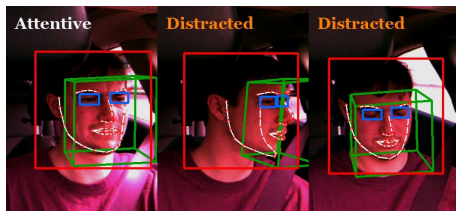
software update

<https://comma-ai.medium.com/towards-a-superhuman-driving-agent-1f7391e2e8ec>

# Openpilots driver monitoring system (DMS)

## Three components

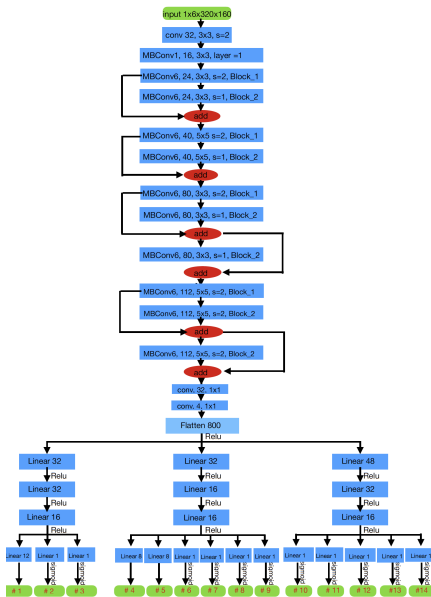
- Face localization
  - ▶ openCV -> cropping
- Feature generation
  - ▶ EfficientNet b0 architecture
  - ▶ Fine tuning
- Decision module
  - ▶ Threshold based decision



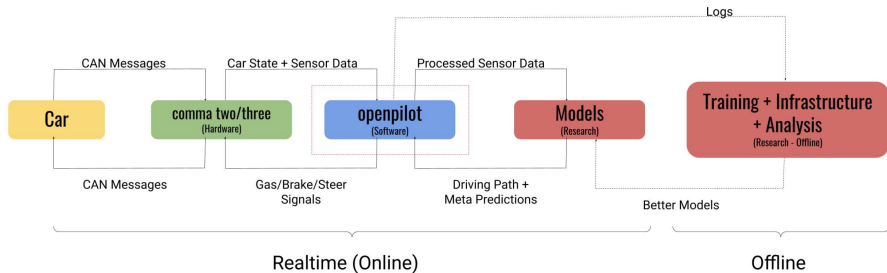
<https://github.com/commaai/openpilot>

# Architecture of the feature generator of openpilot's DMS

- Input: YUV 420 (6 channels)
  - ▶ EfficientNet b0 architecture
  - ▶ Tan et. al. (Google), ICML 2019
- Output: 45-features (03/22)
  - ▶ Face position (12 values)
  - ▶ Eyes positions (8 values)
  - ▶ sunglasses
  - ▶ visible face probability
  - ▶ blinking
  - ▶ ...
- Training data: fine tuning
  - ▶ pytorch inside
  - ▶ Qualcomm Snapdragon 845



# Openpilot's components



AI inside: uses data

(Deep) learning based programming

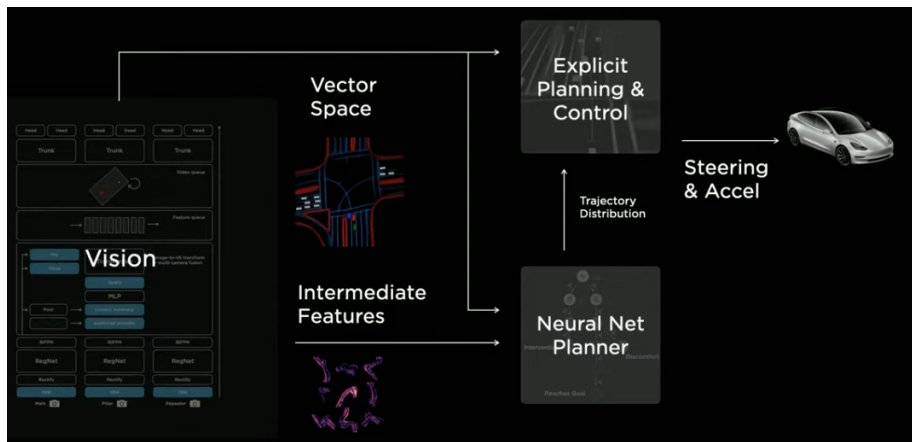
# Tesla's autopilot components

- Driving agent
  - ▶ Automatic lane change
  - ▶ Adaptive cruise control
  - ▶ Autosteer
  - ▶ Navigate on Autopilot (Freeway)
  - ▶ Traffic Light and Stop Sign Control
  - ▶ ...
  - ▶ FSD (limited-access Beta)
- Parking Summon
- Driver monitoring system (DMS)





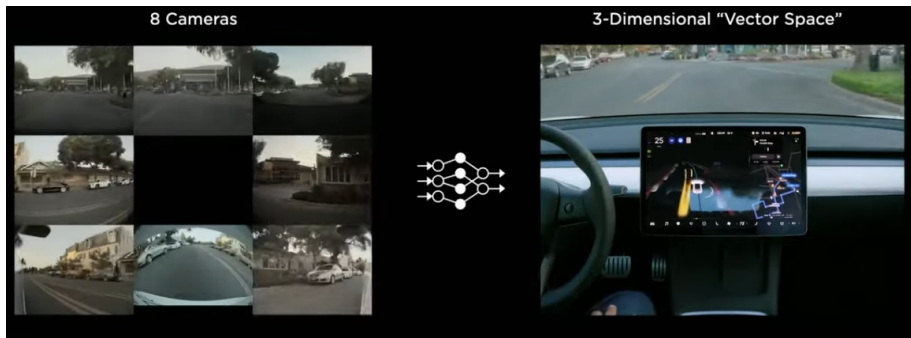
# Summarizing the driving agent architecture



Two AI components = two deep networks

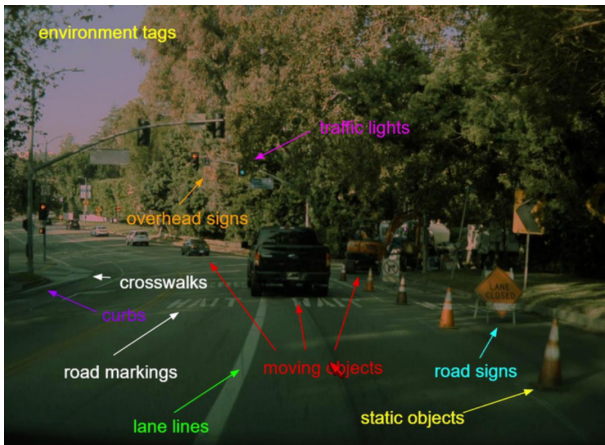
- perception module
- decision module (planner) using deep reinforcement learning

# Tesla's autopilot perception module



- input: 8 cameras
- output:  $640 \times 460$  3D map of the surroundings

# Perception is scene understanding



Scene understanding is

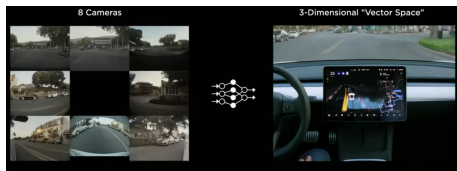
Multi-task learning

# The 5 components of Tesla's perception module

input: 8 cameras

- 1 feature generator: backbone
- 2 multi scale feature fusion
- 3 multi camera fusion
- 4 time filtering
- 5 multi task decision module per pixel on the output map (one per task)
  - ▶ item location (cars, pedestrian...)
  - ▶ traffic signs (Stop sign, traffic light...)
  - ▶ lane prediction
  - ▶ ...

output:  $640 \times 460$  3D map of the surroundings



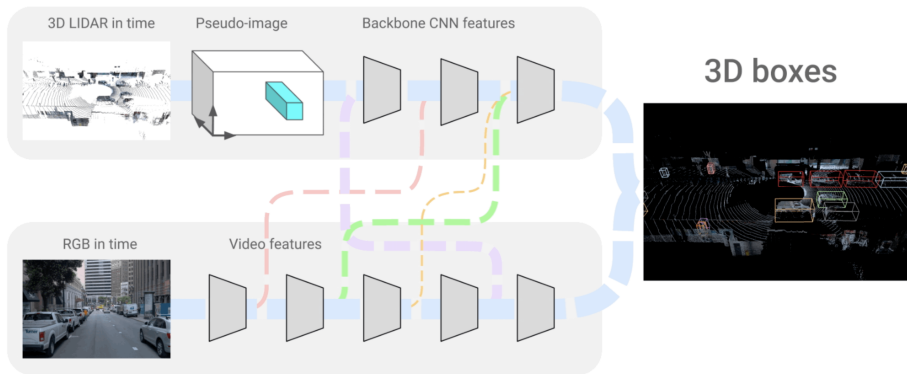
## Tesla perception module

- 1 feature generator: backbone  
ResNet50 (2020), RegNet (2021)  
(from a CVPR 2020 Facebook paper)
- 2 multi scale feature fusion  
EfficientDet  
(from a 2019 Google paper)
- 3 multi camera fusion  
Transformers  
(from a 2020 Facebook paper)
- 4 time filtering  
LSTM (recurrent neural network)
- 5 multi task decision module  
Hydranet
  - ▶ item location (cars, pedestrian...)
  - ▶ traffic signs (Stop sign, traffic light...)
  - ▶ lane prediction
  - ▶ ...

### This perception module contains

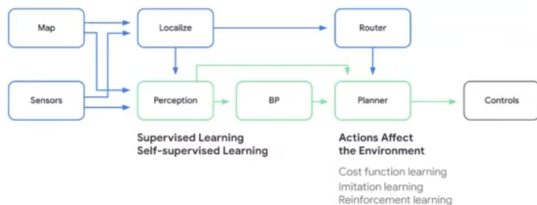
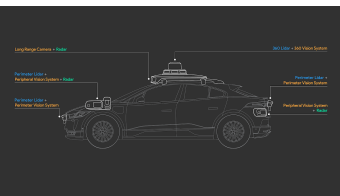
- 1 48 networks, 1,000 outputs tensors, 70,000 GPU h to train
- 2 performs 40 prediction per second

# Perception module at Waymo



"4D-Net for Learned Multi-Modal Alignment", ICCV 2021  
<https://ai.googleblog.com/2022/02/4d-net-learning-multi-modal-alignment.html>

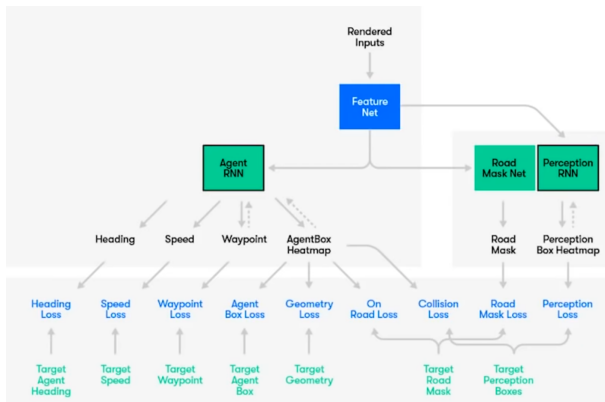
# Waymo's point of view



## Active research

- Stinet: Spatio-temporal-interactive network for pedestrian detection and trajectory prediction, CVPR 2020
- Vectornet: Encoding hd maps and agent dynamics from vectorized representation, CVPR 2020
- Taskology: Utilizing Task Relations at Scale, CVPR 2021
- ChauffeurNet: Learning to Drive by Imitating the Best and Synthesizing the Worst, ICML 2019

# Decision making using deep reinforcement learning



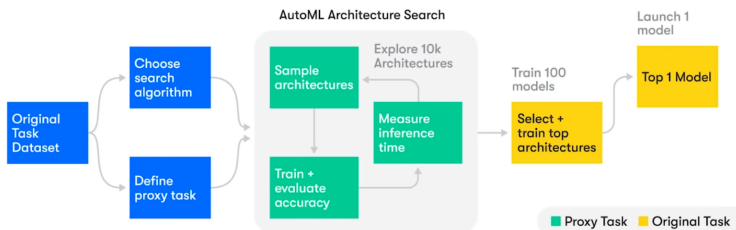
Imitation model providing safety, confort and efficiency

Multi-Task and multi objective learning



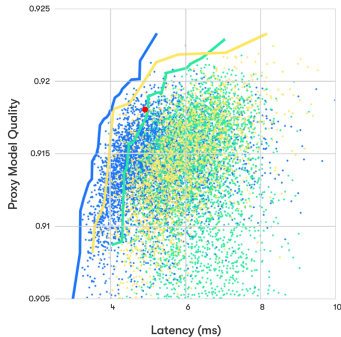
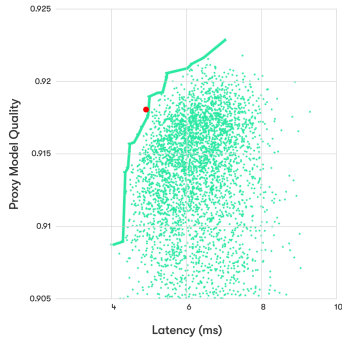
# Waymo's AutoML

## End-to-end architecture search



Proxy end-to-end search: Explore thousands of architecture on a scaled-down proxy task, apply the 100 best ones to the original task, validate and deploy the best of the best architectures on the car.

# Waymo's AutoML



- 1) The first graph shows about 4,000 architectures discovered with a random search on a simple set of architectures. Each point is an architecture that was trained and evaluated. The solid line marks the best architectures at different inference time constraints. The red dot shows the latency and performance of the net built with transfer learning. In this random search, the nets were not as good as the one from transfer learning. 2) In the second graph, the yellow and blue points show the results of two other search algorithms. The yellow one was a random search on a refined set of architectures. The blue one used reinforcement learning as in [1] and explored more than 6,000 architectures. It yielded the best results. These two additional searches found nets that were significantly better than the net from transfer learning.

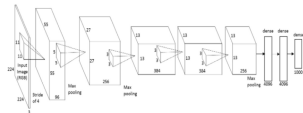
# Road map

1 Today's vehicles with AI

2 Embedded AI

3 Data to train the AI

4 The future of Artificial Intelligence for driving



# Data: the long tail of situations

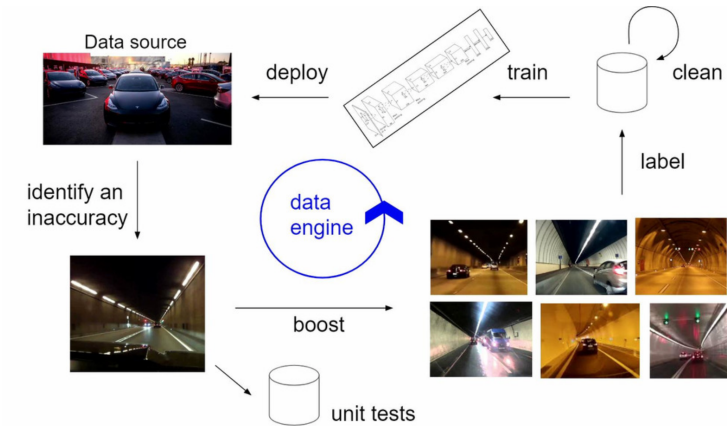


Taiwan, June 2020,



Andrej Karpathy - AI for Full-Self Driving at Tesla, Scaled ML, Feb 2020,

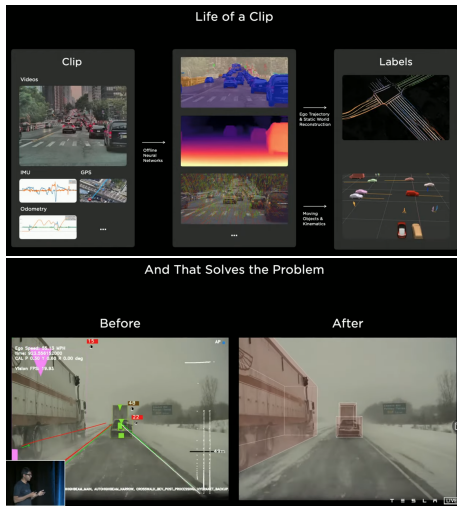
# Improving the autopilot: iterative process



- fleet learning
- testing = shadow mode for more training data

# Tesla's point of view on data

- Gathering process
  - ▶ 221 triggering situations
- manual labelling (1000 person)
  - ▶ 2d -> 3d
  - ▶ reconstruction labelling
- auto labelling
  - ▶ use specifically trained networks
  - ▶ human to clean
- simulation
  - ▶ rare event
  - ▶ sensor robustness
  - ▶ adversarial examples



Tesla's AI day [youtube.com/watch?v=j0z4FweCy4M](https://www.youtube.com/watch?v=j0z4FweCy4M)

# Openpilot : l'étiquetage des données par crowd sourcing

## comma10k

Count and Percentage of Available Images Labeled 6344 out of 9874, 64.25%

This is the first 2,000 images of our internal comma10k dataset. After we clean up these new labels, we'll release more. Learn more from [the Medium post](#), or on the [comma.ai discord](#) in the #comma-pencil channel.



It's 10,000 pngs of real driving captured from the comma fleet. It's MIT license, no academic only restrictions or

<https://github.com/commaai/comma10k>

# Waymo's open data set

A promotional graphic for the Waymo Open Dataset. It features a dark background with a circular inset showing a street scene with a white car and buildings. The scene is overlaid with a grid of blue dots. The text 'WAYMO Open Dataset' is prominently displayed in white. Below it, there are two paragraphs of text and a call to action with a right-pointing arrow icon.

**WAYMO**  
**Open**  
**Dataset**

The field of machine learning is changing rapidly. Waymo is in a unique position to contribute to the research community with some of the largest and most diverse autonomous driving datasets ever released.

Check out the newly released motion dataset in our Waymo Open Dataset and [2021 Challenges!](#)

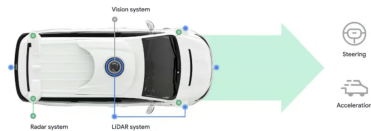
**Access Waymo Open Dataset** 

574 hours of data

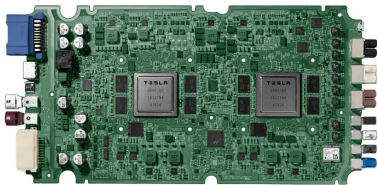
<https://github.com/waymo-research/waymo-open-dataset>



# AI issues in self driving

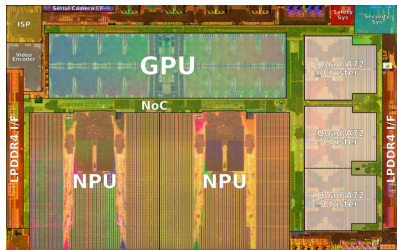


- modular end-to-end differential programming
- multi task, multi objective
- architecture design issues
- scene understanding: the never ending learning (long tails events)
- under budget



Tesla Full self-driving computer

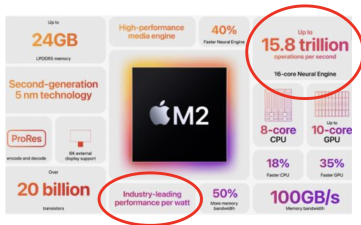
144 TOPS / 2300 Frames per second



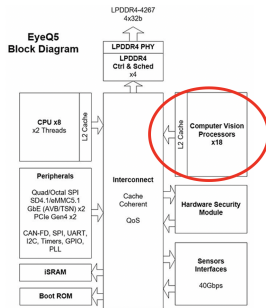
Tesla Full Self-Driving Chip

# AI hardware

Apple  
2022



MobilEye  
2018



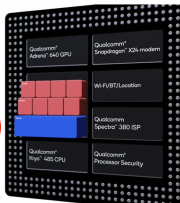
Qualcom  
2019

Qualcomm Snapdragon 855 mobile platform

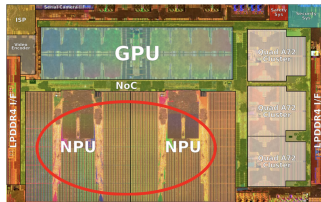
Adreno 640  
50% More ALUs\* FP32 & FP16

Hexagon 690  
New Tensor Accelerator  
- DTI designed  
- Deployed by AI  
- Multidimensional math and integrated nonlinear functions  
4x Vector eXtensions\*

Kryo 485  
New dot product instructions FP32 & INT8



Tesla  
2019



## Automotive tracks – Audi A8 Level 3: Aptiv zFAS controller

(Source: [www.reverse-costing.com](http://www.reverse-costing.com), System Plus Consulting)

### ➤ NVIDIA Tegra K1

Traffic sign recognition  
Pedestrian detection  
collision avoidance warning  
Light detection  
Lane recognition

### ➤ MobilEye EyeQ3

Traffic sign recognition  
360° camera images & processing

Functions: Courtesy of Aptiv



### ➤ Altera Cyclone

Object fusion  
Map fusion  
Parking pilot  
Pre-crash  
Sensor data pre-processing

### ➤ Infineon Aurix Tricore

Traffic jam pilot  
Assistance systems  
Matrix beam  
Road graph

## Comparizon



Company	DL framework	sensors	hardware (chip)
Openpilot	Meta Pytorch?	cameras + radar	Qualcomm (M1?)
Tesla	Meta Pytorch	8 cameras	Tesla's FSD chip
Mobil eye	Tensorflow on AWS	11 cameras (vidar)	ST microelectronic
Waymo	Google Tensorflow	cameras + Lidars + radars	Intel -> Samsung ?
Cruise	Microsoft Azure	4 cameras + Lidar + radar + audio	origin cruise chip

# Road map

- 1 Today's vehicles with AI
- 2 Embedded AI
- 3 Data to train the AI
- 4 The future of Artificial Intelligence for driving



# Tesla vs. Waymo (vs. Openpilot, MobilEye, Appolo...)

	Tesla	Waymo
		
Autonomy objective	Driver assist Hands on the wheel	Self driving no driver
Sensors	Camera Autopilot on highway	Lidar + Camera Taxi in a known area
Data from	the Fleet 3 millions miles per day	Simulation 16 billions simulated miles
AI	multi task deep learning semi supervised	hybrid deep learning reinforcement & Auto ML
Running by now	2 000 000	600
Project started in	2014/2016	2009

# Towards scaling self driving

When will we have more than 10,000 Full Autonomous cars?

- Tesla's strategy of the little steps (improving the ADAS)
- Wyamo strategy including more areas (less specific)
  
- not yet: status quo
  - ▶ driving assistance (automation)
    - ★ increase safety
    - ★ reduces environmental impact
  - ▶ specific applications
  - ▶ communication and equipment
  
- No full autonomy unless... safety is proven
  - ▶ new solution (cf Google)

## Accidents: 14 lethal since 2015 (+1 processing)



[https://en.wikipedia.org/wiki/List\\_of\\_self-driving\\_car\\_fatalities](https://en.wikipedia.org/wiki/List_of_self-driving_car_fatalities)



# Safety Ratings

Safety Assist evaluating driver-assistance and crash-avoidance technologies.

2019 - Notation → A PROPOS DE LA NOTATION EN 2019

Marque et modèle	Équipement de sécurité	Notation globale				
Tesla Model 3	De série	★★★★★	96%	86%	74%	94%
Tesla Model X	De série	★★★★★	98%	81%	72%	94%
Citroën C5 Aircross	Pack sécurité	★★★★★	89%	86%	67%	82%
Volkswagen T-Cross	De série	★★★★★	97%	86%	81%	80%
Audi A1	De série	★★★★★	95%	85%	73%	80%
SEAT Tarraco	De série	★★★★★	97%	84%	79%	79%
Škoda Octavia	De série	★★★★★	92%	88%	73%	79%
Mercedes-Benz GLE	De série	★★★★★	91%	90%	78%	78%
Subaru Forester	De série	★★★★★	97%	91%	80%	78%
VW Golf	De série	★★★★★	95%	89%	76%	78%
Lexus UX	De série	★★★★★	96%	85%	82%	77%

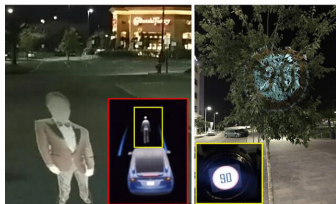
# Attacks against autonomous vehicles



Eykholt et al, Robust Physical-World Attacks on Deep Learning Visual Classification, CVPR 2018

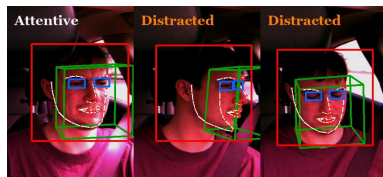


Zhang et al., CAMOU: Learning Physical Vehicle Camouflages to Adversarially Attack Detectors in the Wild, ICLR 2019



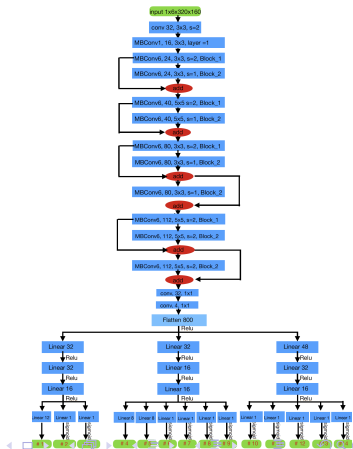
<https://www.mcafee.com/blogs/other-blogs/mcafee-labs/model-hacking-ad-as-to-pave-safer-roads-for-autonomous-vehicles/>  
Nassi et al., Phantom of the ADAS: Securing Advanced Driver-Assistance Systems from Split-Second Phantom Attacks, 2020  
Qayyum, et al., Securing Connected & Autonomous Vehicles: Challenges Posed by Adversarial ML, IEEE Communications, 2019

# Attacking Openpilot 's DMS

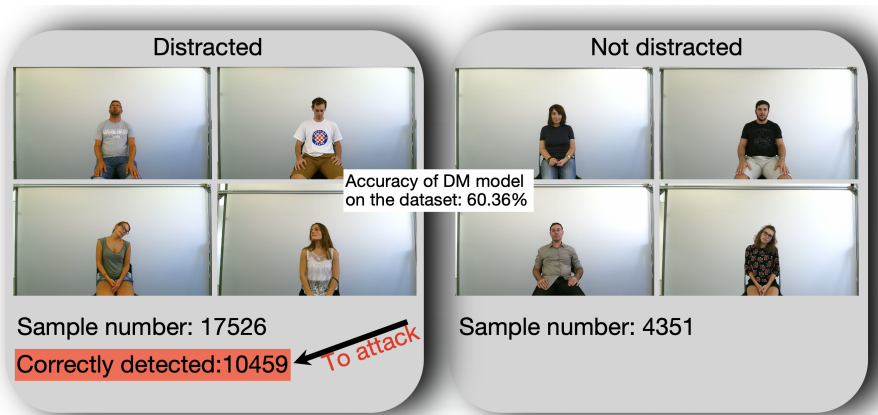


## Three components

- Face localization
  - ▶ openCV -> cropping
- Feature generation
  - ▶ EfficientNet b0 architecture
  - ▶ Fine tuning
- Decision module
  - ▶ Treshold based decision



# Datasets: Pandora (head pose)



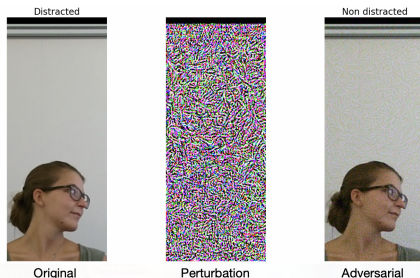
1. Borghi, Guido, et al. "Poseidon: Face-from-depth for driver pose estimation." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.

# Attack performance

- Accuracy on original data: 100%

- Attack settings:

- torchattacks
- $c=1000$  for CW
- $steps = 50$  for CW and Deepfool
- $L_{\infty} 10/255 =$  for all the others



- Accuracy on adversarial data:

Attack models	FGSM	CW	PGD	APGD	AutoAttack	Deepfool
Accuracy(%)	81.85	21.90	13.17	0.057	<b>0.0</b>	6.39

100 % Distracted



100 % Attentive

# Future of AI mobility

- Car manufacturer

IT companies  
AI, hardware & big data

- Prove safety

deep learning theory

- New AI algorithms

common sense (cf Y. LeCun)  
unsupervised learning

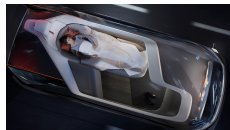
- Co-driving

HM interactions

Communications

- Acceptability

Ethic



Questions?

<https://chaire-raimo.github.io/>