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

Stéphane Canu
https://chaire-raimo.github.io/
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
Autonomous vehicles: when?


Un jour, Elon Musk aura raison.

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

Autonomous Vehicles “Tipping Point”
Transition from human drivers to vehicles driving

Our World Today

<table>
<thead>
<tr>
<th>Level</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Human Only</td>
<td>The driver (human) controls everything: steering, brakes, throttle, power.</td>
</tr>
<tr>
<td>1</td>
<td>Modern Vehicle</td>
<td>Most functions are still controlled by a driver, but some (like braking) can be done automatically by the car.</td>
</tr>
<tr>
<td>2</td>
<td>Modern Plus</td>
<td>At least 2 functions are automated (like cruise control &amp; lane-centering), but the driver must be ready to take control of the vehicle.</td>
</tr>
<tr>
<td>3</td>
<td>Partial Autonomy</td>
<td>Drivers are still necessary, but are not required to monitor the situation as with previous levels.</td>
</tr>
<tr>
<td>4</td>
<td>Full Autonomy (+ Human)</td>
<td>Vehicles perform all safety-critical driving functions and monitor roadway conditions for an entire trip, with option for human driving.</td>
</tr>
<tr>
<td>5</td>
<td>Full Autonomy (No Human)</td>
<td>No option for human driving - no steering wheel or controls.</td>
</tr>
</tbody>
</table>

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

Cox Automotive® Brand

KBB.com

The Blue Book

Monitored Driving
- Driver is continuously exercising longitudinal AND lateral control
- Temporary Hands off

Non-Monitored Driving
- Driver is not required during defined use case
- System can cope with all situations automatically in a defined use case
- No driver required
Level 2/3 = ADAS Ratings

Consumer Reports’ for major Advanced Driver Assistance Systems (2020)

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

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)*

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

* 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 players:
- 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)

https://paperswithcode.com/sota/image-classification-on-imagenet
Detection, tracking and recognition of traffic signs (2011-13)

**Recognition**  German Traffic Sign Recognition Benchmark (GTSRB) data set, containing 51,839 labelled images of real-world traffic signs.

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

and the winner is

→ Deep learning gives very good results on both tasks
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

openpilot is the Android

https://github.com/commaai/openpilot

Works with the push of a button.

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

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
Openpilot’s 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

Andrej Karpathy, Multi-Task Learning in the Wilderness, ICML 2019
https://slideslive.com/38917690/multitask-learning-in-the-wilderness
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 × 460 3D map of the surroundings
Tesla perception module

1. feature generator: backbone
   (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

Mayank Bansal, ChauffeurNet: Learning to Drive by Imitating the Best and Synthesizing the Worst, ICML 2019
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.

Drago Anguelov (Waymo) - MIT Self-Driving Cars lectures
https://medium.com/waymo/automl-automating-the-design-of-machine-learning-models-for-autonomous-driving-141a5583ec2a
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

Karpathy (Tesla) ICML 2019
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
Openpilot : l’étiquetage des données par crowd sourcing

**comma10k**

Count and Percentage of Available Images Labeled: 6,344 out of 9,874, 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

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
AI hardware

Apple
2022

Qualcomm
2019

MobilEye
2018

Tesla
2019
The Audi A8 hardware

Automotive tracks – Audi A8
Level 3: Aptiv zFAS controller
(Source: 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

- 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

Functions: Courtesy of Aptiv
<table>
<thead>
<tr>
<th>Company</th>
<th>DL framework</th>
<th>sensors</th>
<th>hardware (chip)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Openpilot</td>
<td>Meta Pytorch?</td>
<td>cameras + radar</td>
<td>Qualcomm (M1?)</td>
</tr>
<tr>
<td>Tesla</td>
<td>Meta Pytorch</td>
<td>8 cameras</td>
<td>Tesla’s FSD chip</td>
</tr>
<tr>
<td>Mobil eye</td>
<td>Tensorflow on AWS</td>
<td>11 cameras (vidar)</td>
<td>ST microelectronic</td>
</tr>
<tr>
<td>Waymo</td>
<td>Google Tensorflow</td>
<td>cameras + Lidars + radars</td>
<td>Intel -&gt; Samsung ?</td>
</tr>
<tr>
<td>Cruise</td>
<td>Microsoft Azure</td>
<td>4 cameras + Lidar + radar + audio</td>
<td>origin cruise chip</td>
</tr>
</tbody>
</table>
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...)

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

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

<table>
<thead>
<tr>
<th>Marque et modèle</th>
<th>Équipement de sécurité</th>
<th>Notation globale</th>
<th>2019 - Notation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tesla Model 3</td>
<td>De série</td>
<td><strong>⭐⭐⭐⭐</strong></td>
<td>96%</td>
</tr>
<tr>
<td>Tesla Model X</td>
<td>De série</td>
<td><strong>⭐⭐⭐⭐</strong></td>
<td>98%</td>
</tr>
<tr>
<td>Citroën C5 Aircross</td>
<td>Pack sécurité</td>
<td><strong>⭐⭐⭐⭐</strong></td>
<td>89%</td>
</tr>
<tr>
<td>Volkswagen T-Cross</td>
<td>De série</td>
<td><strong>⭐⭐⭐⭐</strong></td>
<td>97%</td>
</tr>
<tr>
<td>Audi A1</td>
<td>De série</td>
<td><strong>⭐⭐⭐⭐</strong></td>
<td>95%</td>
</tr>
<tr>
<td>SEAT Tarraco</td>
<td>De série</td>
<td><strong>⭐⭐⭐⭐</strong></td>
<td>97%</td>
</tr>
<tr>
<td>Škoda Octavia</td>
<td>De série</td>
<td><strong>⭐⭐⭐⭐</strong></td>
<td>92%</td>
</tr>
<tr>
<td>Mercedes-Benz GLE</td>
<td>De série</td>
<td><strong>⭐⭐⭐⭐</strong></td>
<td>91%</td>
</tr>
<tr>
<td>Subaru Forester</td>
<td>De série</td>
<td><strong>⭐⭐⭐⭐</strong></td>
<td>97%</td>
</tr>
<tr>
<td>VW Golf</td>
<td>De série</td>
<td><strong>⭐⭐⭐⭐</strong></td>
<td>95%</td>
</tr>
<tr>
<td>Lexus UX</td>
<td>De série</td>
<td><strong>⭐⭐⭐⭐</strong></td>
<td>96%</td>
</tr>
</tbody>
</table>

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


Attacking Openpilot’s DMS

Three components

- Face localization
  - openCV -> cropping

- Feature generation
  - EfficientNet b0 architecture
  - Fine tuning

- Decision module
  - Threshold based decision
Datasets: Pandora (head pose)

Accuracy of DM model on the dataset: 60.36%

Sample number: 17526
Correctly detected: 10459

Sample number: 4351

Attack performance

- Accuracy on original data: 100%

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

- Accuracy on adversarial data:

<table>
<thead>
<tr>
<th>Attack models</th>
<th>FGSM</th>
<th>CW</th>
<th>PGD</th>
<th>APGD</th>
<th>AutoAttack</th>
<th>Deepfool</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy(%)</td>
<td>81.85</td>
<td>21.90</td>
<td>13.17</td>
<td>0.057</td>
<td>0.0</td>
<td>6.39</td>
</tr>
</tbody>
</table>

100% Distracted

100% Attentive
Future of AI mobility

- Car manufacturer
- IT companies

- Prove safety
- AI, hardware & big data

- New AI algorithms
- deep learning theory

- Co-driving
- common sense (cf Y. LeCun)

- HM interactions
- unsupervised learning

- Acceptability
- Communications

- Ethic
Questions?

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