Adversarial Machine Learning and Crashing Autonomous Vehicles

Mapping the Legal Problems

Steven Van Uytsel – Faculty of Law

Multidisciplinary Perspectives on Algorithms

November 21-23, 2019

KYUSHU UNIVERSITY

https://www.nbcnews.com/tech/innovation/self-driving-uber-car-involved-fatal-accident-arizona-n857941

Elaine Hertzberg

https://www.bbc.com/news/magazine-10987606

Bridget Driscoll

Operator Liable

Human defined as operator

the one engaging the vehicle

the one sitting at the driver seat

the one that pushes the start button

the owner of the vehicle

Human Takes Control

Each vehicle needs to have technology allowing for someone to take control

need to have a driver to take control

vehicle should have a break, accelerator, and steering wheel

override switch

https://www.reifflawfirm.com/whats-worse-sleeping-drivers-self-driving-cars/

Japanese Law on Compensation

Japanese Act on Securing Compensation for Automobile Accidents

"a person that puts an automobile into operational use for that person's own benefit is liable to compensate for damage arising from the operation of the automobile if this results in the death or bodily injury of another person"

Supreme Court on 'put into operation' – person who has control over and benefit from the vehicle

- owner driving herself
- owner employing a driver for herself
- owner employing a driver for business
- owner who allows someone to drive his vehicle (even a system)

(1964 Ferrari GT/L)

What the Future Will Bring

https://www.viatech.com/en/2019/08/a-new-passenger-experience-in-autonomous-vehicles/

Technology in Future Cars

Sensors

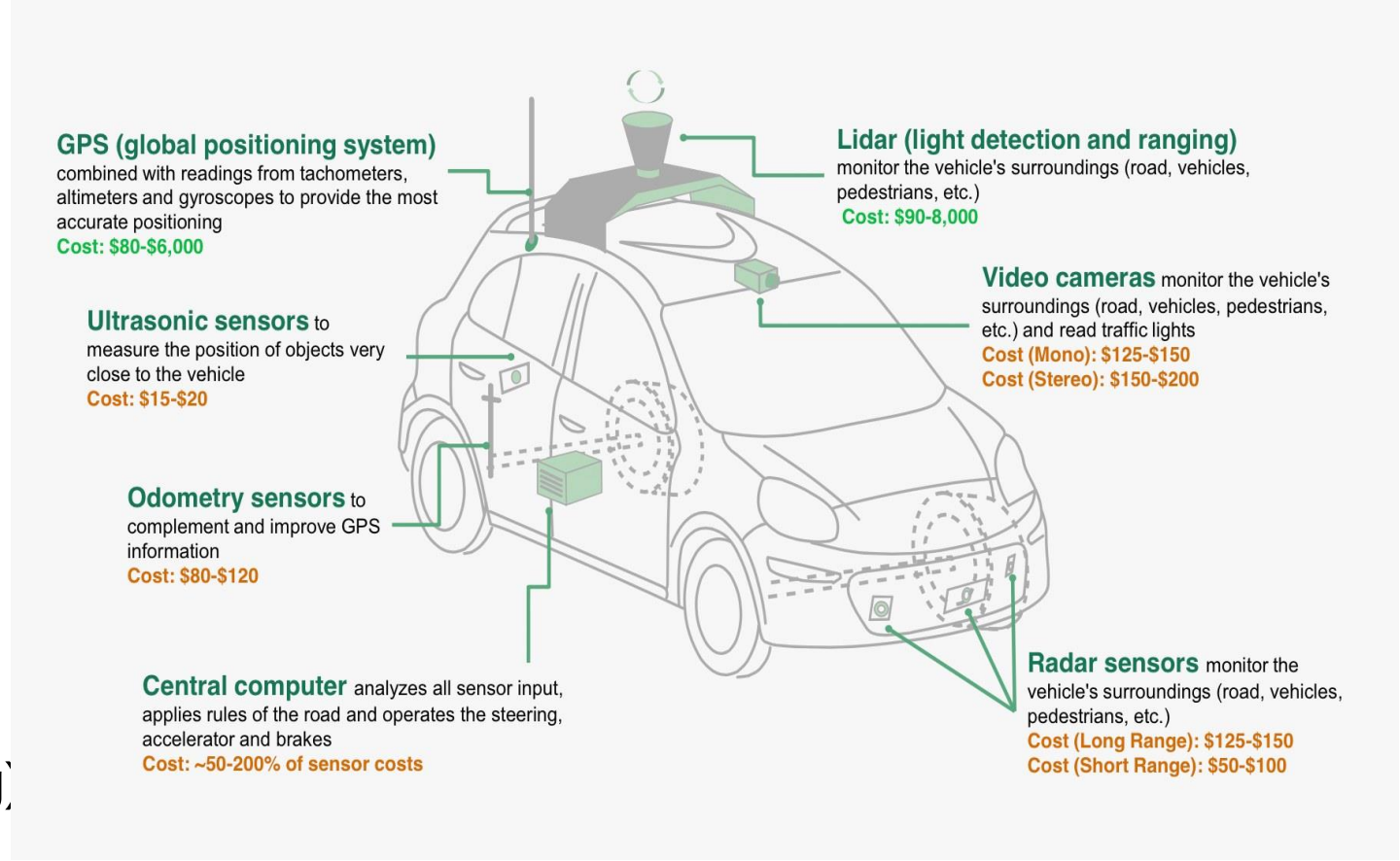
- Radar
- Lidar
- Video Camera
- Sonar

Mapping/Location

- GPS
- Digital Maps / Lidar

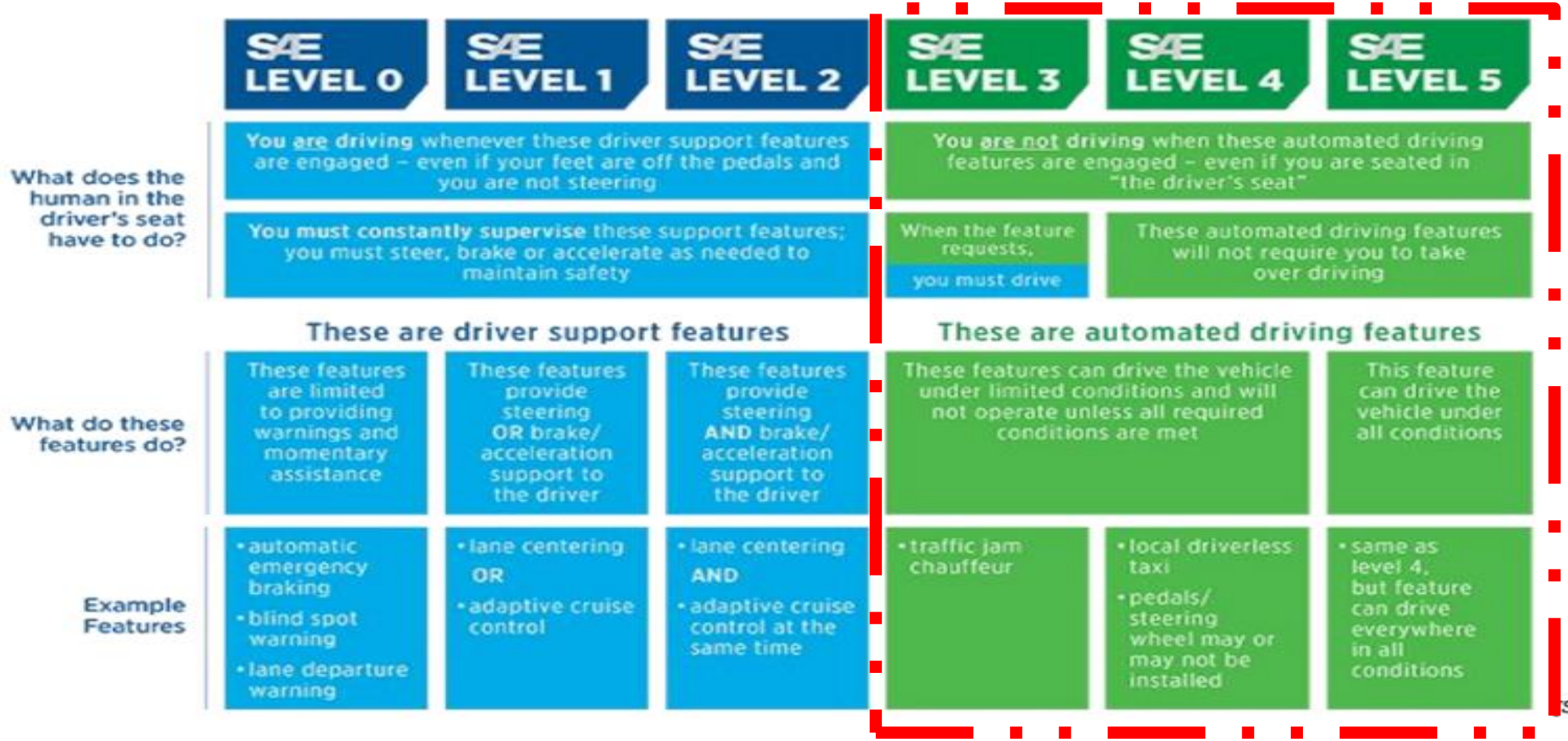
Software

- Object Recognition and Classification (machine learning)
- Planning Software
- Acting Software



https://www.wired.com/2015/04/cost-of-sensors-autonomous-cars/

SAE J3016™ LEVELS OF DRIVING AUTOMATION



[Source: SAE]

Humans and Accidents with Autonomous Vehicles

Accidents may occur – compensation is required (limit to civil liability – payment of compensation)

* human negligence
  - other participants in traffic (sudden movements)
  - vehicle outside the Operational Design Domain (driving in bad weather, no cleaning sensors, …)
  - failure to update software

* mechanical breakdown of the vehicle

But what with the failure of the artificial intelligence (AI)

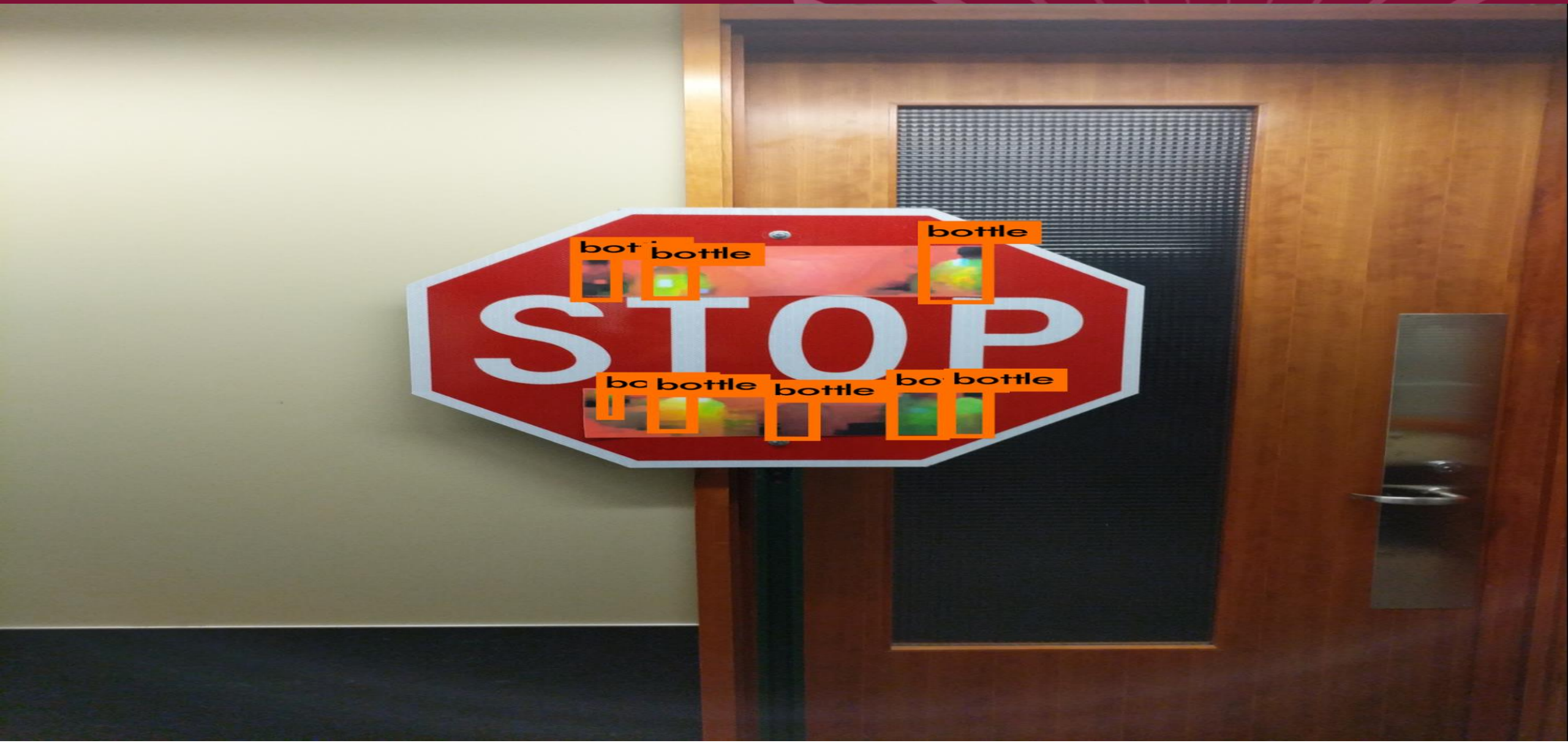
Adversarial Machine Learning

Adversarial examples are inputs to **machine learning** models that an attacker has intentionally designed to cause the model to make a mistake

Su, J., Vargas, D. V., & Kouichi, S. (2017). One pixel attack for fooling deep neural networks. arXiv preprint arXiv:1710.08864.

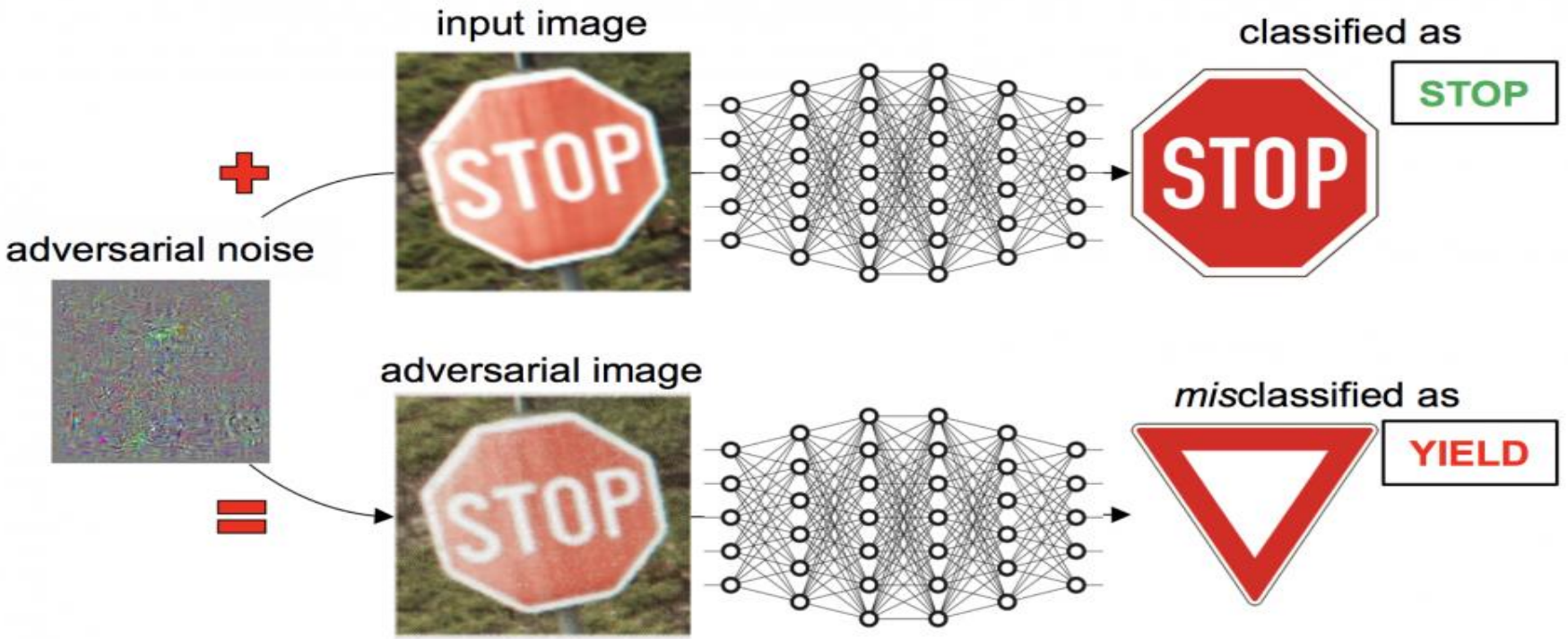

INTERDIT SAUF
ACCÈS PARC DE
STATIONNEMENT
American Apparel


STOP
COAST
GOOGLE Georgia Guidestones
FIGHT THE NEW WORLD ORDER

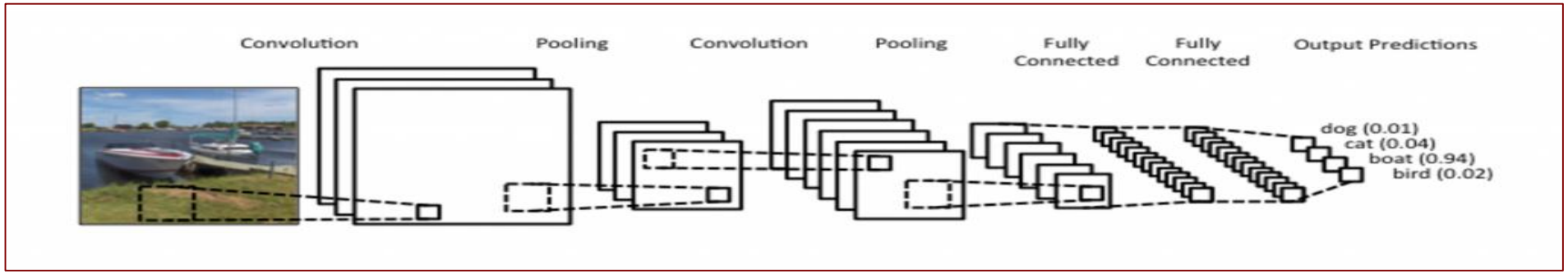


http://bair.berkeley.edu/blog/2017/12/30/yolo-attack/


speedlimit 0.947
STOP



https://blog.csiro.au/vaccinating-machine-learning-against-attacks/



http://www.wildml.com/2015/11/understanding-convolutional-neural-networks-for-nlp/

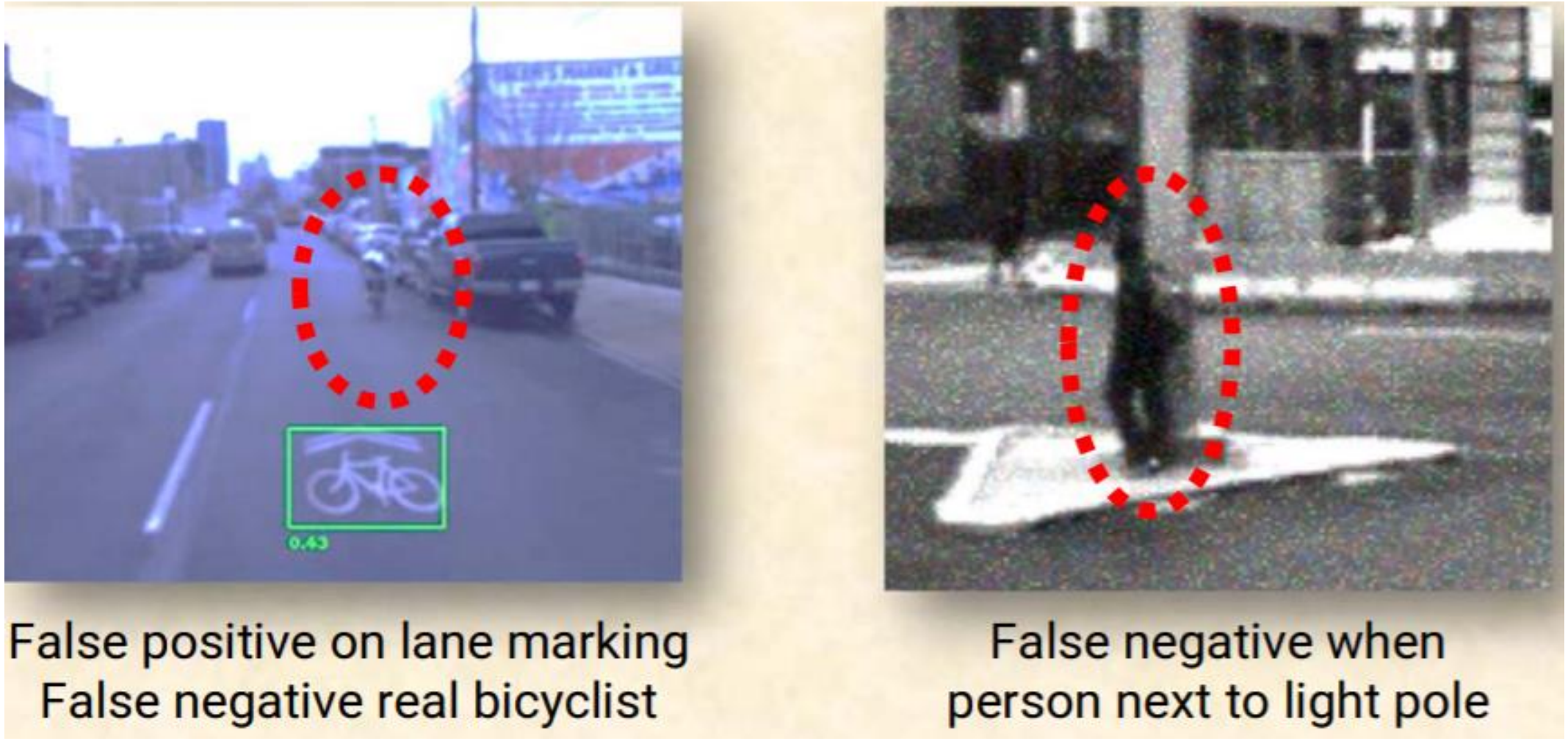
Fig. 2. When the vehicle hits a water puddle, the images captured by the camera will be tilted. As a result, the RGB-based object detector may misclassify objects in scenes.

Source: https://arxiv.org/pdf/1907.05274.pdf: By Liu Letao, Martin Saerbeck, and Justin Dauwels

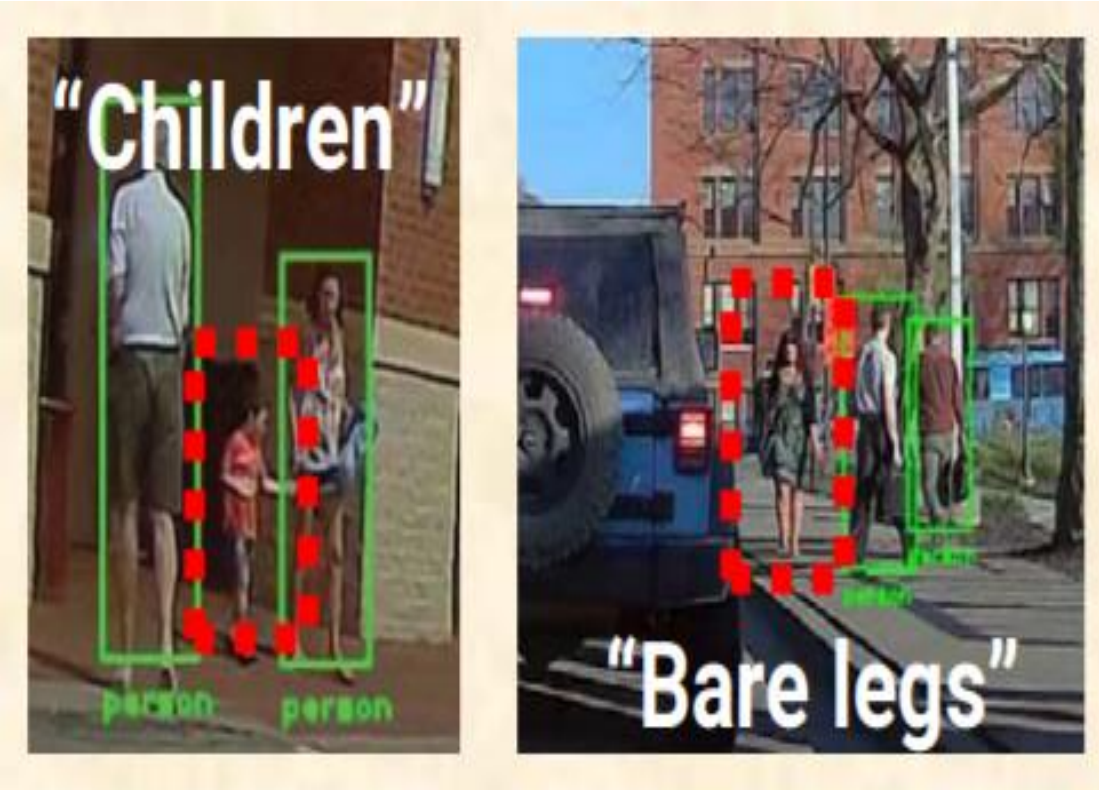
Moving Away from Adversarial Machine Learning towards the Edge Cases

Cases You don’t Think Of

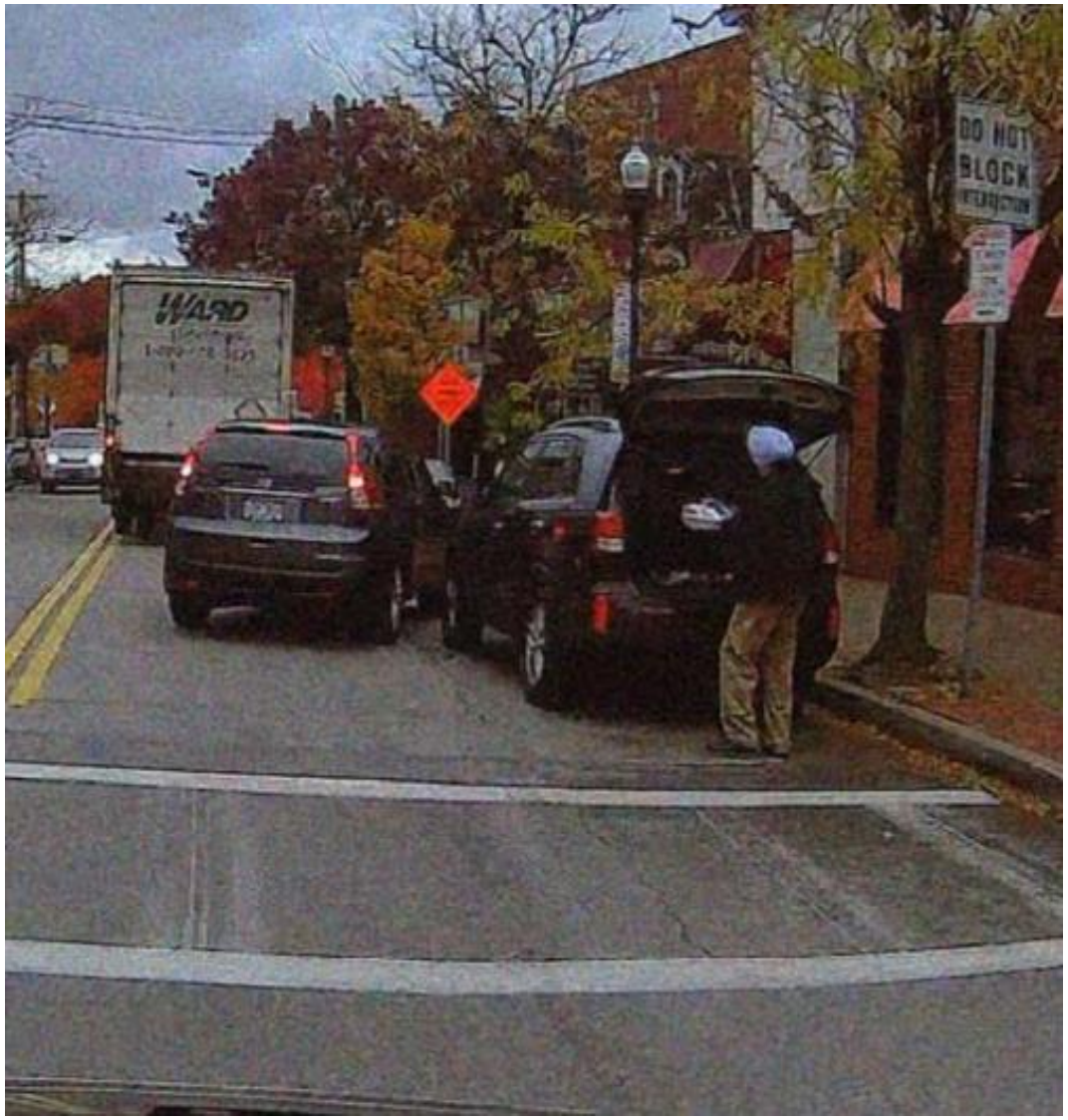
Machine Learning Can have Edges a Human Does Not Expect



http://users.ece.cmu.edu/~koopman/lectures/Koopman19_SSS_slides.pdf



http://users.ece.cmu.edu/~koopman/lectures/Koopman19_SSS_slides.pdf



http://users.ece.cmu.edu/~koopman/lectures/Koopman19_SSS_slides.pdf

Response of Engineers

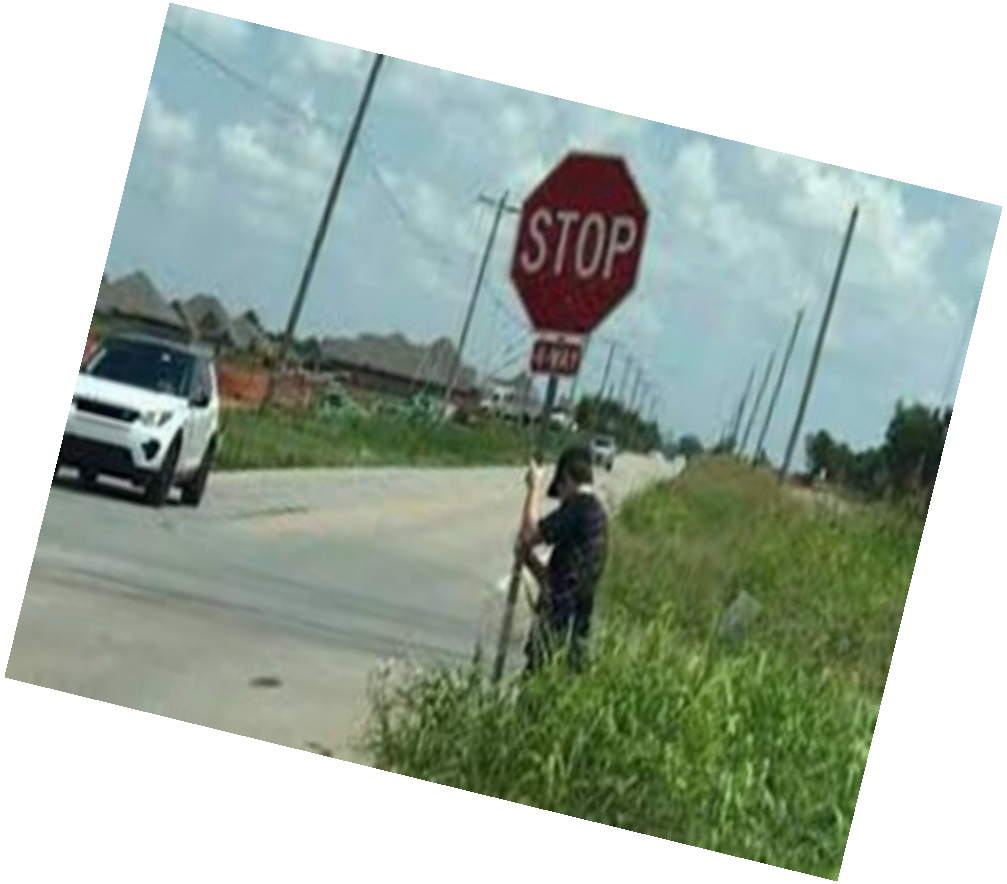
Clear definition of the **operational design domain** + progressive testing

Safety by design:

Advocating for connectivity

Relying on other instruments than only camera


60
40


STOP


100
TEMPORARY


STOP

A Human Operator or the Artificial Intelligence?

Is it fair to look at the ‘human’ operator to request compensation?

SAE level 3 vehicle – human drive when request to regain control

SAE level 4 or above vehicle - human has no role to play

If No, What are the Alternatives?

Product Liability Law

questions: 1) is AI a product? 2) When is AI defective?

problems: lengthy trials?

Negligence

questions: 1) what is the standard of care?

problem: lengthy trial?

No-fault Liability

questions re: organization: 1) insurer, 2) insurance taker

problem: willingness of the insurance sector to change

Additional Reasons for Reconsidering Liability Rules

Automated driving and vehicle ownership

*shift from individual ownership to fleet ownership

*fleet owner as operator – liable for compensation (maybe fair)

=> Market access prohibitively high

Infrastructure enabled automotive driving

increase of actors, products and services

=> complex web of potential defendants for a victim of a traffic accident to sue (negligence or defect)

In conclusion

Need to rethink the liability rules for autonomous vehicles

*think of a victim that has to receive a smooth compensation

*think of deter OEMs from releasing faulty products

This may be even more necessary

1) if we move to changed ownership structures?

2) if we move to infrastructure enabled autonomy?

Factors Complicating Product Liability Law: Product

Is AI a product?

unsettled issue – no case law or legislation

tendency in the literature to treat AI as a product

Fumio Shimpo, Member Cabinet Office Advisory Board on AI

> "for an example of the current legal dilemma, I will refer the reader to an accident involving a robot which was caused by inaccurate information or software defect malfunction. At present, the questioning of the product liability of the information itself, which was the main cause of the accident, is outside the range of the current Japanese Product Liability Act."

=> Revert back to negligence: victim has to pinpoint the exact fault in the system

Conceptualization of a No-Fault Scheme

Different formats

insurance taker – possessor of vehicle, manufacturer, both, or any other person that may be justifiable in the future, etc.

insurer – state, third-party private insurance company, car manufacturer, etc.

Adjust the format according to the aims

compensation – easy compensation to victims

prevention of accidents – make manufacturers contribute to the fund

tort law remains option if not choose for the fund

Moving in the Direction of No-Fault Compensation

Volvo CEO

- publicly indicated that they want to take responsibility for accidents with their self-driving vehicles

Internalization of the costs of the accidents

- manufacturers prevent negative publicity
- manufacturers prevent unpredictable court decisions
  - (partly selfish – if a victim is not able to get compensation – trust in autonomous vehicles will disappear)

The concept of “driver” under the Road Traffic Act now includes an “Autonomous Driving Device” – a system that takes the place of recognition, prediction, judgment and operation of humans entirely.

Vehicles must be fitted with a data recorder (an Operating Condition Recorder), which may be required by the police in the event of an accident.

Drivers may only use Level 3 features when their vehicles satisfy the operational design domain set by the manufacturer.

Drivers are not permitted to sleep, consume alcohol or to sit in the back seat, but may use a mobile phone, read a book, eat or watch an in-built entertainment system, provided that they are ready to immediately take back control of the vehicle.

Approvals will be required through a permit system for wireless updates to self-driving programs, aimed at preventing the risk of cybersecurity breaches.

Manufacturers will be required to provide technical information to facilitate inspections of autonomous driving equipment, and special certification will be required for autonomous vehicle maintenance providers to carry out operations.