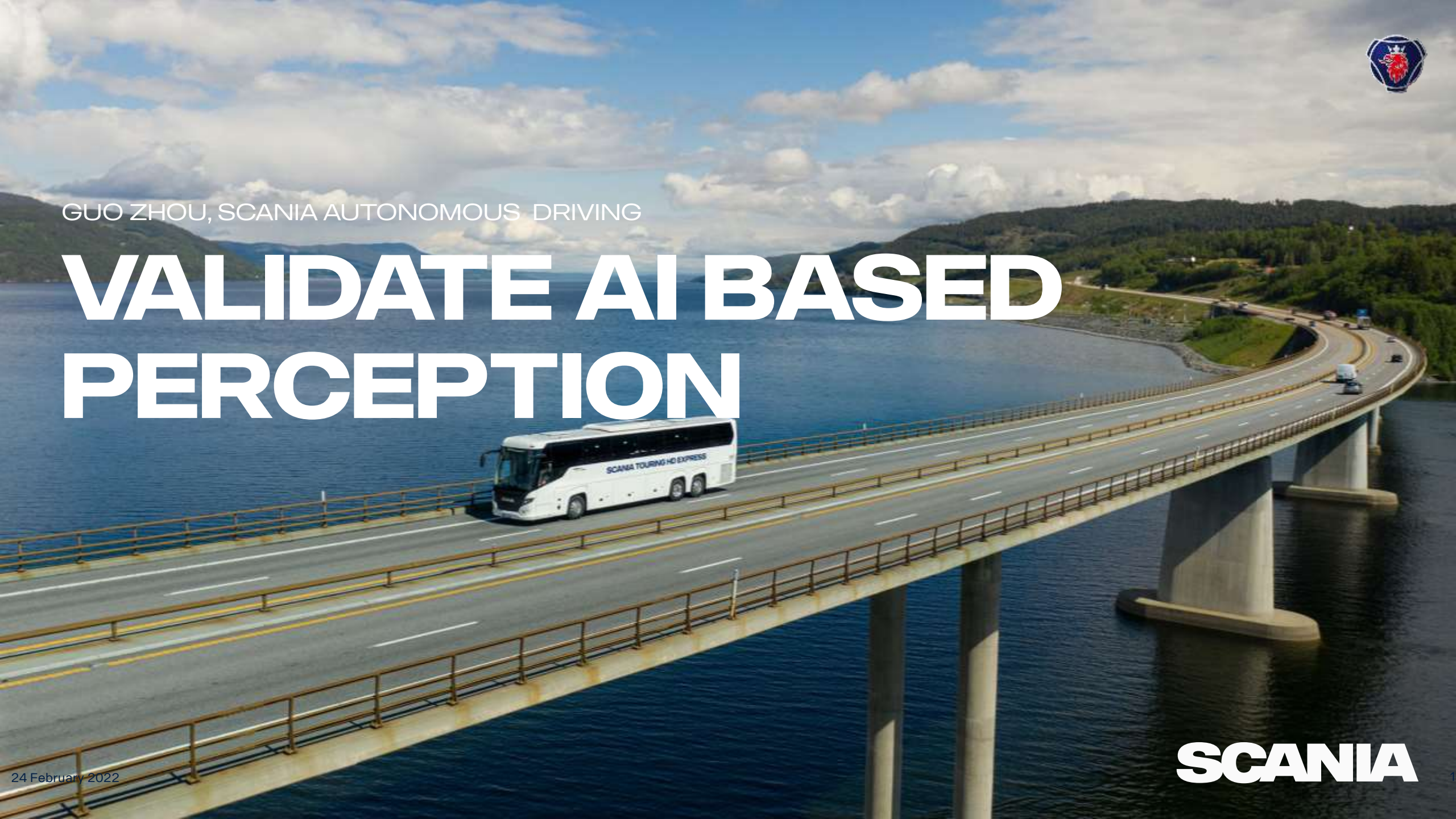


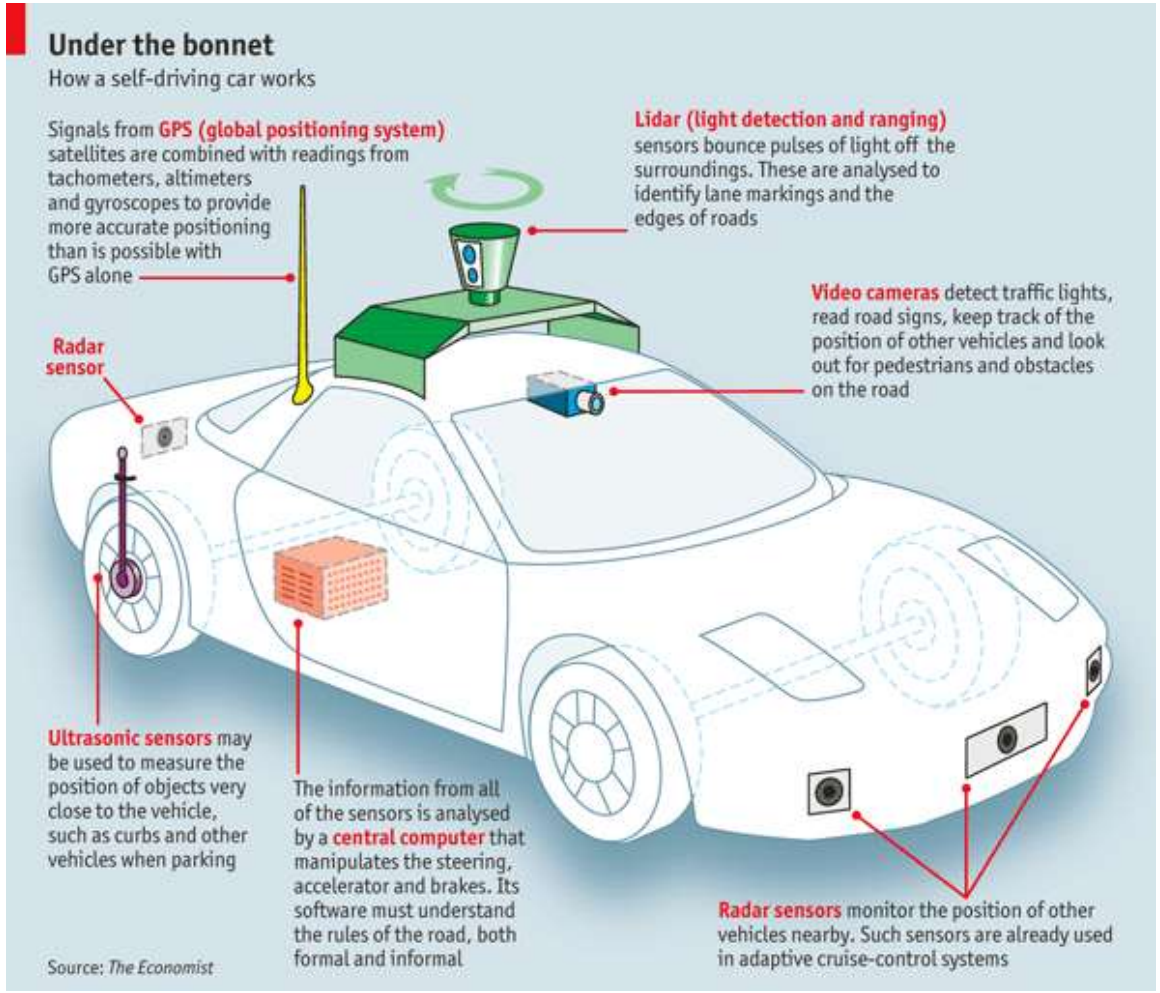
GUO ZHOU, SCANIA AUTONOMOUS DRIVING

VALIDATE AI BASED PERCEPTION

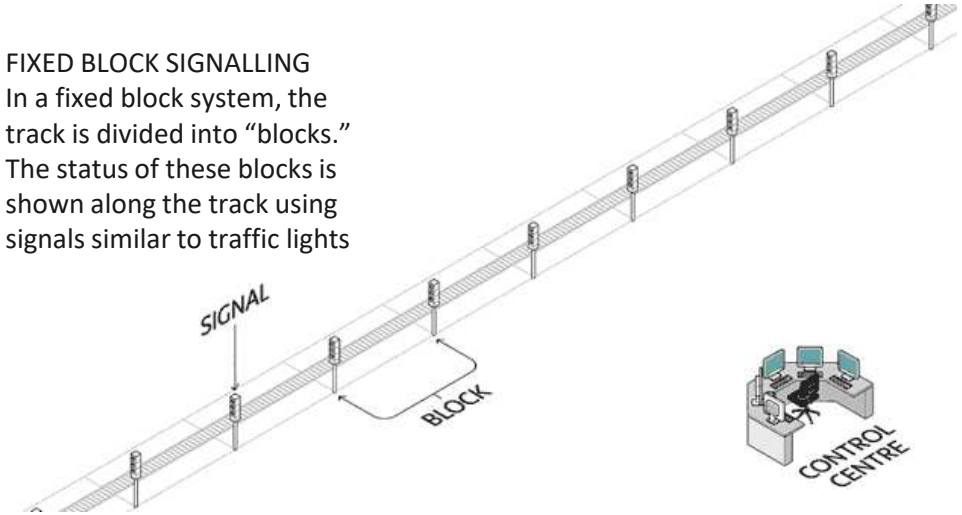
CONTENT

SCANIA

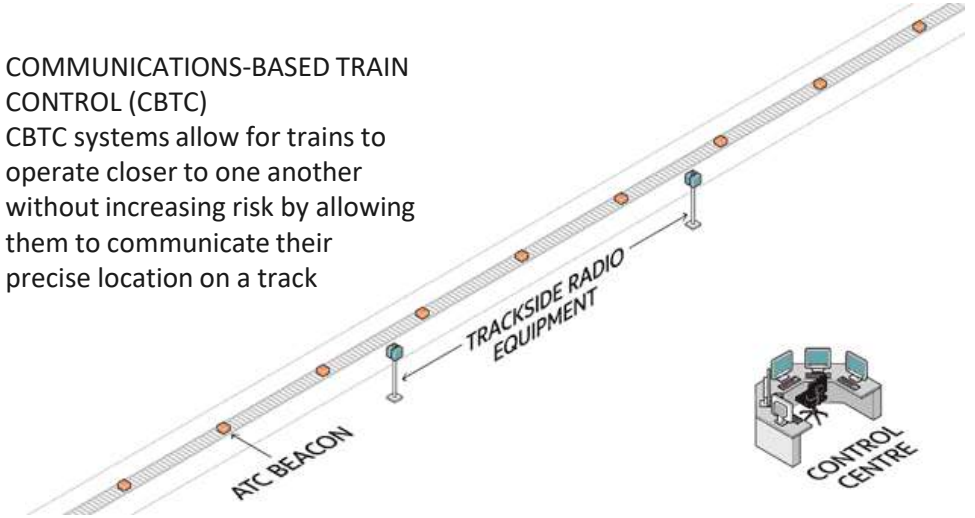
Autonomous Driving vs. Railway Signaling



FIXED BLOCK SIGNALLING
In a fixed block system, the track is divided into “blocks.” The status of these blocks is shown along the track using signals similar to traffic lights

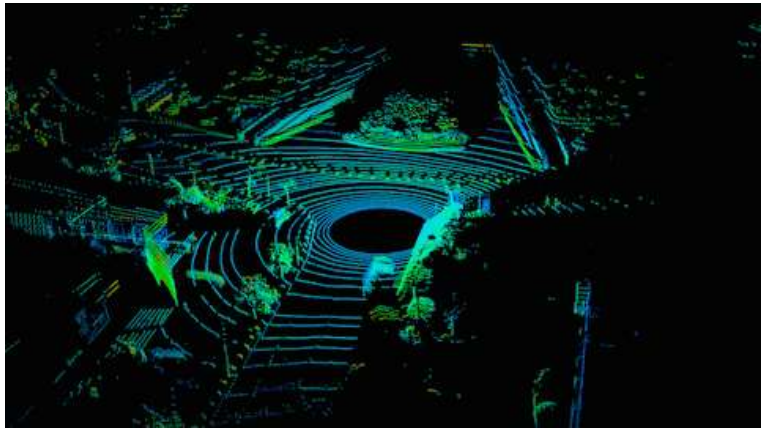


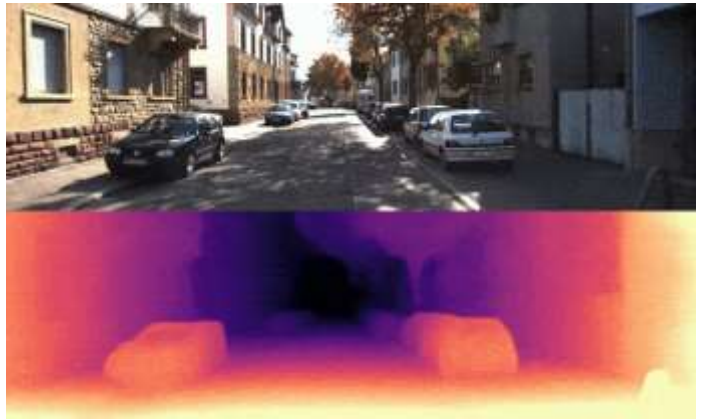
COMMUNICATIONS-BASED TRAIN CONTROL (CBTC)
CBTC systems allow for trains to operate closer to one another without increasing risk by allowing them to communicate their precise location on a track



https://www.theglobeandmail.com/news/toronto/ttc-subway-signalling/article35174381/

AI Based Perception





Challenges in Validating AI Based Perception Performance Limitation

The model cannot perform well in the task in conditions that are considered as normal for humans

- The pedestrian was alternatively detected as a vehicle (and then considered as traveling in the other lane), and as an unknown (static) object.
- 2.5 seconds before impact, it was seen as a bicycle, and 1.2 seconds before as being on the path of the car.
- An alarm was raised, and 0.02s before impact the operator took control of the wheel.
- One observation made is that *the system design did not include a consideration for jaywalking pedestrians*.

Uber accident in Arizona 2018

https://www.youtube.com/watch?v=XtTB8hTgHbM

https://cctedu.net/lustersoft_file/ueditor/file/20191111/1573436942679082425.pdf

Challenges in Validating AI Based Perception Performance Limitation

- The vehicle plows directly into the top of a large white truck lying on its side.
- The driver states the vehicle was in Autopilot mode.
- The driver did not hit the brakes himself until far too late, indicating he was probably not paying attention.
- The road has light traffic and visibility is very good. Nobody was injured.

Tesla accident in Taiwan 2020

https://www.youtube.com/watch?v=LfmAG4dk-rU

https://www.forbes.com/sites/bradtempleton/2020/06/02/tesla-in-taiwan-crashes-directly-into-overturned-truck-ignores-pedestrian-with-autopilot-on/?sh=4bf6cbe258e5

Challenges in Validating AI Based Perception Robustness Limitation

- The model performs well but has vulnerabilities that may lead to malfunctions in specific conditions (chewing gum or bird's dropping, etc.)

- These malfunctions may appear either naturally during the execution of the program or be intentionally provoked by an adversary with malicious intentions.

Adversarial examples for a neural network trained on the GTSRB dataset.

After a slight perturbation of Euclidean distance 0.88, the image classification changes from “go right or straight” to “go left or straight”.

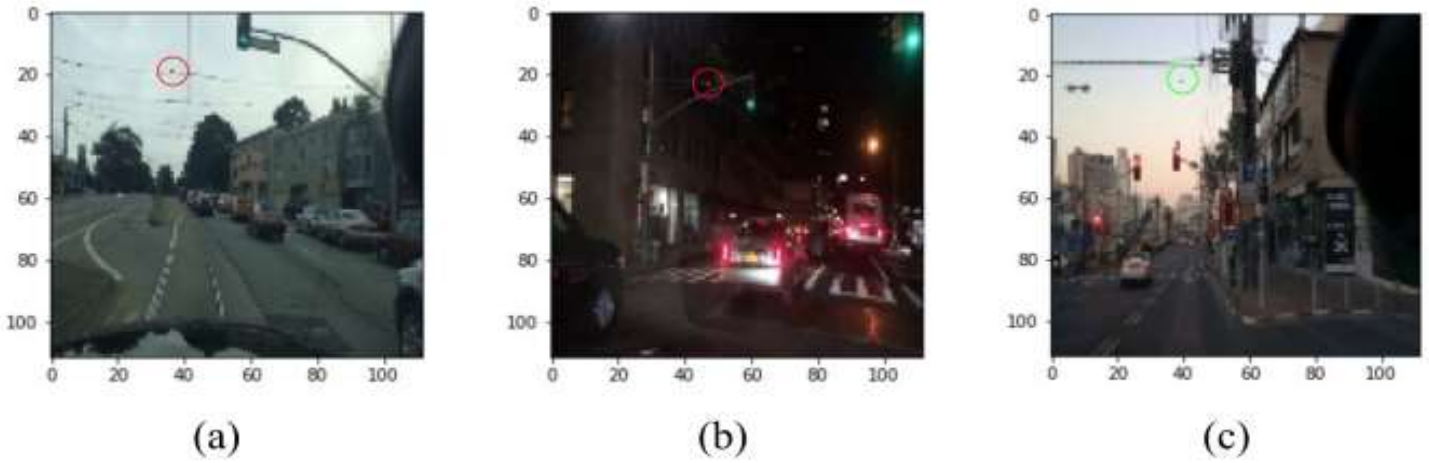
Speed limit 80” misclassified into “speed limit 60.

“Danger” misclassified into “pedestrian crossing”.

https://arxiv.org/pdf/1807.03571.pdf

Challenges in Validating AI Based Perception Robustness Limitation

Adversarial examples generated on Nexar data demonstrate a lack of robustness.



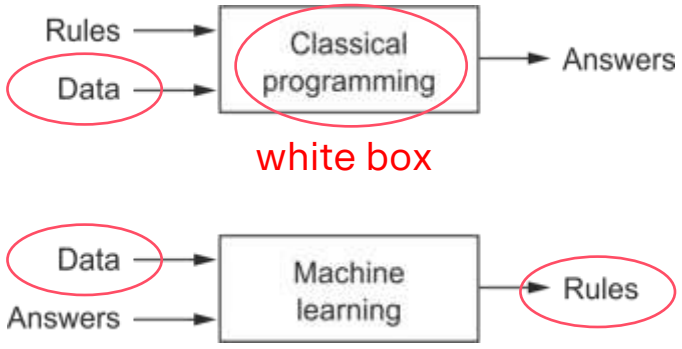
(a) Green light classified as red with confidence 56% after one pixel change.

(b) Green light classified as red with confidence 76% after one pixel change.

(c) Red light classified as green with 90% confidence after one pixel change.

Method Overview
Validation Targets

Comparison



ML Validation

Data validation

1.preprocess various data sources; sensor fusion if necessary

2.transform data to standardized format

3.validation: ground truth comparison, data feature skew/drift check

Model Validation

1.performance

2.robustness

3.edge cases and corner cases

4.data independent with training

5.formal verification

Challenges in Data Validation

Data validation/safety is neither a new problem nor unique to ML, but rethink in ML and autonomous system context

- What is the definition of data quality? and how to measure it?

feature engineering, dataset dissimilarity, class distribution, etc. No benchmark, no standard.

- What are the requirements for data validation?

relevant, complete, balanced, accurate, consistent, etc. Facilitate in CI/CD pipeline

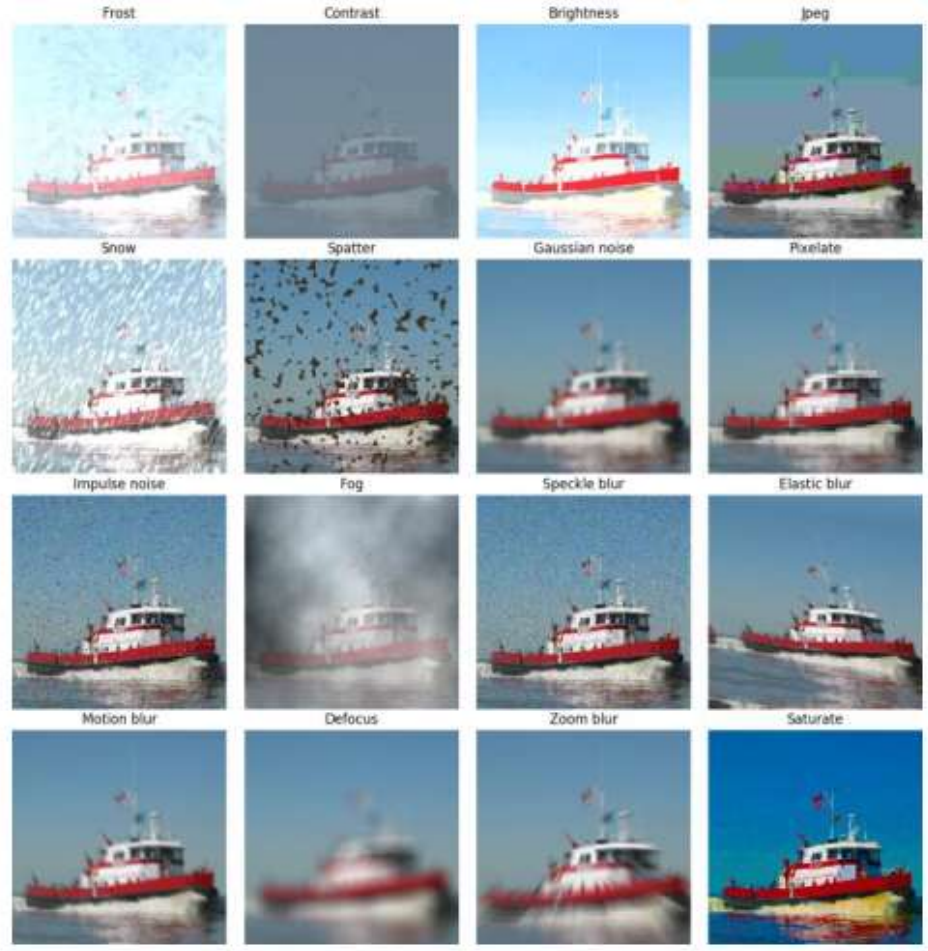
- How to evaluate the risk for data safety?

data properties, data error modes, risk analysis model, safety requirements, mitigation measures

Challenges in Model Validation

Test Based Validation

- Utilizes the validation data to demonstrate that the model generalizes to cases not present in the model learning stage.
- Specifically, those safety requirements associated with ensuring the robustness of models are evaluated on the independent validation data set
- The performance is maintained in the presence of adversarial conditions or signal perturbations.
- The test team should examine those cases which lie on boundaries/edge cases, or which are known to be problematic within the context to which the model is to be deployed.

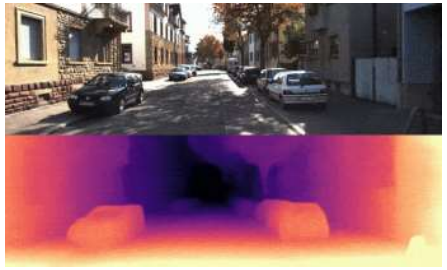


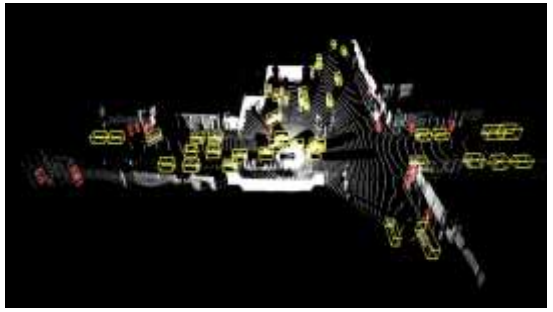
Augmented data with perturbations

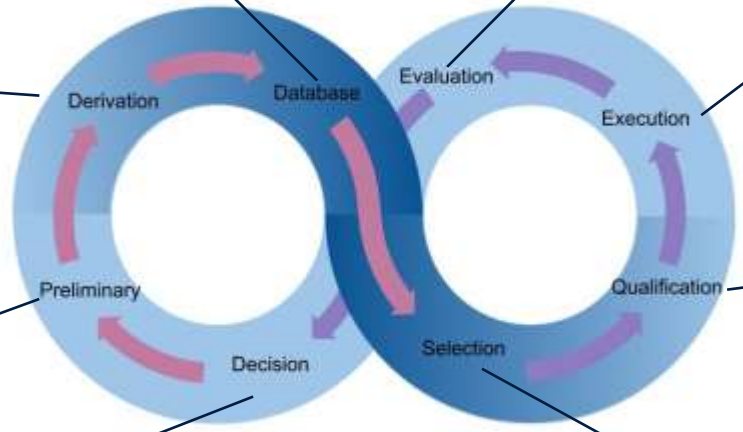
https://publications.jrc.ec.europa.eu/repository/handle/JRC119336

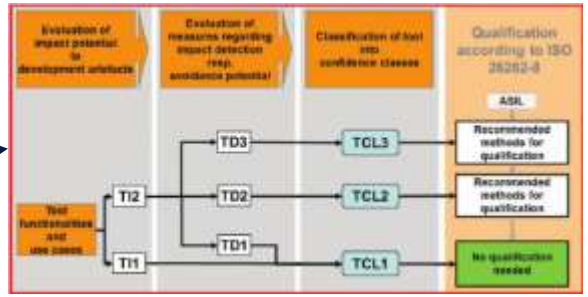
Method Overview

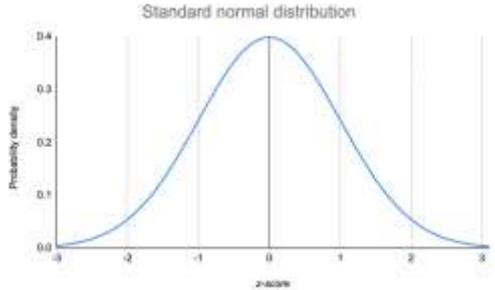
Scenario Based Approach for AI Based Perception Validation











https://blog.waymo.com/2019/08/waymo-open-dataset-sharing-our-self.html

THANKS