

# Reliable AI applied to perception systems in autonomous vehicles

*Cristiano Premebida*

*cpremebida@deec.uc.pt*

# Table of contents

- Intelligent vehicles (IV) progress: last 7 years
- Artificial / sensory perception systems for intelligent vehicles
- Reliable AI/ML for perception systems: critical-safe applications
- *Post-hoc* approaches for explainable AI
- Open problems, challenges, future perspectives

@ Disclaimer

This presentation is for academic and research purposes. The contents of this presentation are for non-commercial use. Some of the images/pictures shown in this presentation come from public sources.

# VEHITS 2018

4<sup>th</sup> International Conference on Vehicle Technology  
and Intelligent Transport Systems

Funchal, Madeira - Portugal · 16 - 18 March, 2018

## “AUTOCITS Pilot in Lisbon perspectives, challenges and approaches” \*

\* C.Premebida, et al.



AUTO  
C-ITS



Co-financed by the European Union  
Connecting Europe Facility

indra



POLITÉCNICA  
"Ingeniamos el futuro"



INSIA



Inria  
Inventeurs du monde numérique

ANSR  
AUTORIDADE NACIONAL  
SEGURANÇA RODoviÁRIA

IPN  
INSTITUTO PEDRO NUNES



UNIVERSIDADE DE COIMBRA



### Pilots

#### MADRID



- 4 connected vehicles
- 2 autonomous vehicles
- 3 Day 1 C-ITS services

#### LISBON



- 2 connected vehicles
- 2 autonomous vehicles
- 2 autonomous shuttles
- 3 Day 1 C-ITS services

#### PARIS



- 4 connected vehicles
- 1 autonomous vehicles
- 3 Day 1 C-ITS services



# From 2018 to today: key achievements in IV

Advancements in Aut. Driving

Vehicle-to-Everything (V2X)

Electric Vehicle (EV) Growth

AI and ML in Automotive Systems

Regulations and Safety Standards

Lidar Technology Advancements

Automated Delivery Vehicles

Sustainability and Green Technology in IV

AI-Based Traffic Management Systems

Advanced In-Car Entertainment and Experience



ChatGPT



Advancements in Autonomous Driving (AD) and Advanced Driver-Assistance Systems (ADAS)

Increased Connectivity and V2X Communication

Electric Vehicle (EV) Integration and Development

Software and Over-the-Air (OTA) Updates



# Autonomous vehicles – “not only cars”



## ‘Robot vehicles’





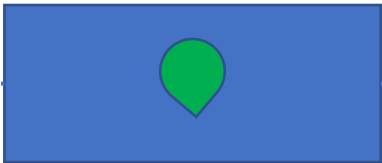
# Intelligent (robotic) vehicles: 20+ years of history



2004



2007



2010



Today

Big players in AD & ADAS:

- W...
- N...
- M...
- P...
- ...

## Sensors for perception system



# Perception systems

In automated-autonomous vehicles, or simply **intelligent vehicles (IV)**, **perception** designates a system that endows the vehicle with the ability to perceive, comprehend, and reason about the surrounding environment.

Perception systems are designed to cope with environment/surrounding\* understanding and are crucial for decision-making - sometimes in real-time – in **tasks** such as obstacle avoidance, lane detection, object detection, ADAS, and so on.

(\*) in real-world applications the surrounding / environment is subject to changes, disturbance, noise, interference – plus varying weather/environmental conditions (e.g., rain, dust, light, and son on).

Perception relies strongly on

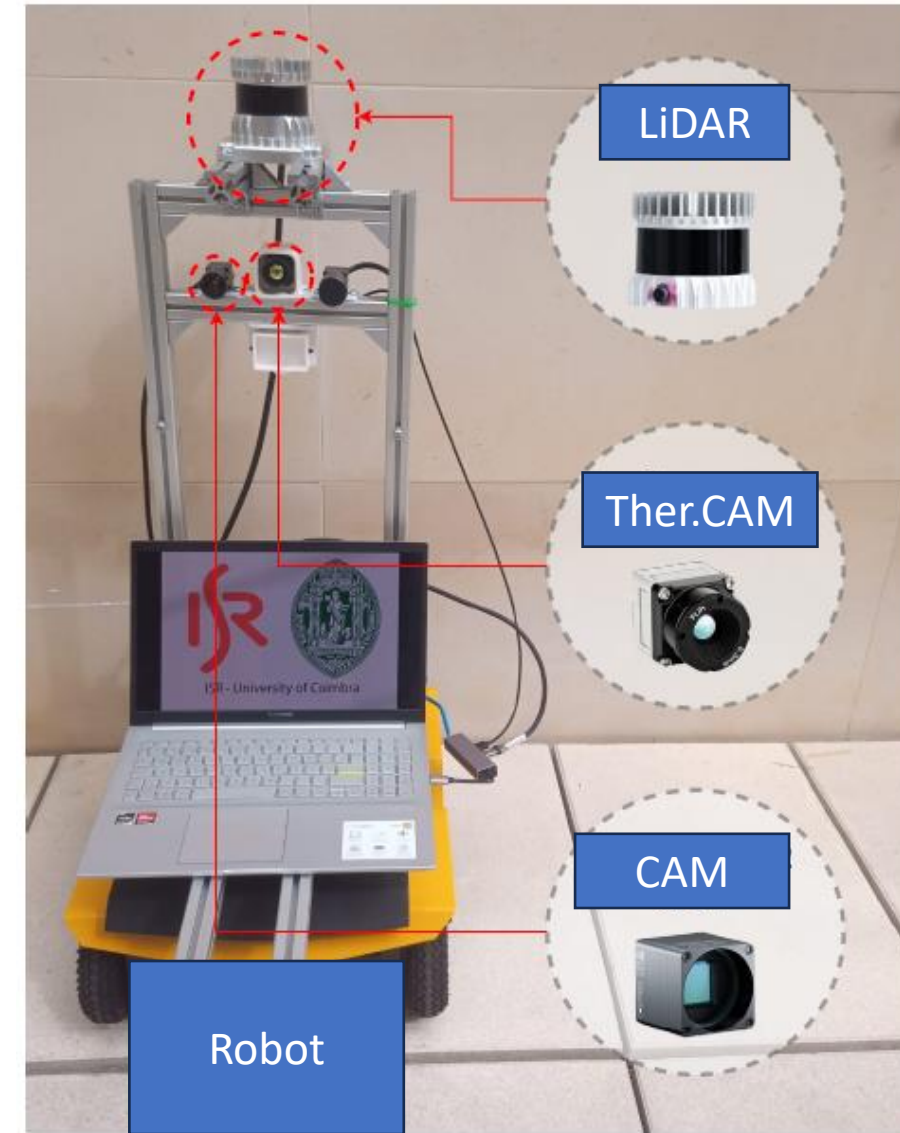
1. **sensory data** (from onboard sensors and some from the infrastructure) and
2. **Software**, algorithms, techniques: here **AI/ML** is the key element

# Perception systems - sensory data

In advanced vehicles, typically the **on-board** sensor are:

- Cameras (visual data)
- LIDAR (Light Detection and Ranging) (3D spatial mapping)
- Radar (detecting distance, speed, and position of objects)

Onboard sensory data can be complemented by **infrastructure-based** sensors:





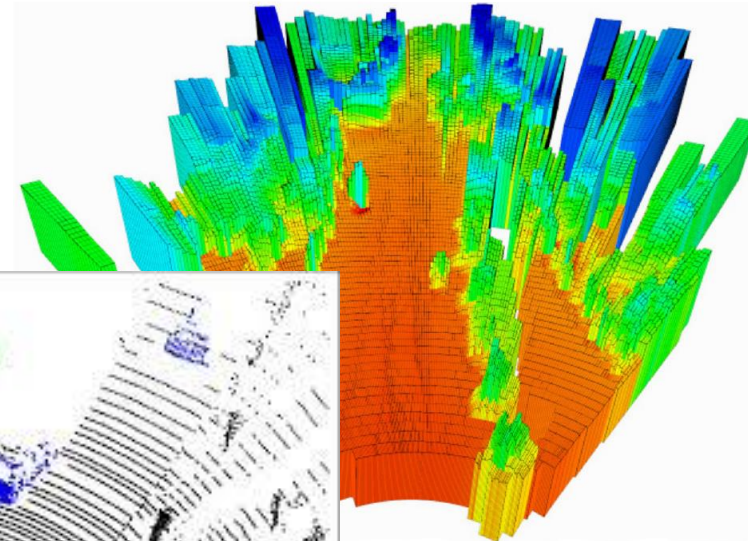
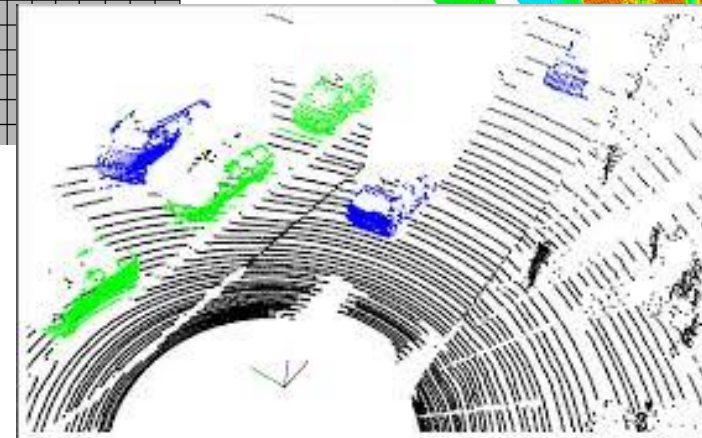
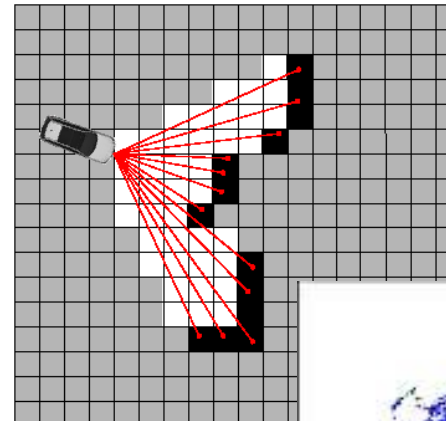
# Perception systems - sensory data

Data collected by the sensors should be processed in order to obtain a Representation i.e., **data representation mapping**.

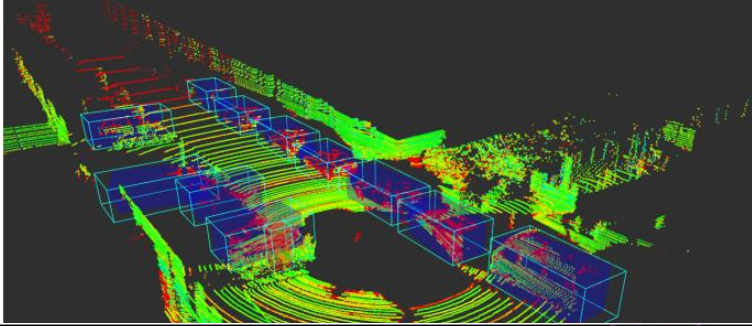
- Occupancy grid mapping (2D representation): Radar, LiDAR
- Bird-eye view (BEV): LiDAR
- Calibration of cameras: common ref. system
- 3D representation e.g, voxels

**Appropriated representation and mapping depends on:**

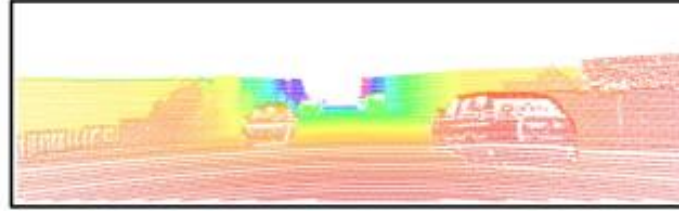
- Time alignment / synch.
- Spatial alignment i.e., calibration
- Common reference system
  - Data fusion



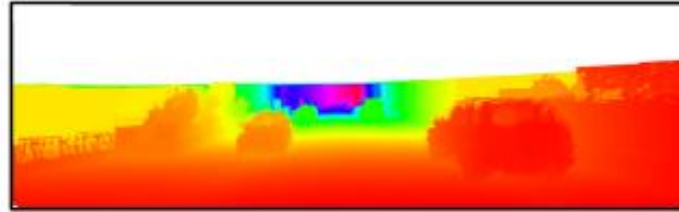
# LiDAR representation



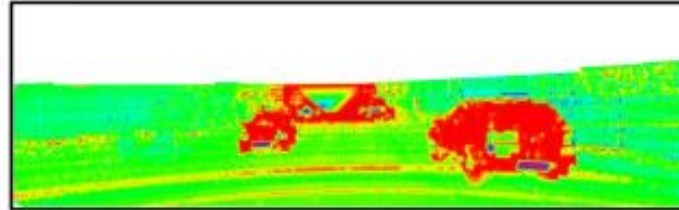
(a) RGB camera image



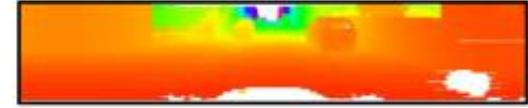
(b) LiDAR sparse depth map



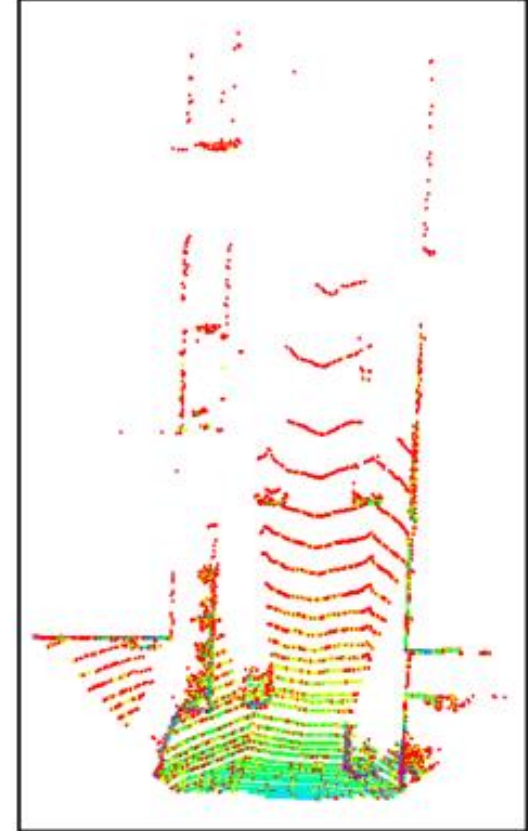
(c) LiDAR dense depth map



(d) LiDAR dense intensity map



(e) LiDAR spherical map



(f) LiDAR BEV density map

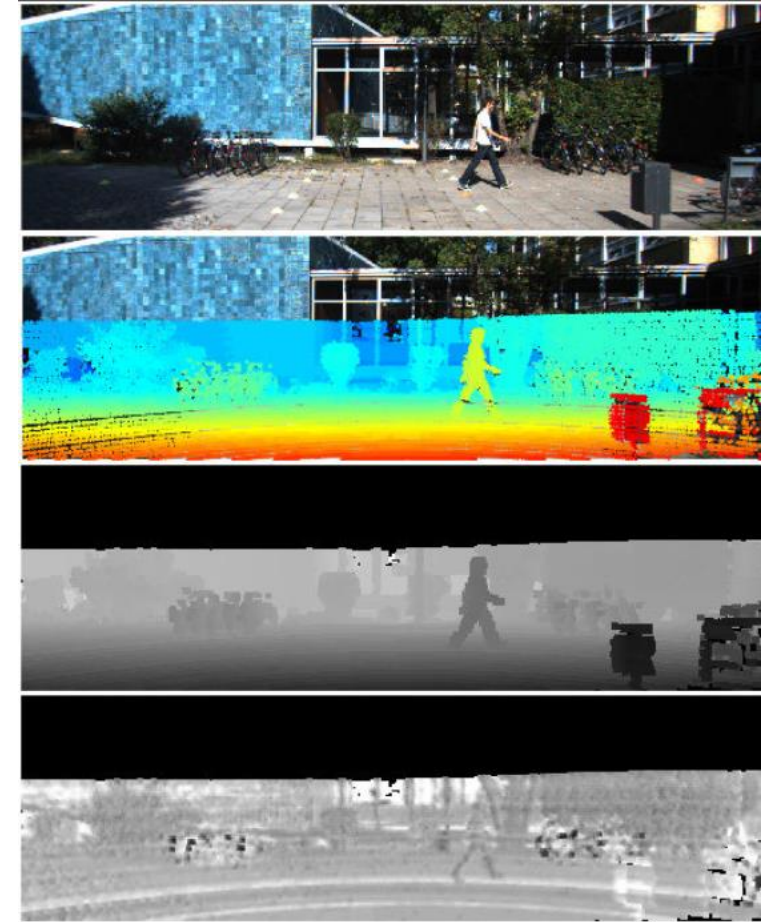
# Perception: task-specific + AI approaches

## Perception systems for intelligent vehicles (including robotic-vehicles)

- (multisensory – multimodal) Object detection
- Detection and tracking objects
- 3D object detection
- Environment representation / mapping (2D, 2.5D, 3D representations)
- Sensor-fusion (e.g., camera + LiDAR + radar data)

## Machine learning/AI (as a key-component in perception systems)

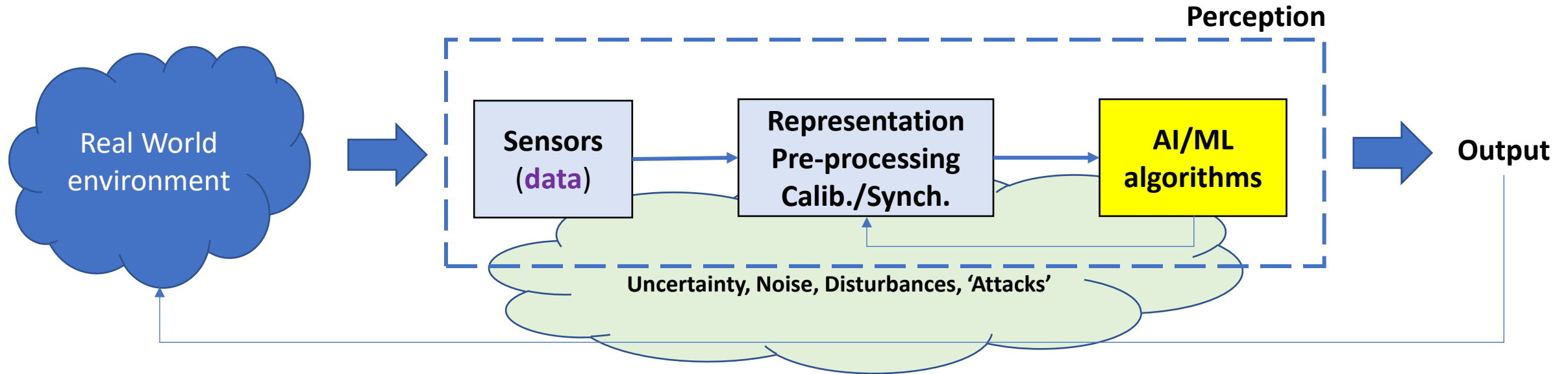
- Segmentation / clustering (i.e., unsupervised learning)
- Generative methods (based on conditional distribution e.g., BayesNN)
- Discriminative methods (e.g., RandForest, SVM, LDA, MLP)
- **Deep learning** (CNN-based architectures)
- Combination/mixture of AI-experts
- Reinforcement learning





# Perception systems

Main goal: to extract meaningful information from the measurements (**data**) and/or info (higher-level data) from *exteroceptive*\* **sensors** mounted on-board the robot and/or from the 'infrastructure'.



How to model/characterize the uncertainties which are inherent to the sensors, data, and consequently the AI models ?

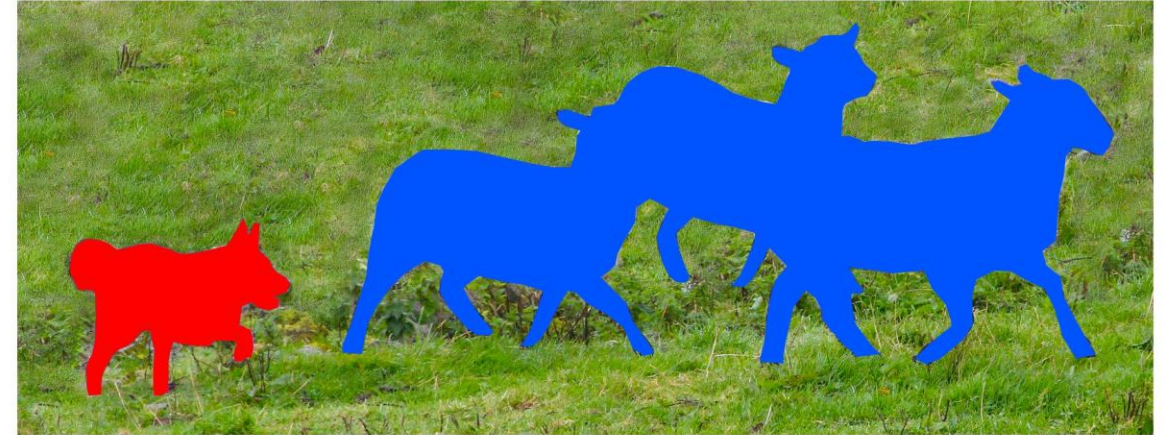
\* The so called **proprioceptive** sensors are, for example, : encoders, IMU, INSS.

# Perception - SW, algorithms, models, techniques

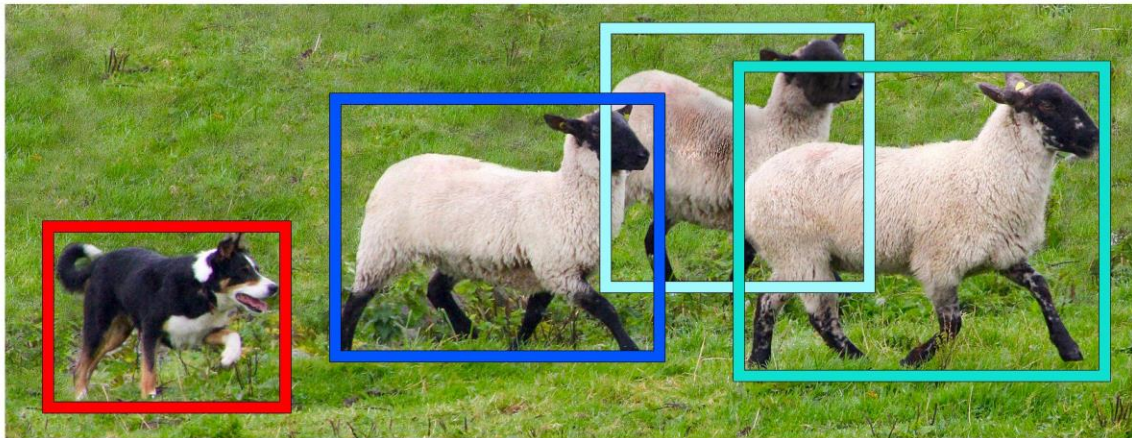
Perception is not only about colour vision/camera systems and object detection...



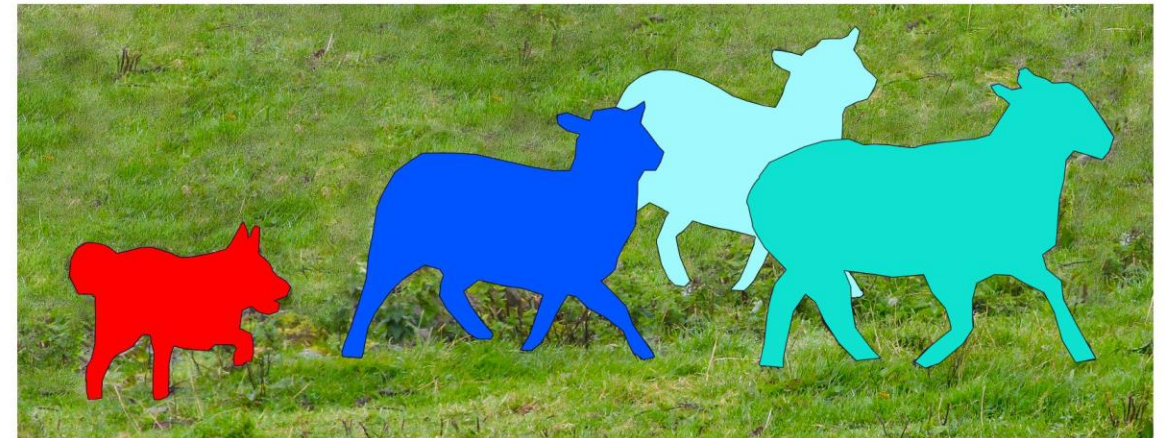
**Image Recognition**



**Semantic Segmentation**



**Object Detection**



**Instance Segmentation**

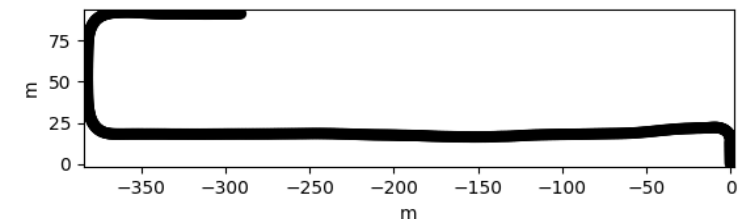


# Perception - SW, algorithms, models, techniques

@credits: Tiago Barros

... it is present in many applications:

- ADAS (e.g.: lane detection, parking assist., traffic signal recognition)
- Mapping, SLAM, localization
- Place recognition
- Lane detection
- Agents' intention/interaction prediction
- Sophisticated ACC
- ... besides cameras, other sensors modalities are involved.

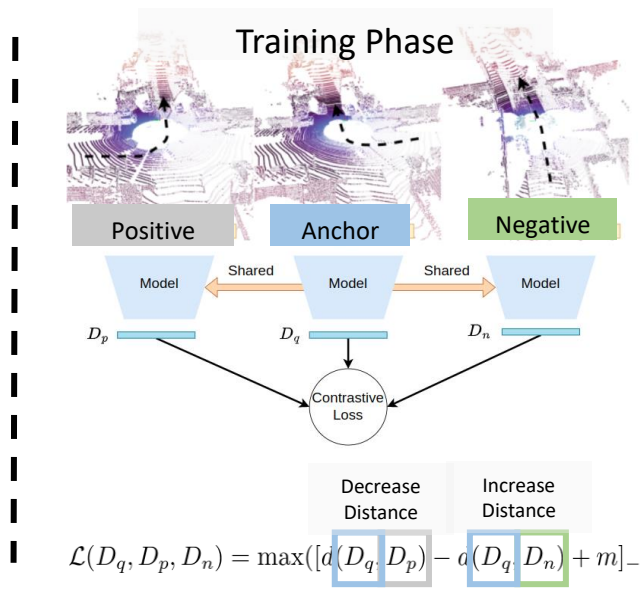
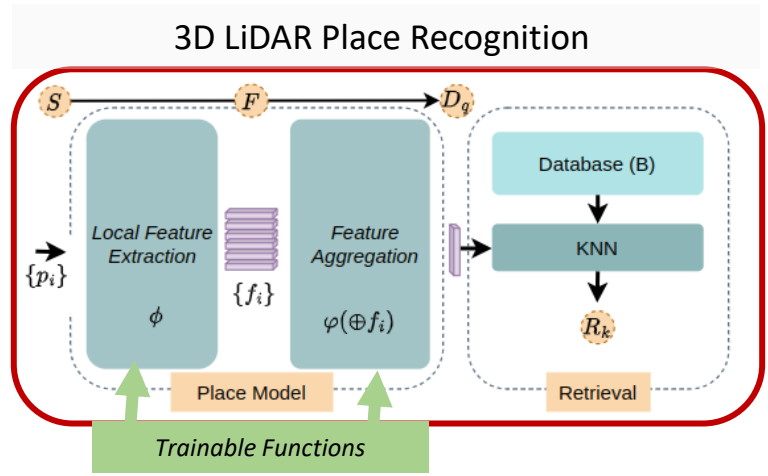
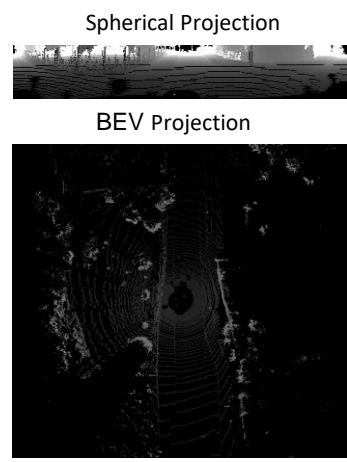
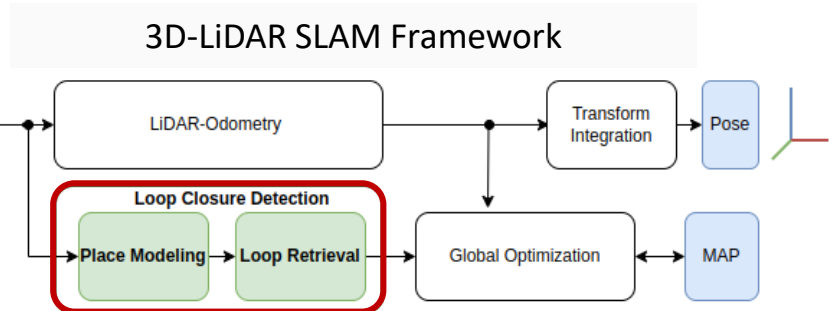
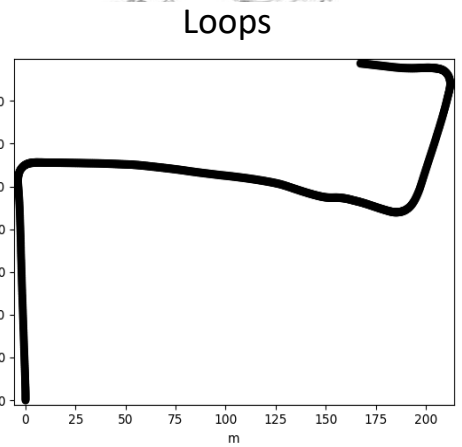


Perception systems -> critical-safety applications -> reliable solutions



# Deep Learning-based Place Recognitions

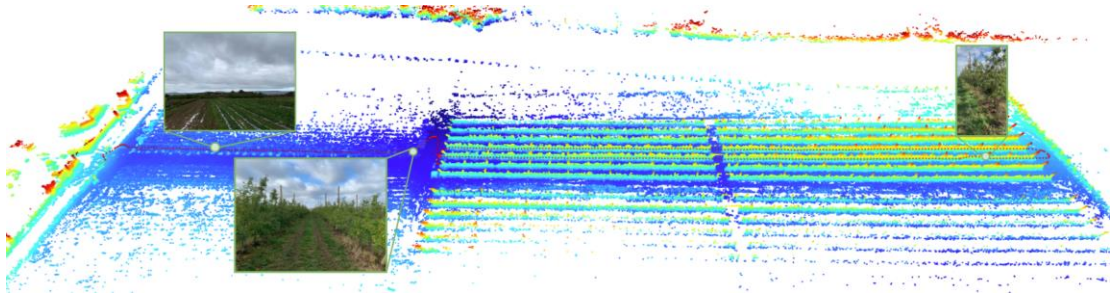
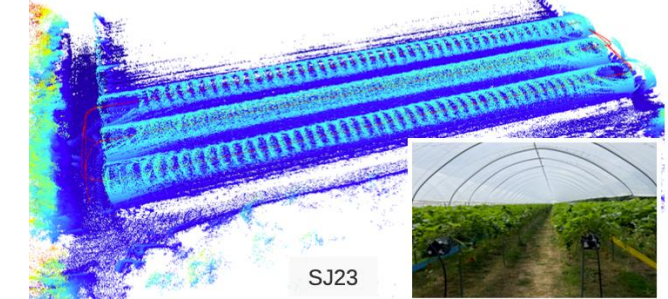
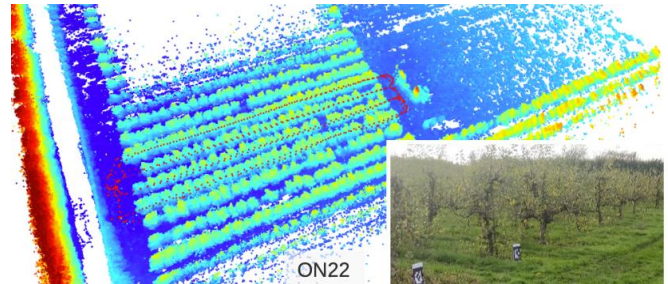
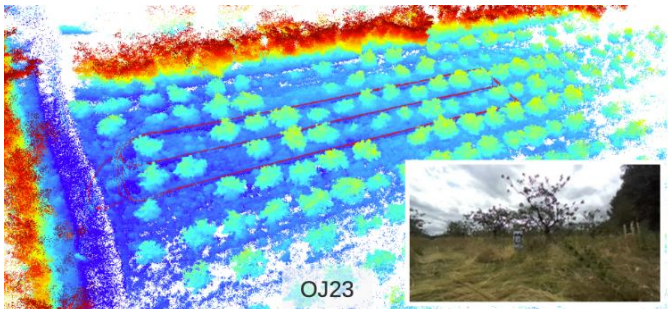
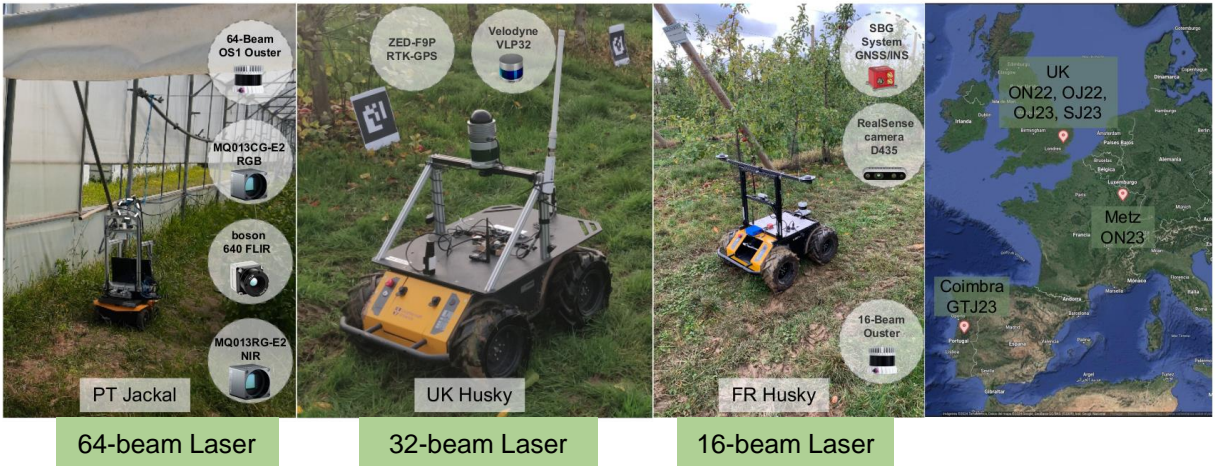
Place recognition is a perception-based global localization approach that finds revised places by matching descriptors instead of point clouds.



[1] Barros, Tiago, Ricardo Pereira, Luís Garrote, Cristiano Premebida, and Urbano J. Nunes. "Place recognition survey: An update on deep learning approaches." arXiv preprint arXiv:2106.10458 (2021).  
[2] Barros, Tiago, Luís Garrote, Martin Aleksandrov, Cristiano Premebida, and Urbano J. Nunes. "TRer: A Lightweight Transformer Re-Ranking Approach for 3D LiDAR Place Recognition." In IEEE ITSC, pp. 2843-2849. 2023.  
[3] T. Barros, L.Garrote, R.Pereira, C.Pemebida, U.J. Nunes. "Attdlnet: Attention-based deep network for 3d lidar place recognition." In Iberian Robotics conference, pp. 309-320. Cham: Springer International Publishing, 2022.



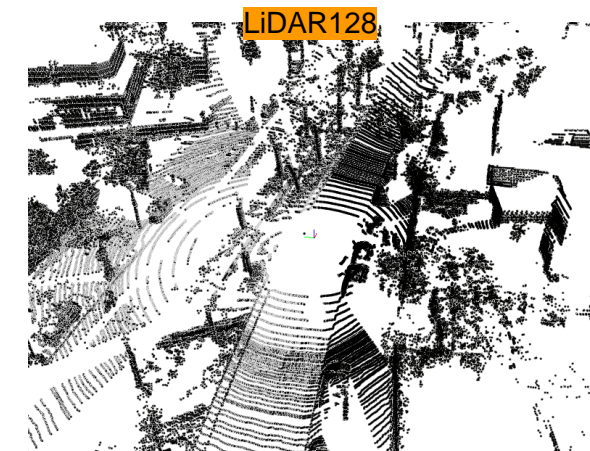
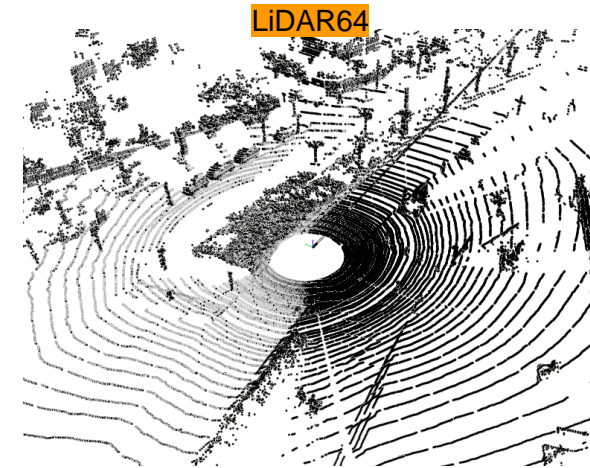
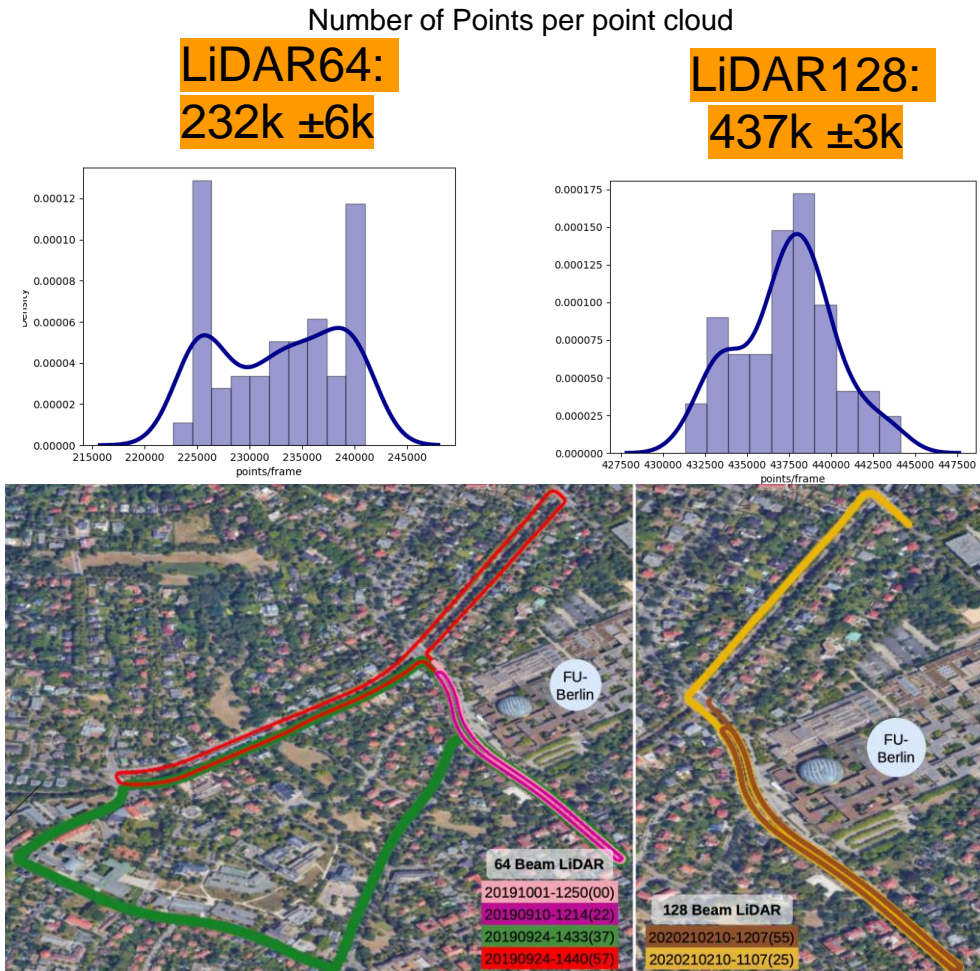
# Perception - LiDAR having distinct resolution





# Perception - LiDAR having distinct resolution

What is the impact of 3D LiDAR resolution in SLAM algorithms and DL-based perception models?

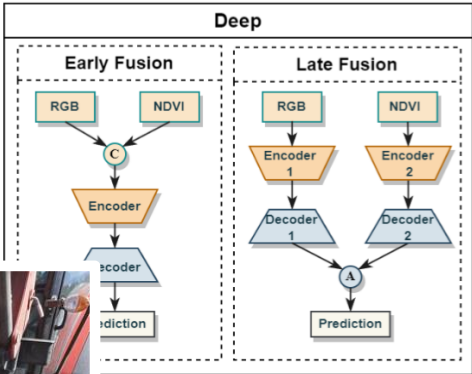
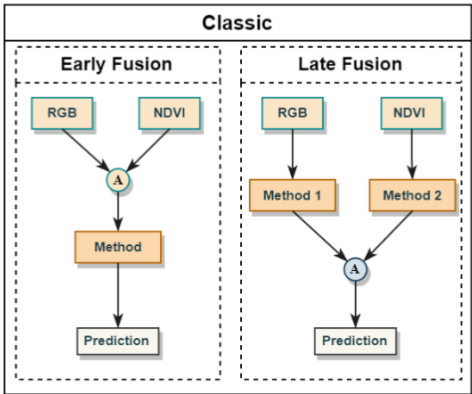




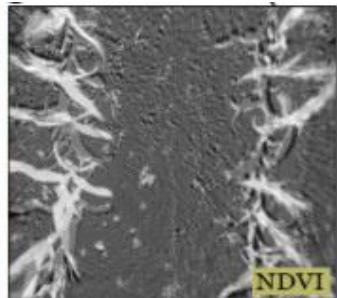
# Perception – agricultural vehicles-robots



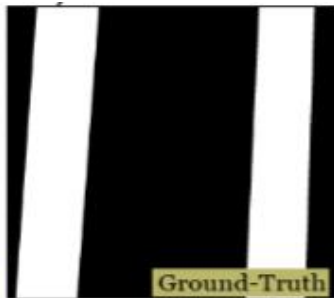
(a)



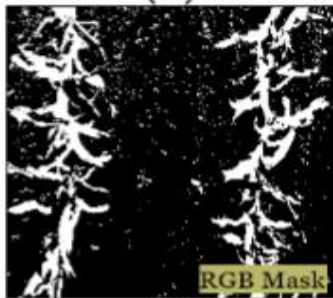
(a)



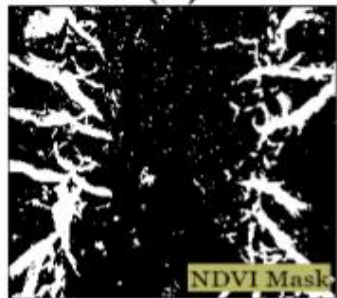
(b)



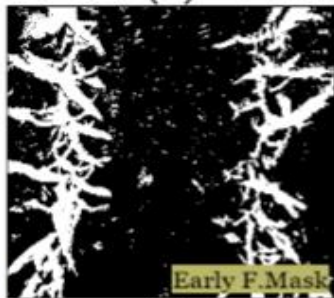
(c)



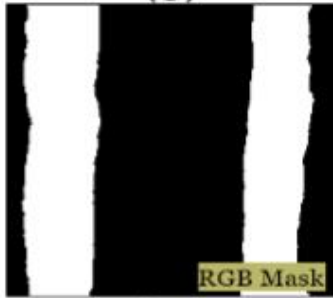
(g)



(h)



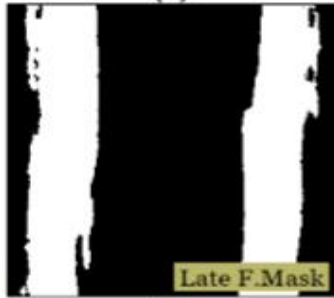
(i)



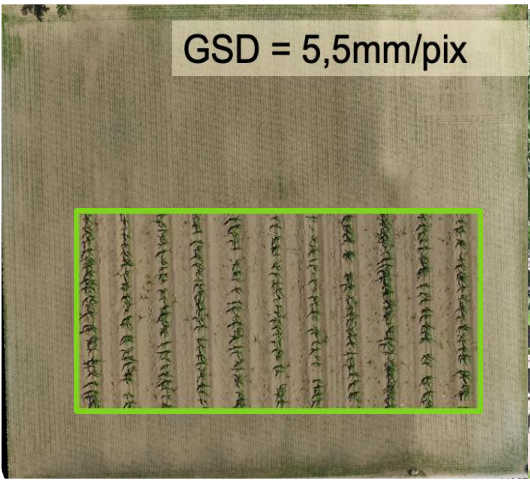
(m)



(n)

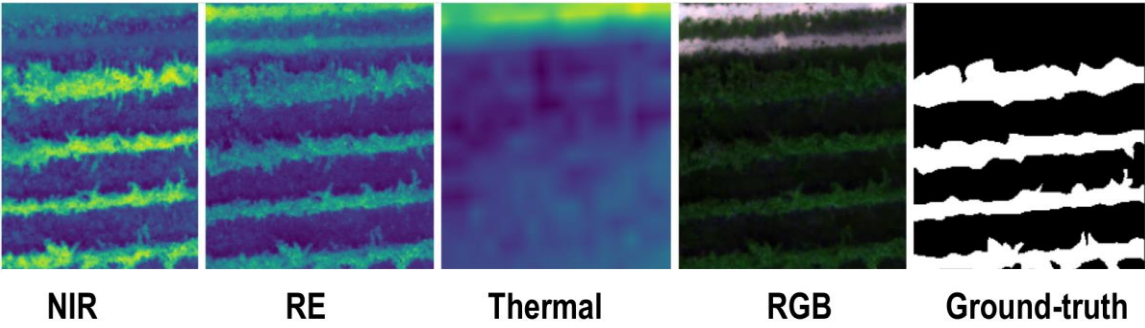
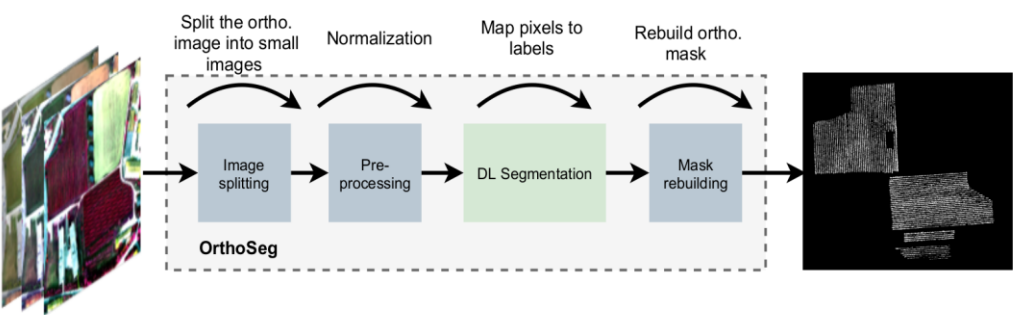
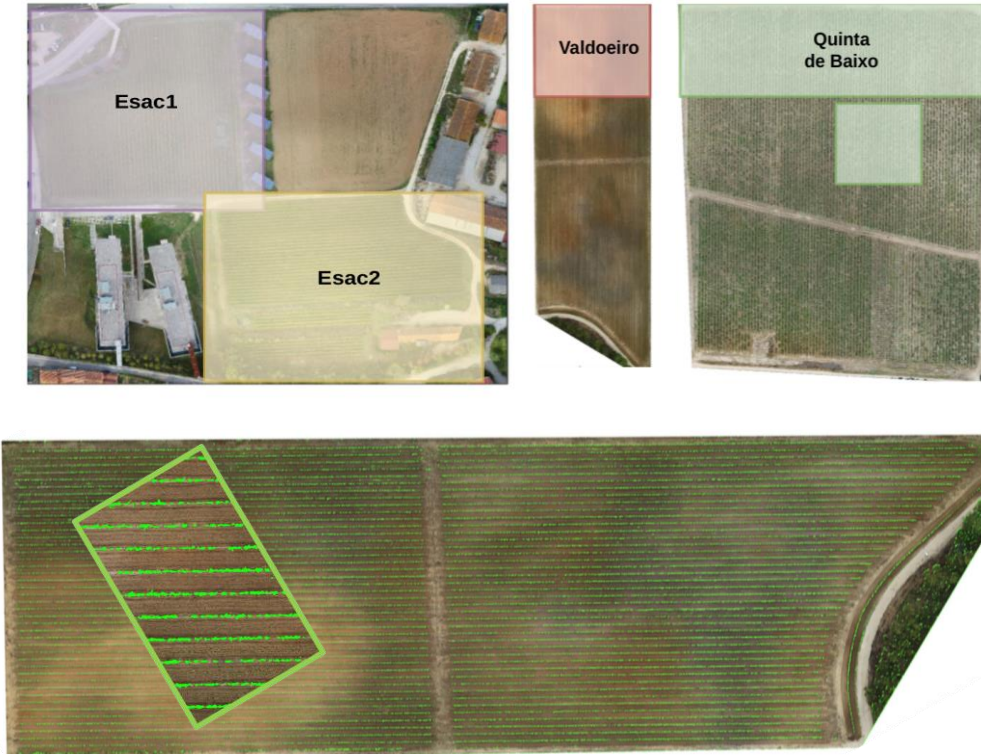
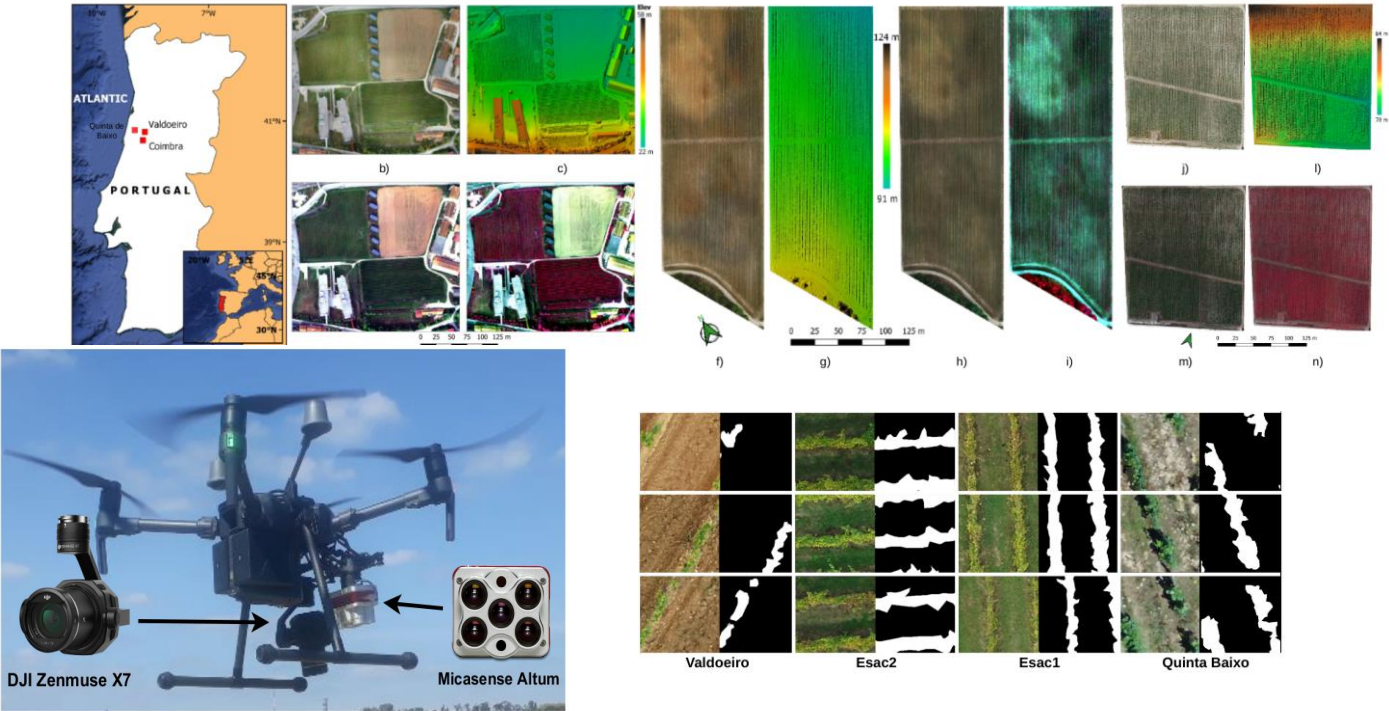


(o)





# Perception – UAV / drones



Cybonic/DL\_vineyard\_segmentation\_study



## Part II

Uncertainty calibration, Overconfidence – Reliable Deep architectures

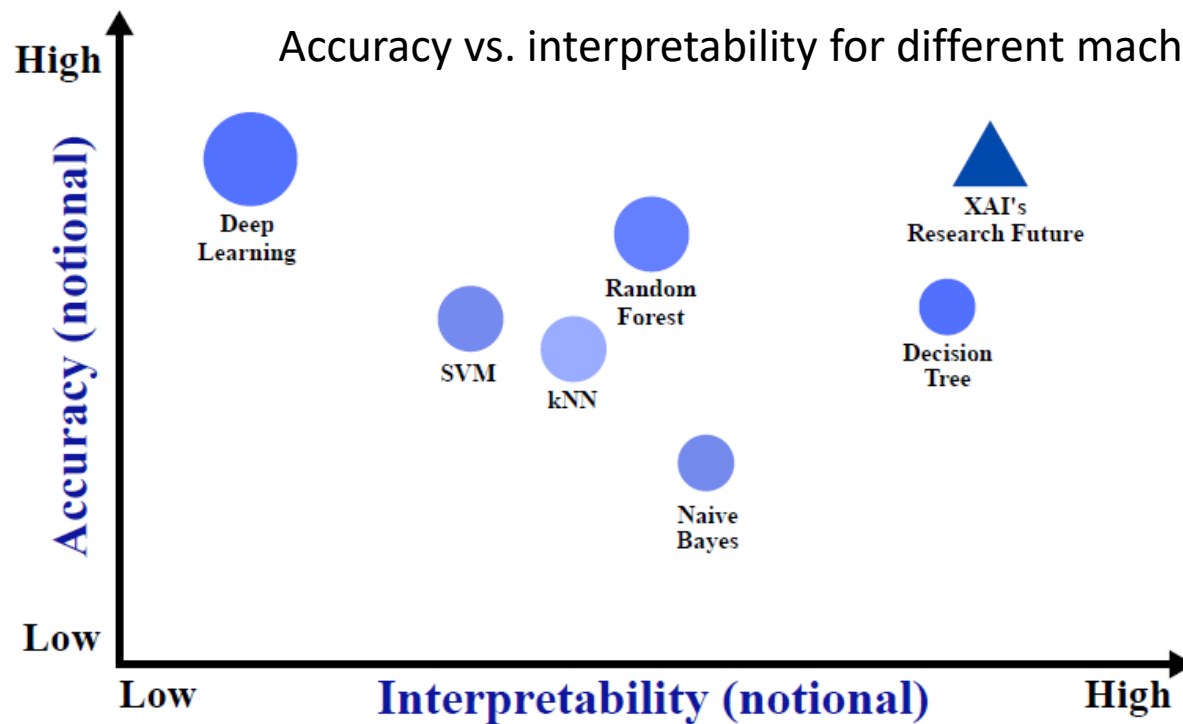


# Reliable AI



## Trustworthy AI -> Intelligent Vehicles & Robotics

- + Explainable AI (**XAI**): methods and techniques for making AI systems more transparent and understandable to humans.
- + Ethical considerations in AI: addressing the ethical implications of AI, such as bias, privacy, and autonomy.
- + Safety and security in AI and autonomous systems: exploring the risks and **challenges of AI and autonomous systems**, and methods for mitigating them.



- transparency and explainability of advanced AI and ML models
- interpretability tends to be low in most DL approaches

From [\*\*]

- Interpretability and explainability have escaped a clear universal definition
- Other terms that are synonymous to interpretability: intelligibility, and understandability
- More recently (XAI): is closely tied with interpretability; and many authors do not differentiate between the two
- [\*\*\*] interpretable ML focuses on designing models that are inherently interpretable; whereas **explainable ML** tries to provide post hoc explanations for existing black box models

[\*] Plamen P. Angelov, E.A. Soares, R. Jiang, N. I. Arnold, and P. M. Atkinson. "Explainable artificial intelligence: an analytical review." *WIREs Data Mining and Knowledge Discovery*, 2021.

[\*\*] R. Marcinkevics, Julia E. Vogt. "Interpretability and Explainability: A Machine Learning Zoo Mini-tour". *ArXiv*, 2023.













[\*\*\*] C. Rudin, "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead," *Nature Machine Intelligence*, 2019.



# Machine learning – DL: undesirable aspects

- Lack of proper Uncertainty quantification
- Overconfidence problem in Deep Neural Networks (DNNs)
- The DNNs' outputs are commonly "normalized scores" but it does not guarantee of being proper probabilities.

DNNs tend to be **over-confident** in their predictions  
... and in most cases, we do not know why.

	confidence	prediction	correct
	 80%  20%	healthy	✓
	 80%  20%	healthy	✓
	 80%  20%	healthy	✓
	 80%  20%	healthy	✗

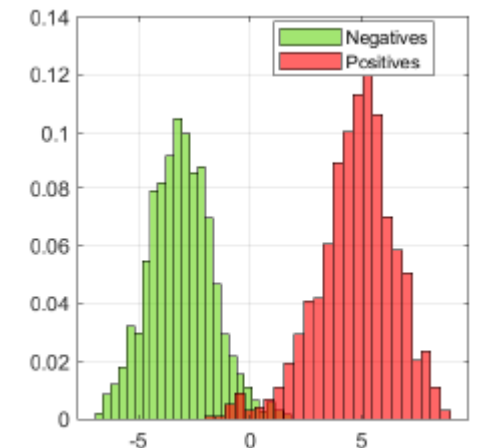
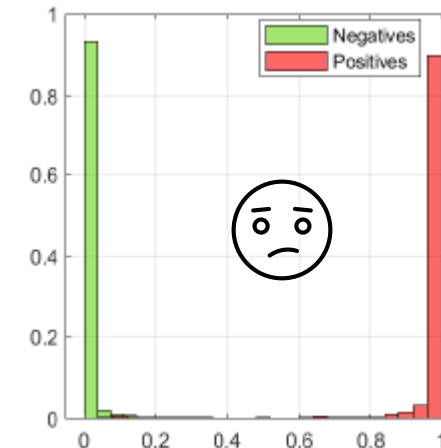
# Machine learning – DL: undesirable aspects

## Why probability is important in ML-based perception for IV / AD?

- Most of the modern deep learning (DL) algorithms, and available software packages, tend to lack explainability in terms of probability
- They might generalize in unforeseen and **overconfident** ways on out-of-training-distribution [\*].

So, the presumable inability of ANNs to answer “I don’t know”

... is problematic in fields where their predictions have **critical implications**, such autonomous driving, ADAS, robotics.



[\*] “Hands-on Bayesian Neural Networks - a Tutorial for Deep Learning Users”

LV Jospin; et. al. – 2020

<https://arxiv.org/abs/2007.06823>



# Reliable ML applied to IV perception

## Calibration of ML/DL models

[\*]

“Real-world applications of machine learning (ML) systems require a thorough look into the reliability of the learning models and consequently to their uncertainty calibration (also referred as confidence calibration or simply calibration).

In addition to having highly accurate classification models, the user should be able to "trust" their predictions, specially when dealing with critical application domains, where wrong decisions can result in potentially dramatic consequences.”

Examples:

- **Autonomous driving**
- Robotics
- Medical diagnosis

[\*]

P.Conde, C.Premebida (2022). “Adaptive-TTA: accuracy-consistent weighted test time augmentation method for the uncertainty calibration of deep learning classifiers”. In. Proc. 33rd British Machine Vision Conference (BMVC).



Classes	Car	Cyc	Ped
LeNet CNN	99.05%	1.95%	0.00%
AlexNet CNN	99.98%	0.2%	0.00%
InceptionV3 CNN	99.76%	0.15%	0.09%
EfficientNetB1 CNN	98.85%	0.12%	0.03%
Vision Transform	99.98%	0.02%	0.00%
MLP Mixer	91.54%	8.43%	0.03%

# Reliable ML applied to IV perception

SOTA on object recognition and detection use deep architectures; DNNs provide normalized prediction scores (the outputs) via a SoftMax or Sigmoid layer i.e., the prediction values are in the interval of  $[0, 1]$ .

Usually, such models/architectures are implemented through deterministic neural networks thus, the prediction itself does not consider uncertainty for the predict class of an object during the decision-making.

Therefore, evaluating the prediction confidence or uncertainty is crucial in decision-making whereas an erroneous decision may have severe implications.

Techniques to mitigate the overconfident problem:

- **Calibration**
- Regularization

Calibration acts directly in the network output prediction (**post-hoc** calibration\*), while regularization aims at penalizing network weights through a variety of methods, adding parameters or terms directly to the cost/loss function.

\* “...adjusts the output logits of a pre-trained model...”



# Reliable ML – confidence calibration

[\*] “confidence calibration is the problem of predicting probability estimates representative of the true correctness likelihood”.

Intuitively, the idea of calibration can be formulated as follows: let  $h$  to be a ML model  $h(X) = (\hat{Y}, \hat{P})$ .

Considering a distribution generated over the  $K$  possible classes of the model for a given input  $X$ , where  $\hat{Y}$  is the predicted class with an associated predicted confidence  $\hat{P}$ .

The *perfect calibration* is given by:

$$\mathbb{P}(\hat{Y} = Y | \hat{P} = p) = p, \quad \forall p \in [0,1]$$

The expression above can be better understood by a toy example [\*]:

“given 100 predictions, each with confidence of 0.8, we expect that 80 should be correctly classified.”

Thus, for every subset of predicted samples of a given class with score values equal to  $S$ , the proportion of samples that actually belongs to that class is  $S$ .

# Reliable ML – confidence calibration

[\*]

“perfectly calibrated models are those for which the predicted confidence for each sample is equal to the model accuracy” ...

“an **over-confident** model tends to yield predicted confidences that are larger than its accuracy,

whereas an **underconfident** model displays lower confidence than the model’s accuracy.”

The calibration algorithm is an approximation process that depends on a calibration measure, which can be obtained by separating the predictions into multiple bins, as **Reliability Diagram**.

The scores (predicted values) are grouped into M bins (histogram) in reliability diagrams. Each example (classification score of an object) is allocated within a bin according to the maximum prediction value (prediction confidence).

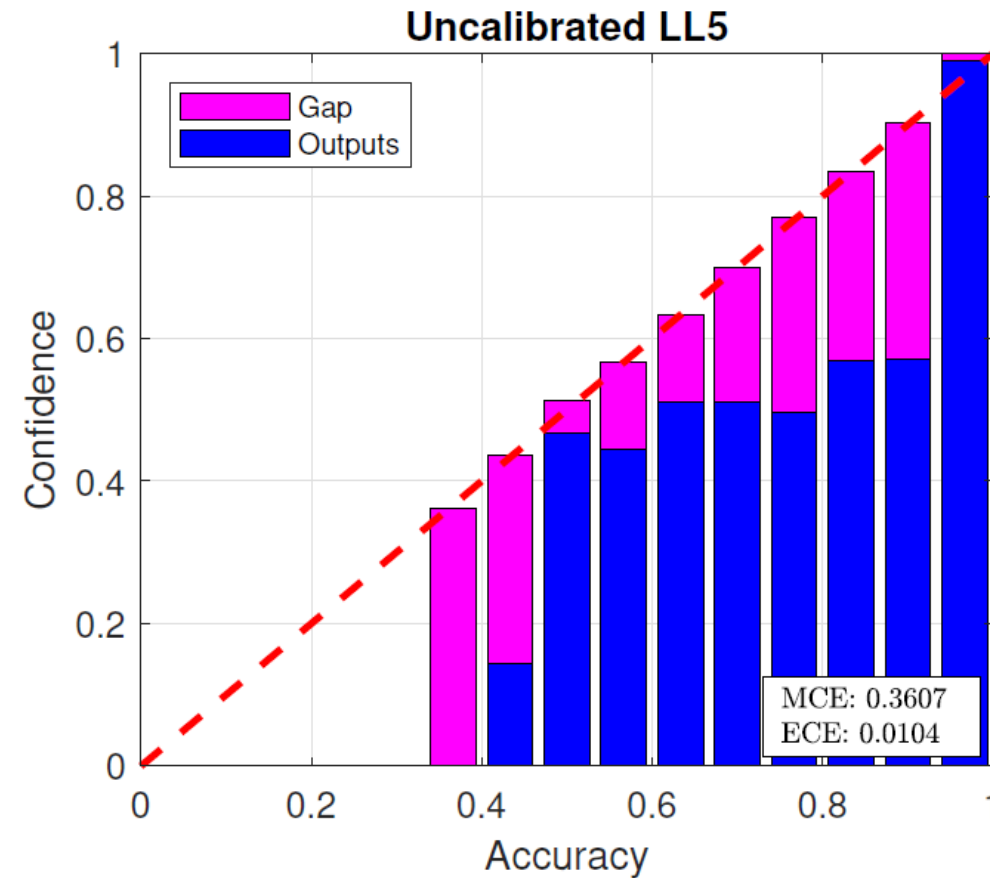


# Reliable ML – reliability and calibration errors

## Reliability Diagram

[\*] “Typically, post-calibration predictions are analysed in the form of **reliability diagram** representations, which illustrate the relationship of the model’s prediction scores regarding the true correctness likelihood/probability.”

Reliability diagrams show the expected accuracy of the samples as a function of confidence i.e., the maximum value of the prediction function.”



[\*] G Melotti, C Premebida, JJ Bird, DR Faria, N Gonçalves (2022). “Reducing Overconfidence Predictions in Autonomous Driving Perception”. IEEE Access.

# Reliable ML – reliability and calibration errors

## Reliability Diagram – toy example

	i=0	1	2	3	4	5	6	7	8	9
$P(y_i = 0 x_i)$	0.1	0.8	0.3	0.6	0.2	0.9	0.8	0.2	0.5	0.1
$P(y_i = 1 x_i)$	0.9	0.2	0.7	0.4	0.8	0.1	0.2	0.8	0.5	0.9

Considering  $P(y_i = 1|x_i)$ , the probabilities are then partitioned into K subsets, in which each subset represents a disjoint interval of probabilities between 0 and 1.

If K=3, then we have 3 sets: [0.0 – 0.33), [0.33-0.66), [0.66 – 1.0]

Partitioned sets	
Set1	( $i = 1, 5, 6$ ) -> (0.2 , 0.1 , 0.2)
Set2	( $i = 3, 8$ ) -> (0.4 , 0.5)
Set3	( $i = 0, 2, 4, 7, 9$ ) -> (0.9 , 0.7, 0.8, 0.8, 0.9)

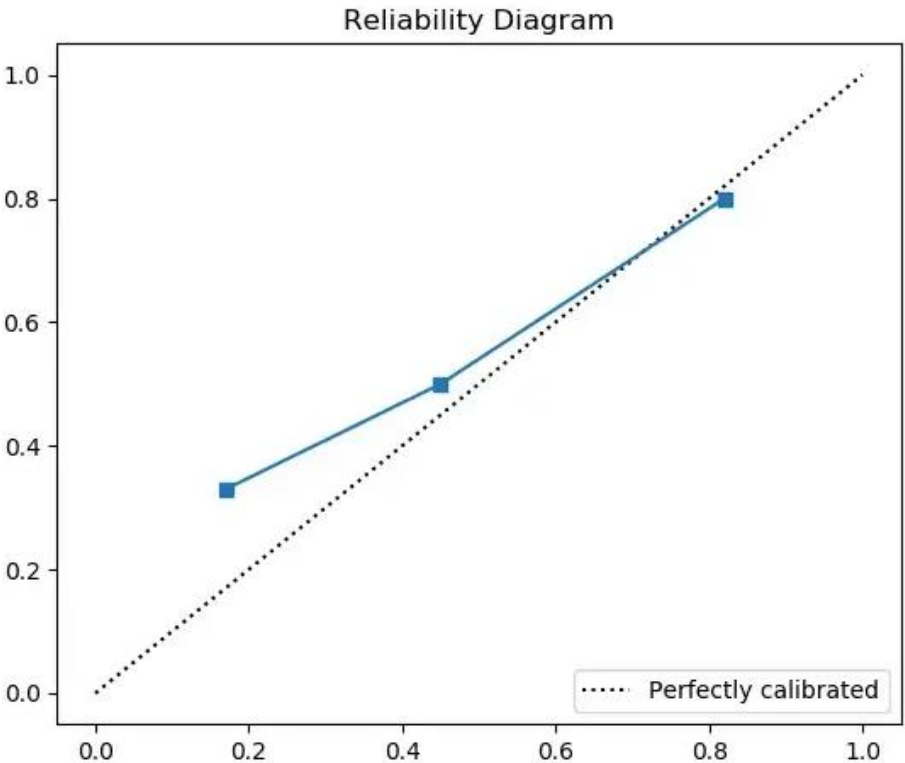
[Partially @Credits] Xiang Jiang (2020); “A brief introduction to uncertainty calibration and reliability diagrams”, online: <https://towardsdatascience.com/introduction-to-reliability-diagrams-for-probability-calibration-ed785b3f5d44>

# Reliable ML – reliability and calibration errors

## Reliability Diagram – toy example

For each  $K^{\text{th}}$  subset, two estimates are computed: (a) average of the predicted probabilities, (b) the relative frequency of positive examples (i.e., **Accuracy** in ML applications).

Partitioned sets	
Set1	( $i = 1, 5, 6$ ) -> (0.2, 0.1, 0.2)
Set2	( $i = 3, 8$ ) -> (0.4, 0.5)
Set3	( $i = 0, 2, 4, 7, 9$ ) -> (0.9, 0.7, 0.8, 0.8, 0.9)



Sets:	1	2	3
Average predictions	0.17	0.45	0.82
Relative Freq. of "1"	1/3	0.50	0.80

- (i) when the average predictive probability is 0.17, about 33% of the predictions are positive;
- (ii) when the average predictive probability is 0.45, about 50% of the predictions are positive;
- (iii) when the average predictive probability is 0.82, 80% of the predictions are positive.



# Reliable ML –calibration errors

## ECE – Expected Calibration Error

Expected Calibration Error, Overconfidence Error, Max. calib. Error, ...

### Notations

Predictions/probabilities from a model are grouped into  $M$  interval bins of equal size

$B_m$  is the set of samples whose prediction scores fall into bin  $m$

$y_i$  and  $\hat{y}_i$  are true label vector and prediction vector, respectively

$\hat{p}_i$  is the confidence/“probability” (winning score) of sample  $i$

$n$  is the total number of samples in all the bins

The accuracy and confidence of  $B_m$  are defined as

$$\text{acc}(B_m) = \frac{1}{|B_m|} \sum_{i \in B_m} \mathbf{1}(\hat{y}_i = y_i)$$

$$\text{conf}(B_m) = \frac{1}{|B_m|} \sum_{i \in B_m} \hat{p}_i$$

# ECE – Expected Calibration Error

$$\text{acc}(B_m) = \frac{1}{|B_m|} \sum_{i \in B_m} \mathbf{1}(\hat{y}_i = y_i)$$

$$\text{conf}(B_m) = \frac{1}{|B_m|} \sum_{i \in B_m} \hat{p}_i$$

- $\text{conf}(B_m)$  is just the average confidence/probability of predictions in that bin
- $\text{acc}(B_m)$  is the fraction of the correctly classified examples  $B_m$

The **Expected Calibration Error (ECE)** is then defined as:

$$\text{ECE} = \sum_{m=1}^M \frac{|B_m|}{n} \left| \text{acc}(B_m) - \text{conf}(B_m) \right|$$

Maximum Calibration Error (**MCE**):

$$\text{MCE} = \max_{m \in \{1, \dots, M\}} |\text{acc}(B_m) - \text{conf}(B_m)|$$

**Overconfidence Error (OE)**

$$\text{OE} = \sum_{m=1}^M \frac{|B_m|}{n} \left[ \text{conf}(B_m) \times \max \left( \text{conf}(B_m) - \text{acc}(B_m), 0 \right) \right]$$

penalizes predictions by the weight of the confidence but only when confidence exceeds accuracy ie, overconfident bins incur a high penalty.

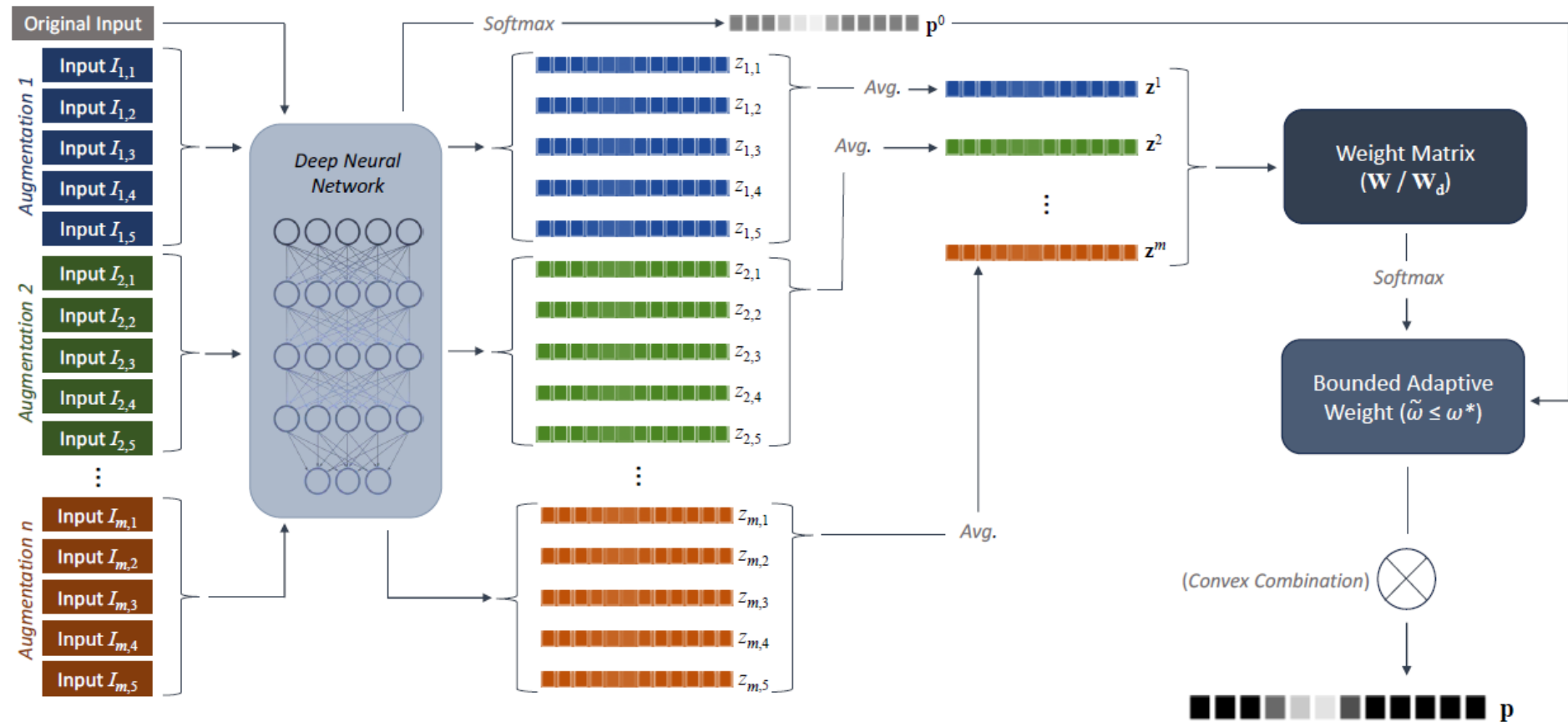
<Code>

# Post-hoc calibration techniques

Post-hoc calibration techniques: designed to address uncertainty calibration without the need of retraining the DL models.

- Temperature scaling (an extension of the Platt scaling algorithm)
- Histogram binning
- Isotonic regression

- **Test Time Augmentation**



P.Conde, C.Premebida. "Adaptive-TTA: accuracy-consistent weighted test time augmentation method for the uncertainty calibration of deep learning classifiers". In. BMVC, 2022.



# Open problems, challenges, future perspectives

- AD, ADAS, AI-based perception for: buses, trains, ships, ... military vehicles/systems

<https://www.metromondego.pt/pt/metrobus>



<https://railway-news.com/atc-3540-3750-computers-transport-intelligently/>



<https://www.europarl.europa.eu/topics/en/topic/artificial-intelligence>



Photo by the European Defence Agency

# Open problems, challenges, future perspectives

- More and more data vs representativeness
- Heterogeneity of data source: cameras, LiDARs, Radars, V2X, and so on
  - Distinct **resolution**, plus diverse functional operation and performance
  - Difference in **time** acquisition / frame-rate
  - Data **representation**
- Generalization to new/unseen conditions
  - Distribution shift problem
  - Continuous update of models
- AI + autonomous systems -> ...
  - AI/ML is evolving exponentially fast
  - When combined (embodied) with vehicles-robots: several opportunities to explore
  - Defence/military systems will be of concern?



# THANK YOU

## Questions?

Acknowledgement:



*Tiago Barros*



*Pedro Conde*



*Kennedy Mota*



*Wilgo Cardoso*



*Prof. Urbano Nunes*

