Recognizing New Visual Categories in Real Open World Environments

Doctoral Examination Committee:

Prof. Bottino Andrea
Prof. Ciliberto Carlo
Prof. Iocchi Luca
Prof. Montrucchio Bartolomeo
Prof. Zanuttigh Pietro

Supervisor:

Prof. Caputo Barbara

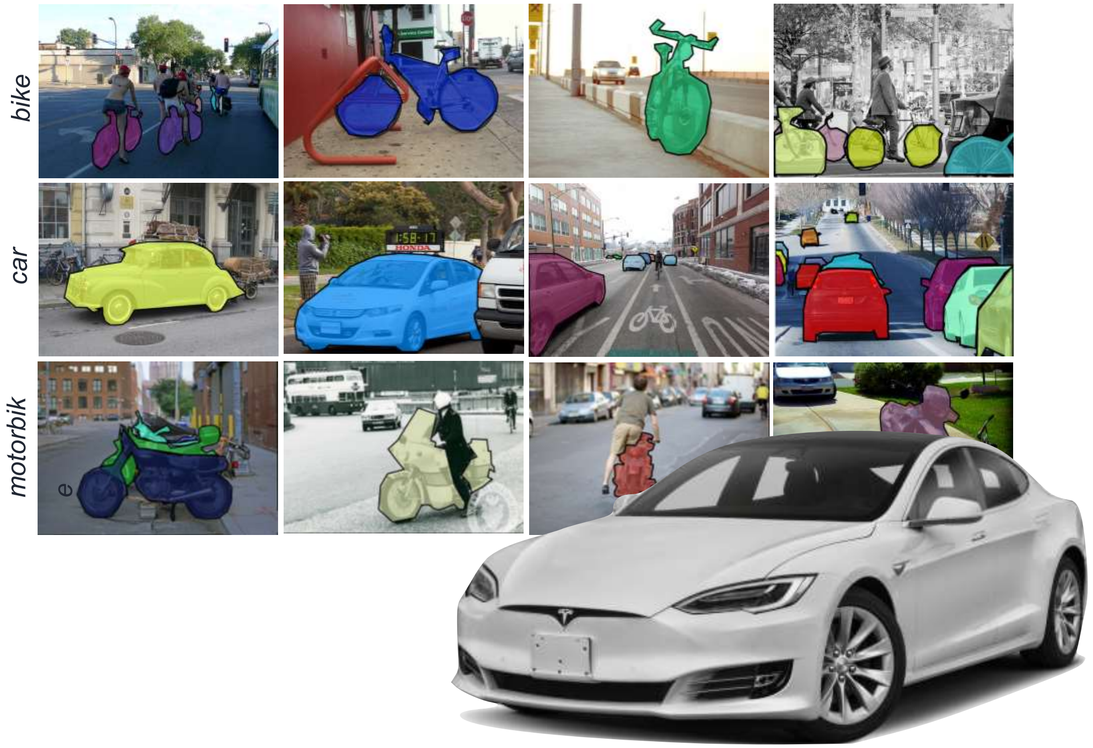
Candidate:

Fontanel Dario

Preliminaries

Closed world assumption

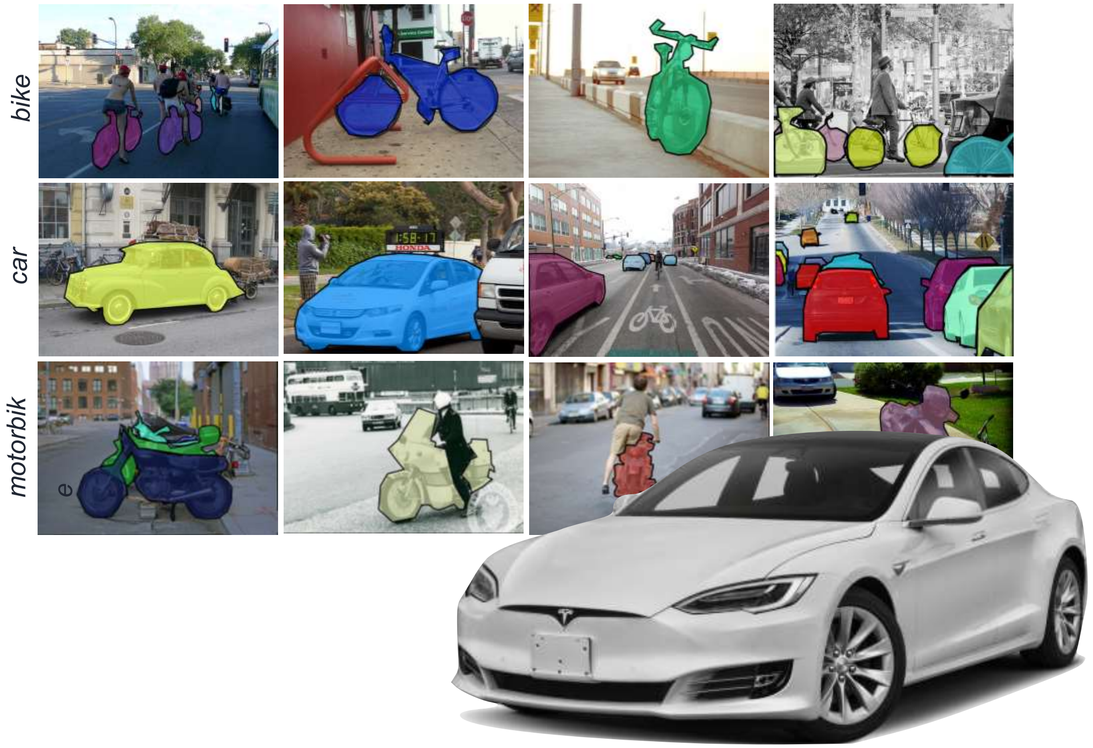
Training



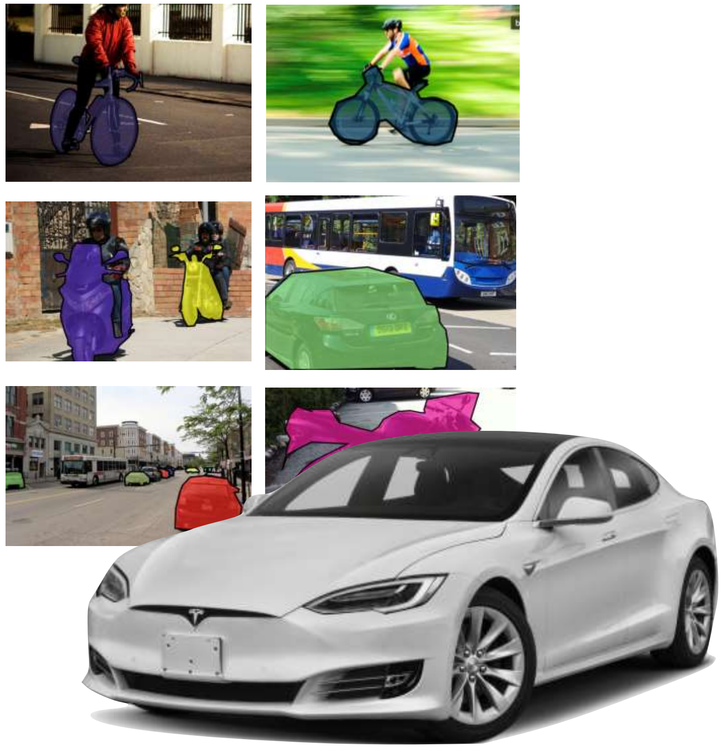
Preliminaries

Closed world assumption

Training



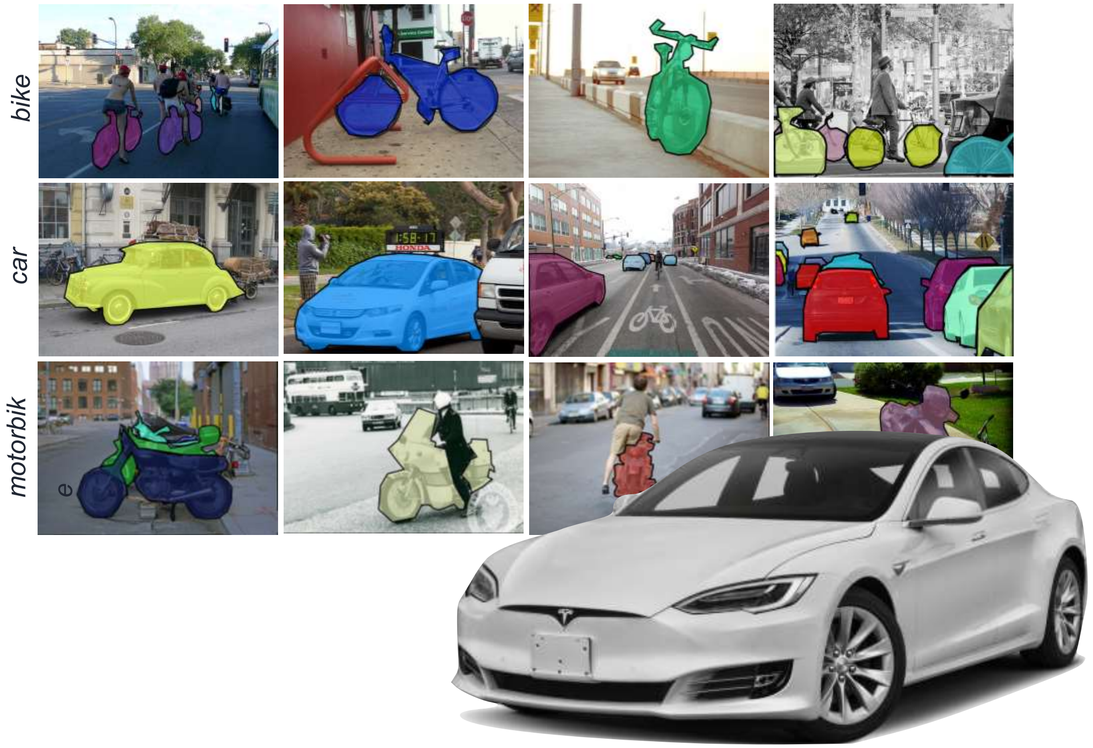
Deployment



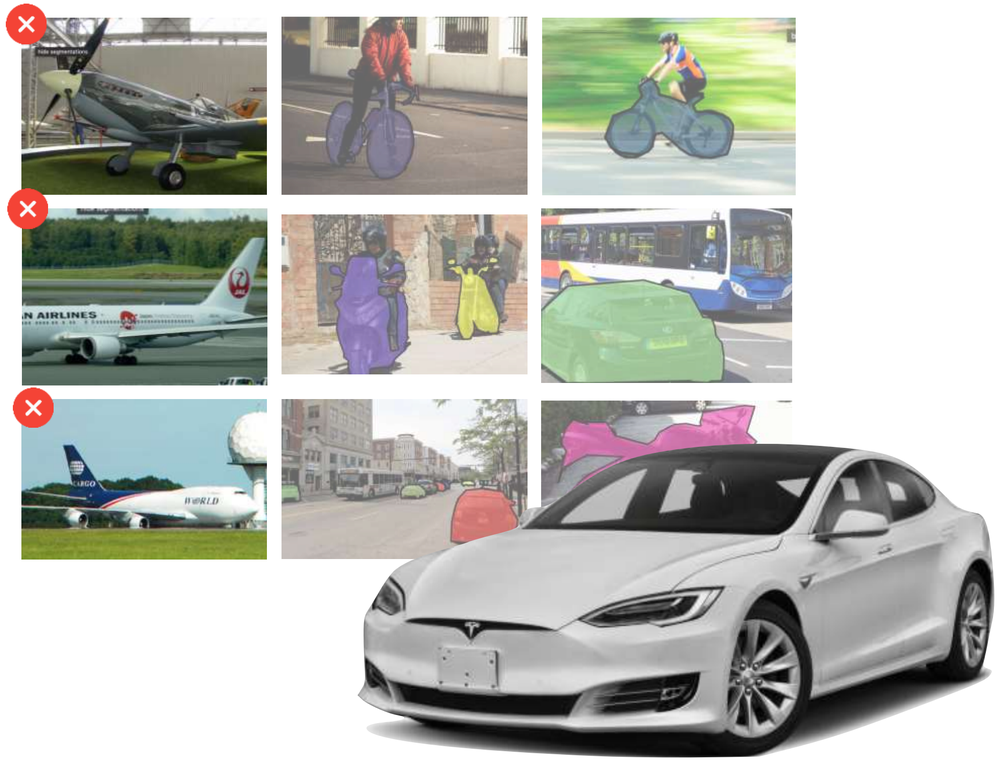
Preliminaries

Closed world assumption

Training



Deployment



Preliminaries

Open world

Train

Time

Knowledge (t)

Preliminaries

Open world

Train

Test

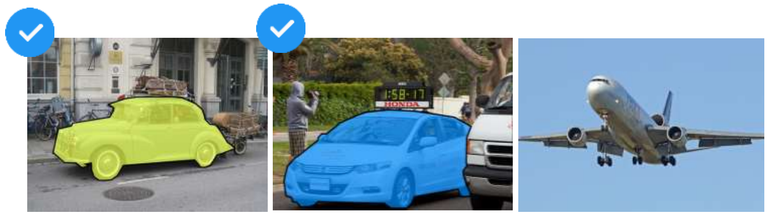
Time

Knowledge (t)

Preliminaries

Open world

Train



Test

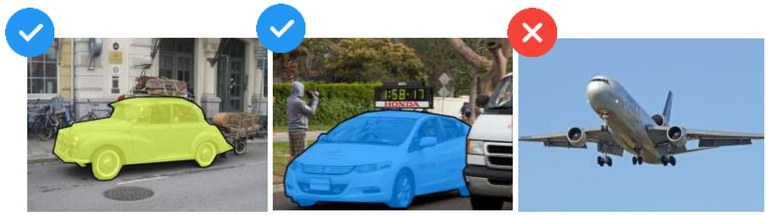
Time

Knowledge (t)

Preliminaries

Open world

Train



Test

Time

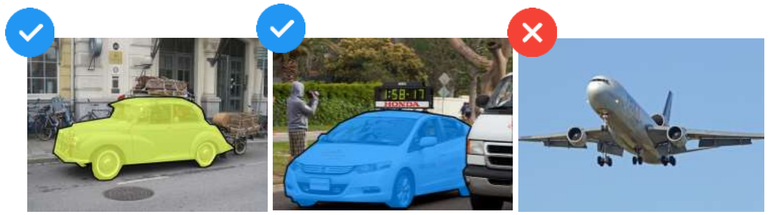
Knowledge (t)

Unknown category found

Preliminaries

Open world

Train



Test

Train

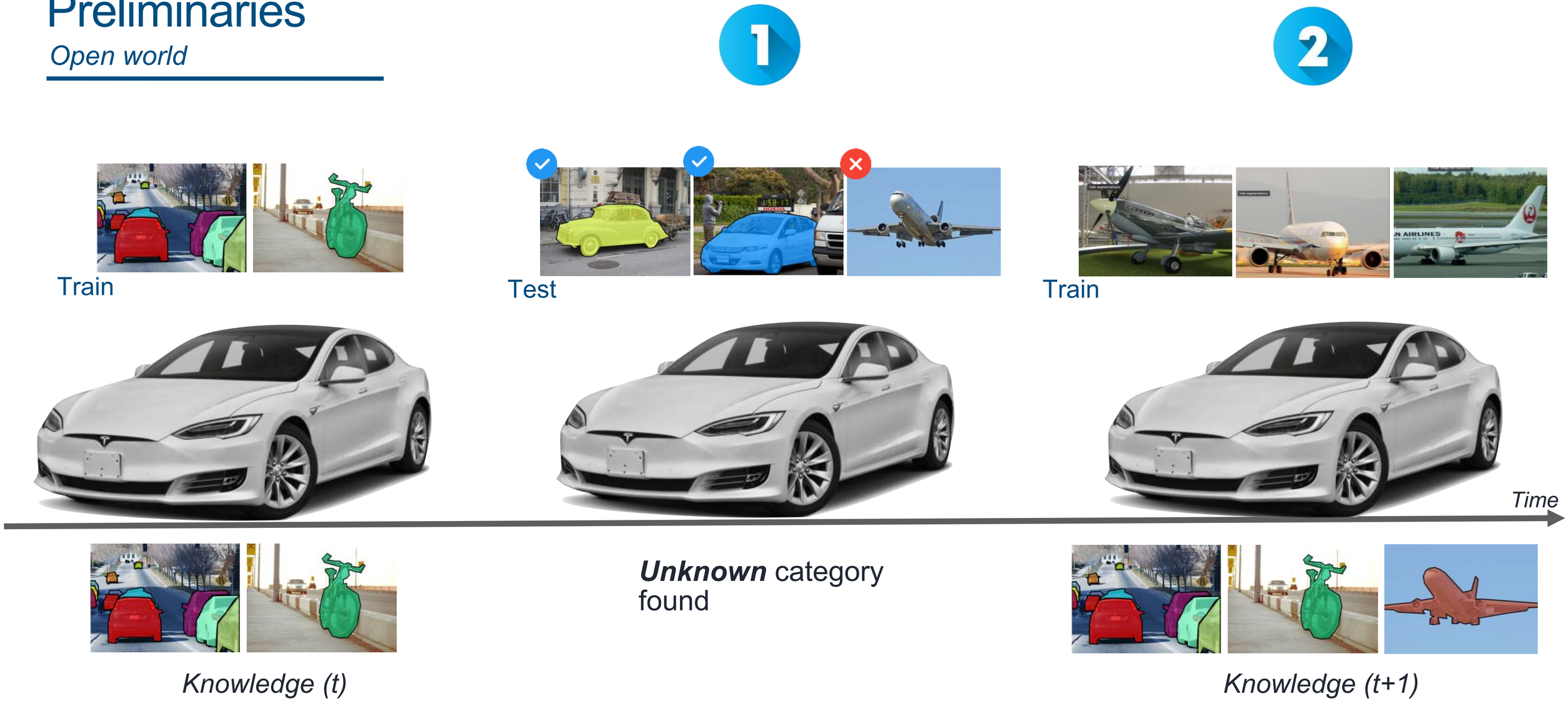
Time

Knowledge (t)

Unknown category found

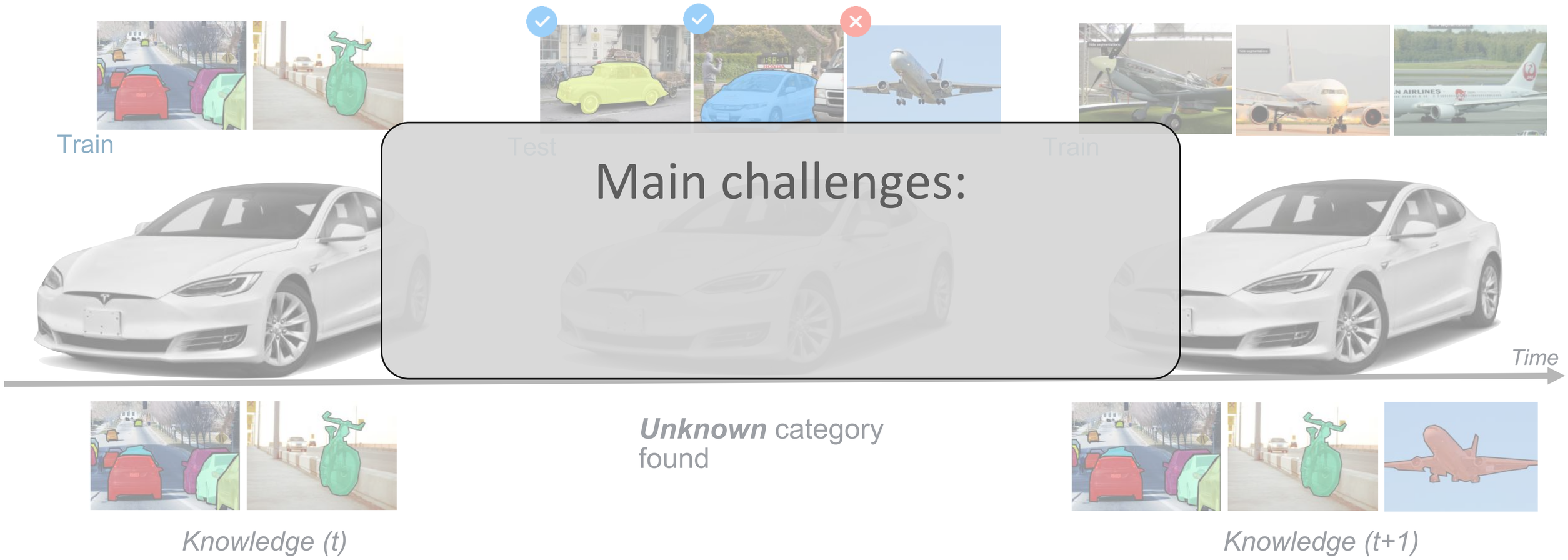
Preliminaries

Open world



Preliminaries

Open world

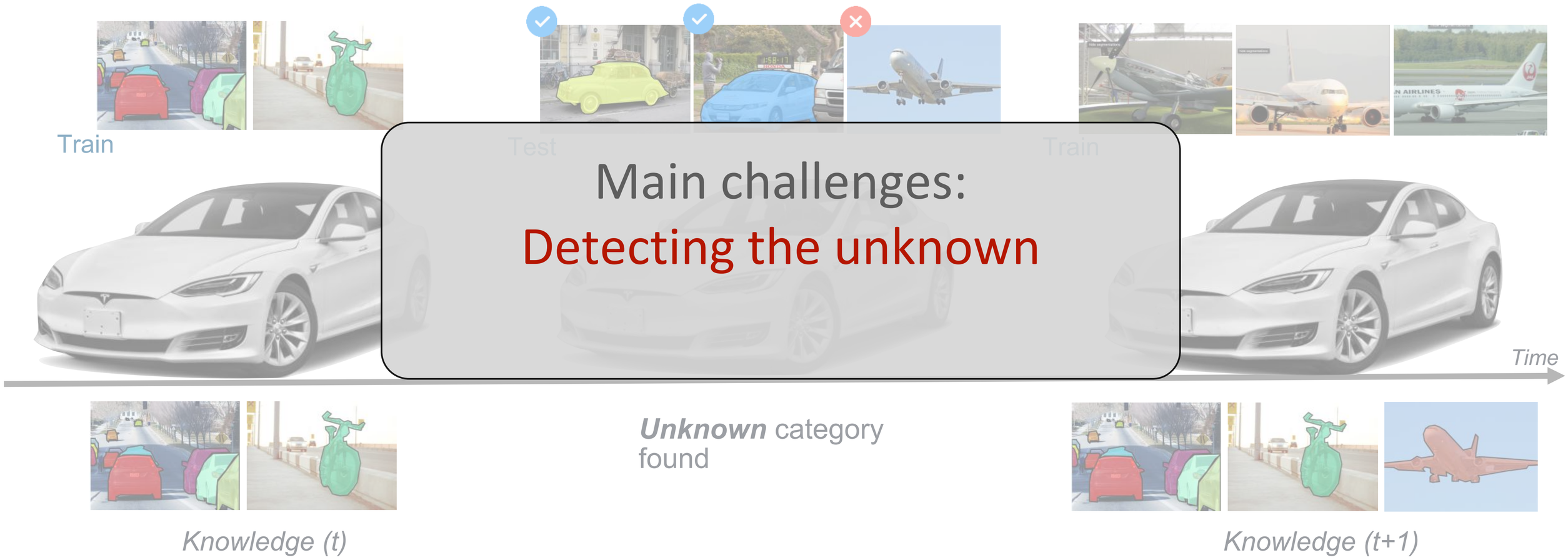


Main challenges:

Unknown category found

Preliminaries

Open world



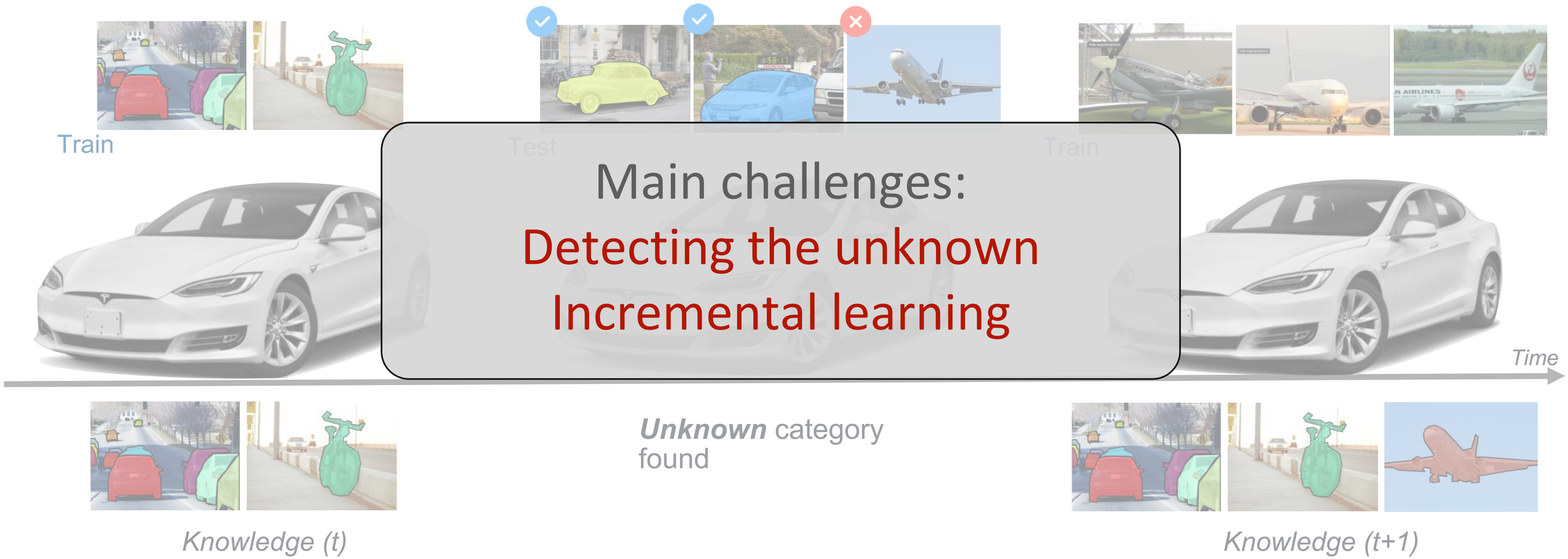
Main challenges:

Detecting the unknown

Unknown category found

Preliminaries

Open world



Main challenges:

Detecting the unknown

Incremental learning

Outline

Recognizing New Visual Categories in Real Open World Environments

Detection of the unknown

"PAnS: Prototopycal Anomaly Segmentation".

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Barbara Caputo

CVPR 21

Incremental learning

"Weakly supervised Incremental Learning for Semantic Segmentation"

Fabio Cermelli, Dario Fontanel, Antonio Tavera, Marco Ciccone, Barbara Caputo

CVPR 22

Unified framework

"Boosting Deep Open World Recognition by Clustering"

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Samuel Rota Bulo', Elisa Ricci, Barbara Caputo

IROS 20

Impact of domain shift

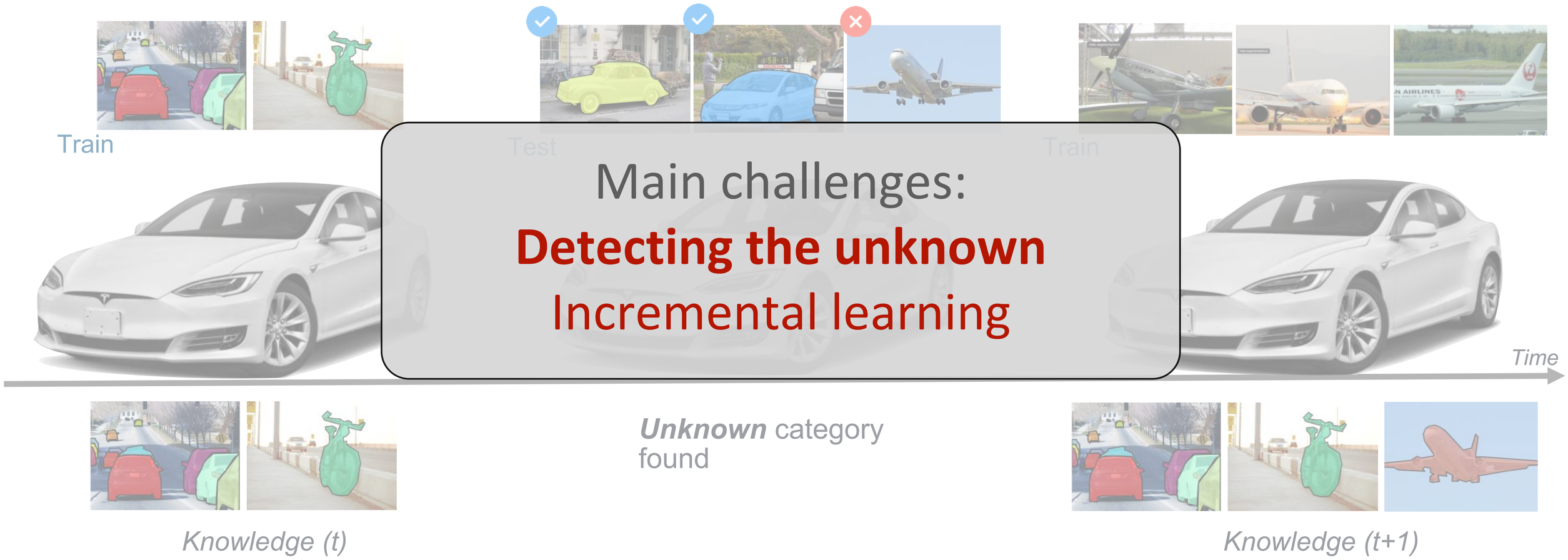
"On the Challenges of Open World Recognition under Shifting Visual Domains"

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Barbara Caputo

ICRA 21

Preliminaries

Open world



Main challenges:

Detecting the unknown

Incremental learning

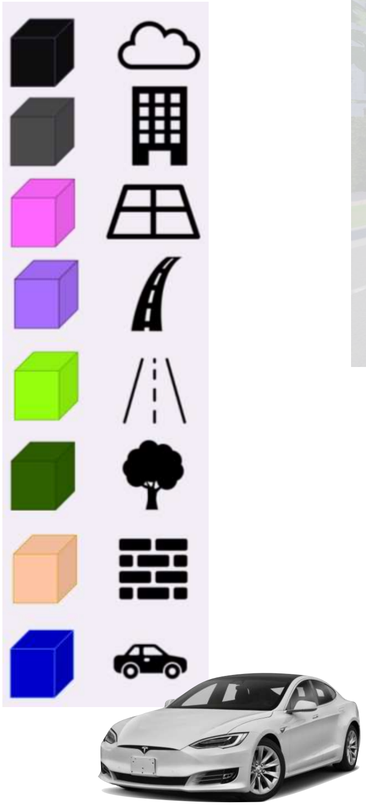
“PAnS: Prototopycal Anomaly Segmentation”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Barbara Caputo

IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2021, Nashville, USA

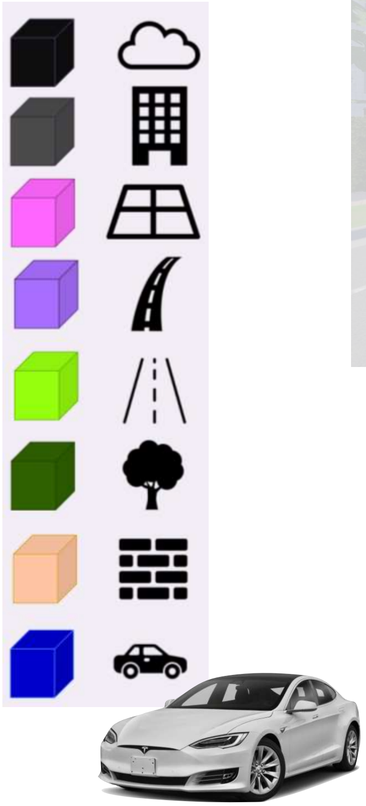
Anomaly Segmentation

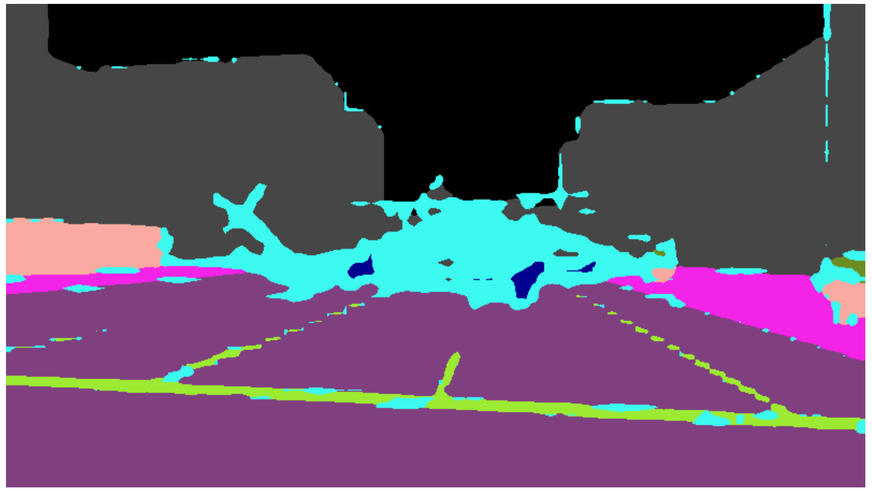
It refers to the task of identifying the pixels in an image that belong to unknown categories.



Anomaly Segmentation

It refers to the task of identifying the pixels in an image that belong to unknown categories.

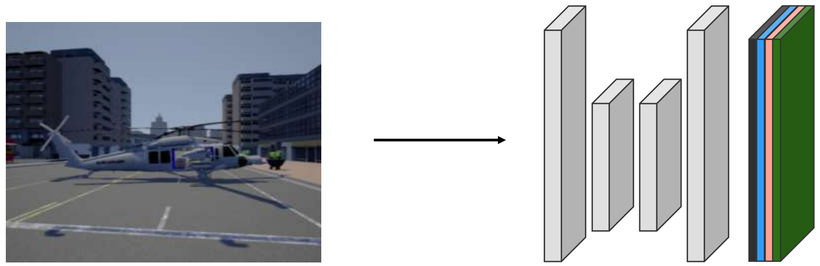




“PAnS: Prototopycal Anomaly Segmentation”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Barbara Caputo

Traditional approaches



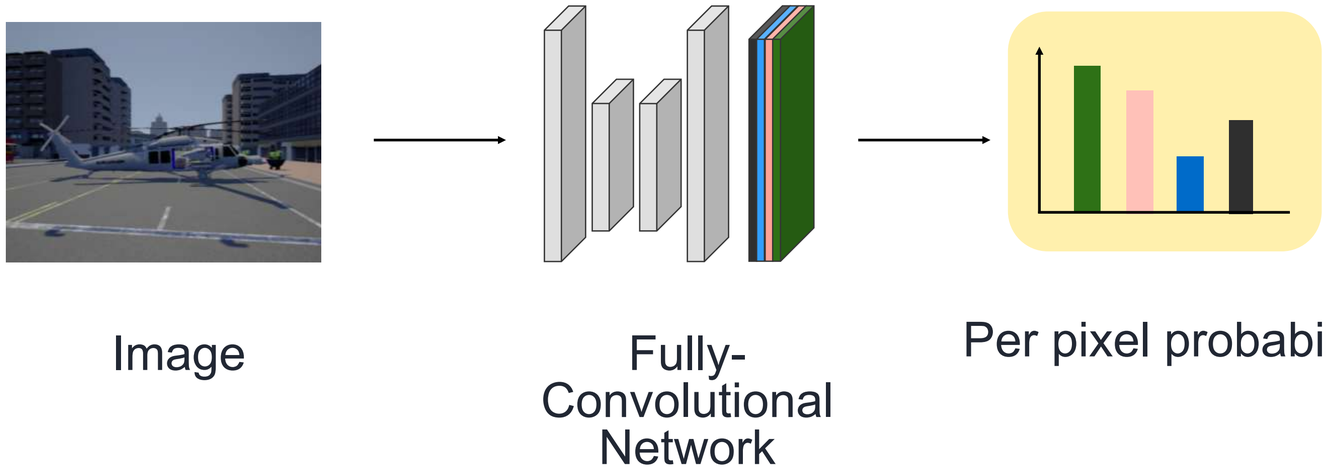
Image

Fully-Convolutional Network

“PAnS: Prototopycal Anomaly Segmentation”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Barbara Caputo

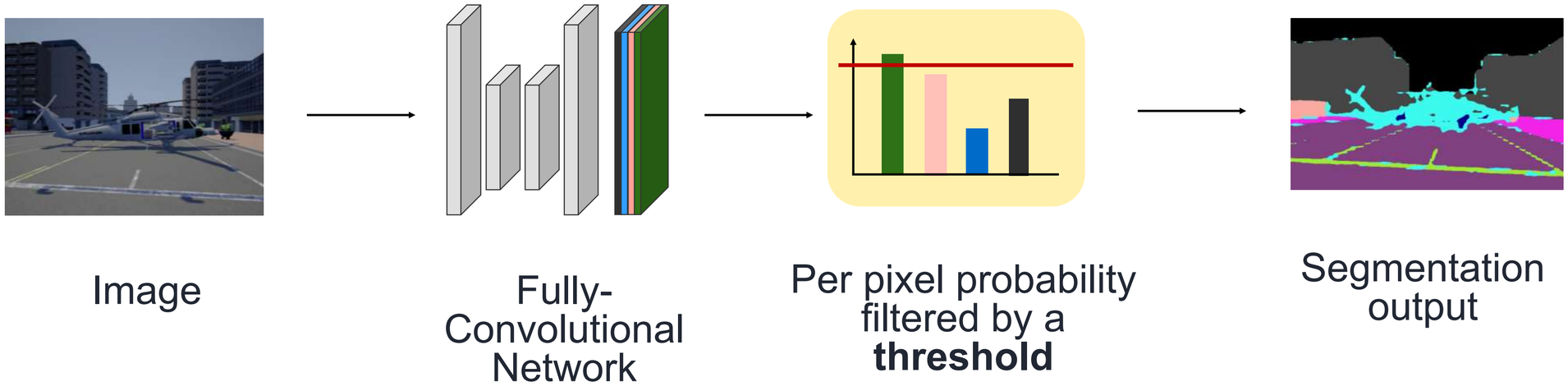
Traditional approaches



“PAnS: Prototopycal Anomaly Segmentation”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Barbara Caputo

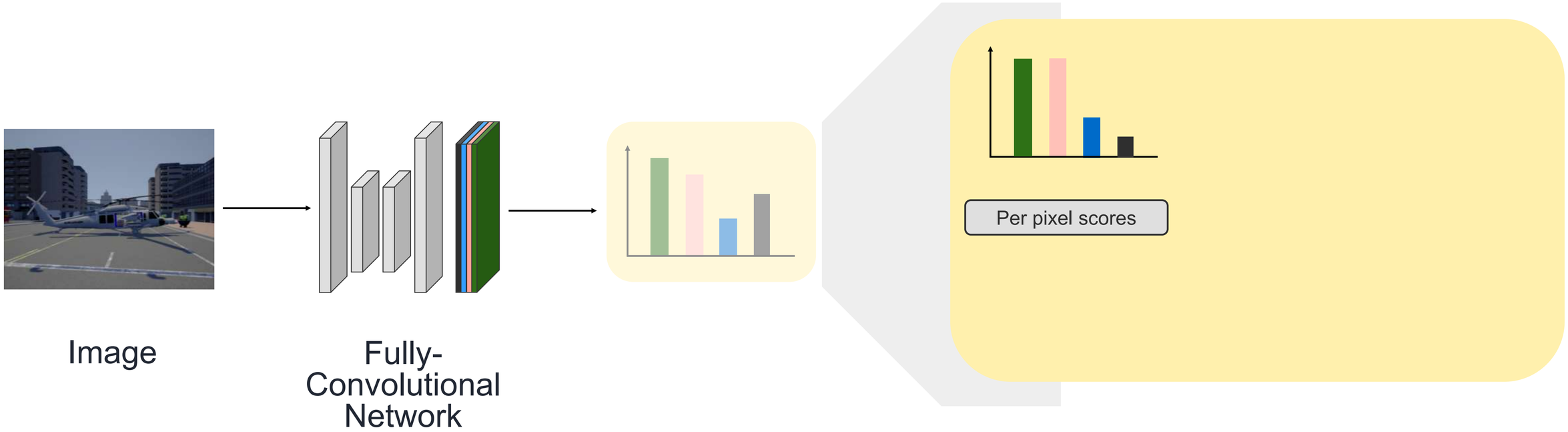
Traditional approaches



“PAnS: Prototopycal Anomaly Segmentation”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Barbara Caputo

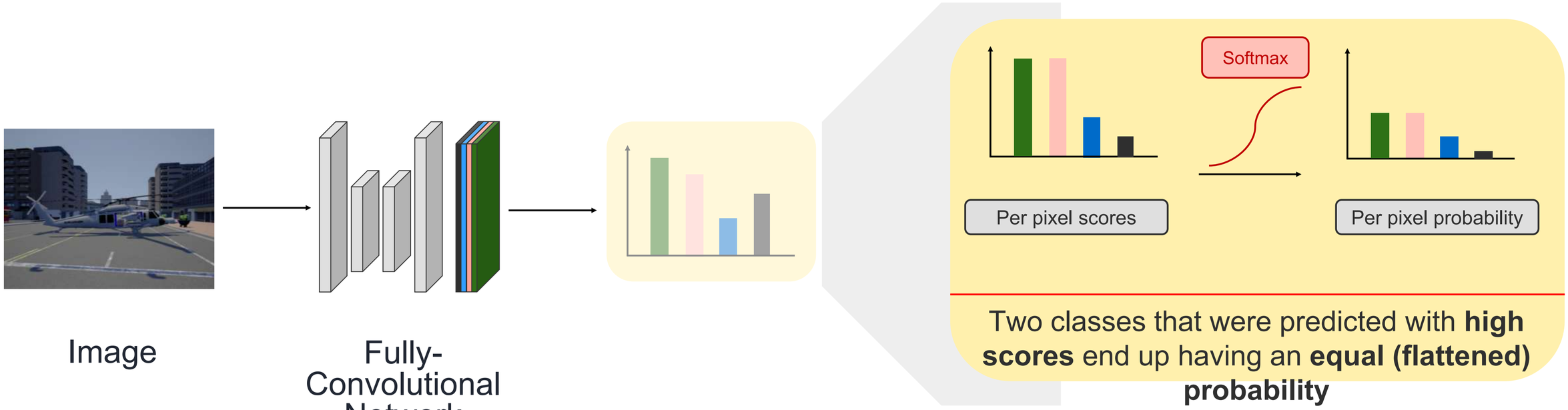
Traditional approaches



“PAnS: Prototopycal Anomaly Segmentation”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Barbara Caputo

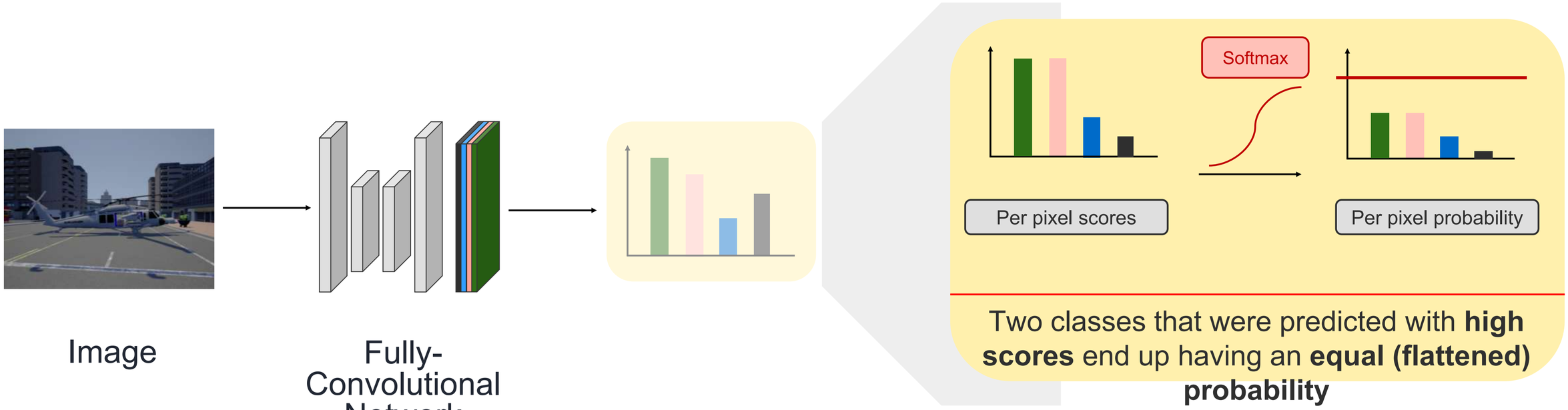
Traditional approaches



“PAnS: Prototopycal Anomaly Segmentation”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Barbara Caputo

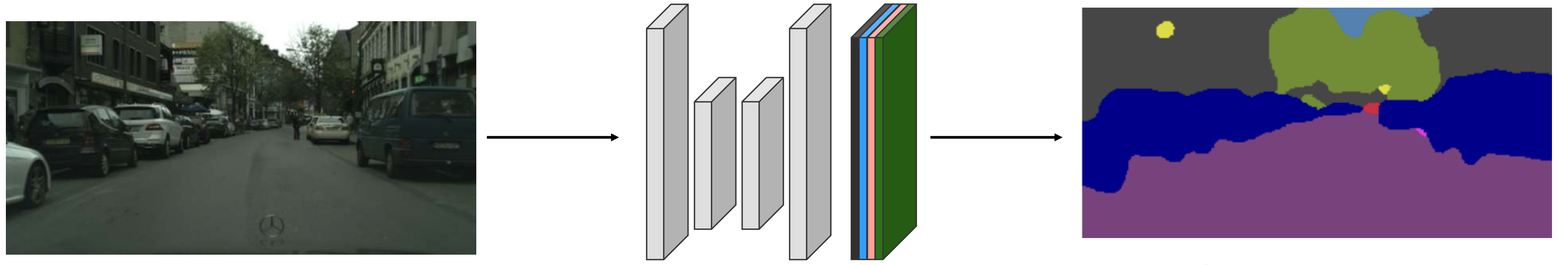
Traditional approaches



“PAnS: Prototopycal Anomaly Segmentation”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Barbara Caputo

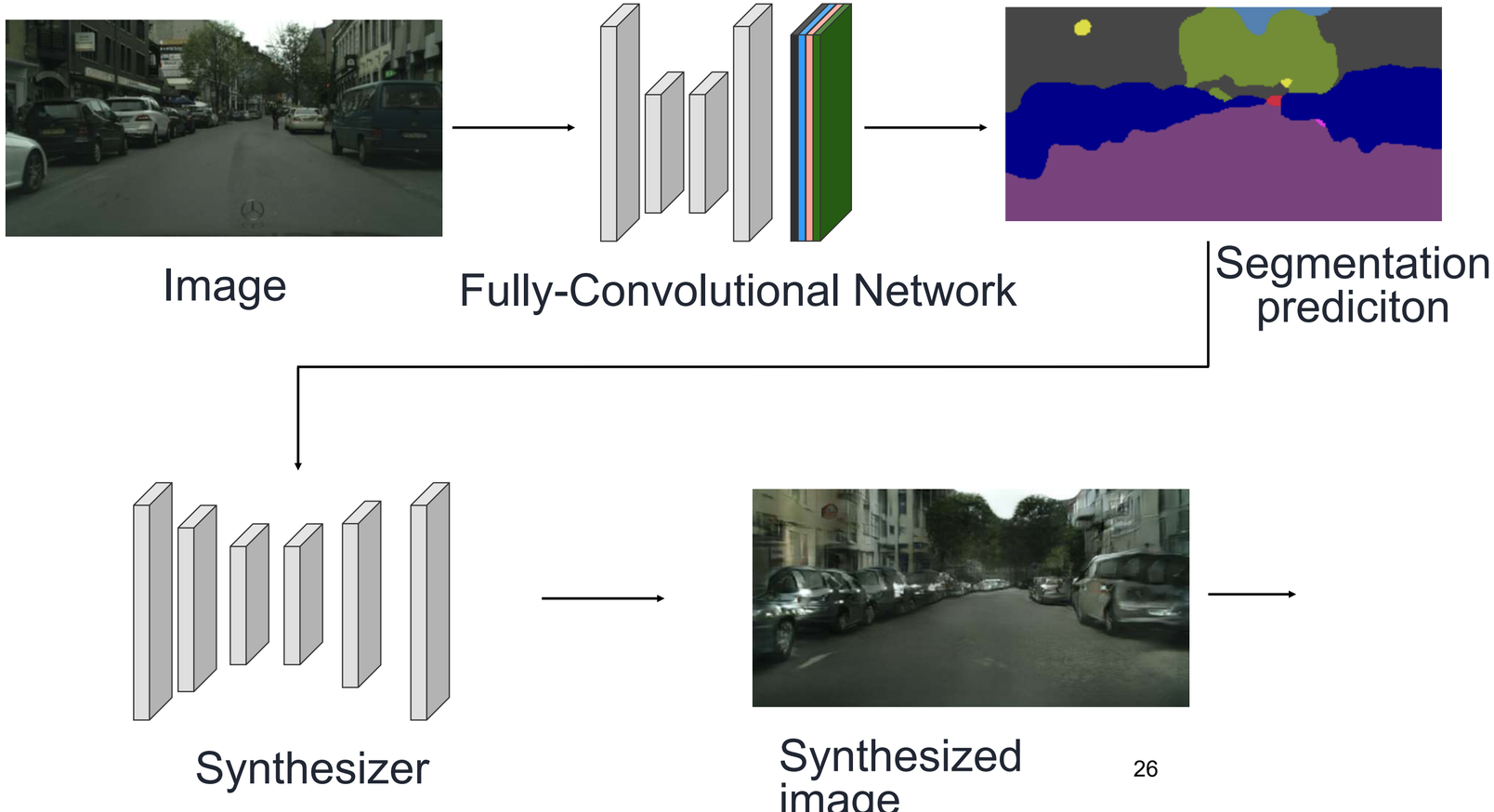
Traditional approaches



“PAnS: Prototopycal Anomaly Segmentation”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Barbara Caputo

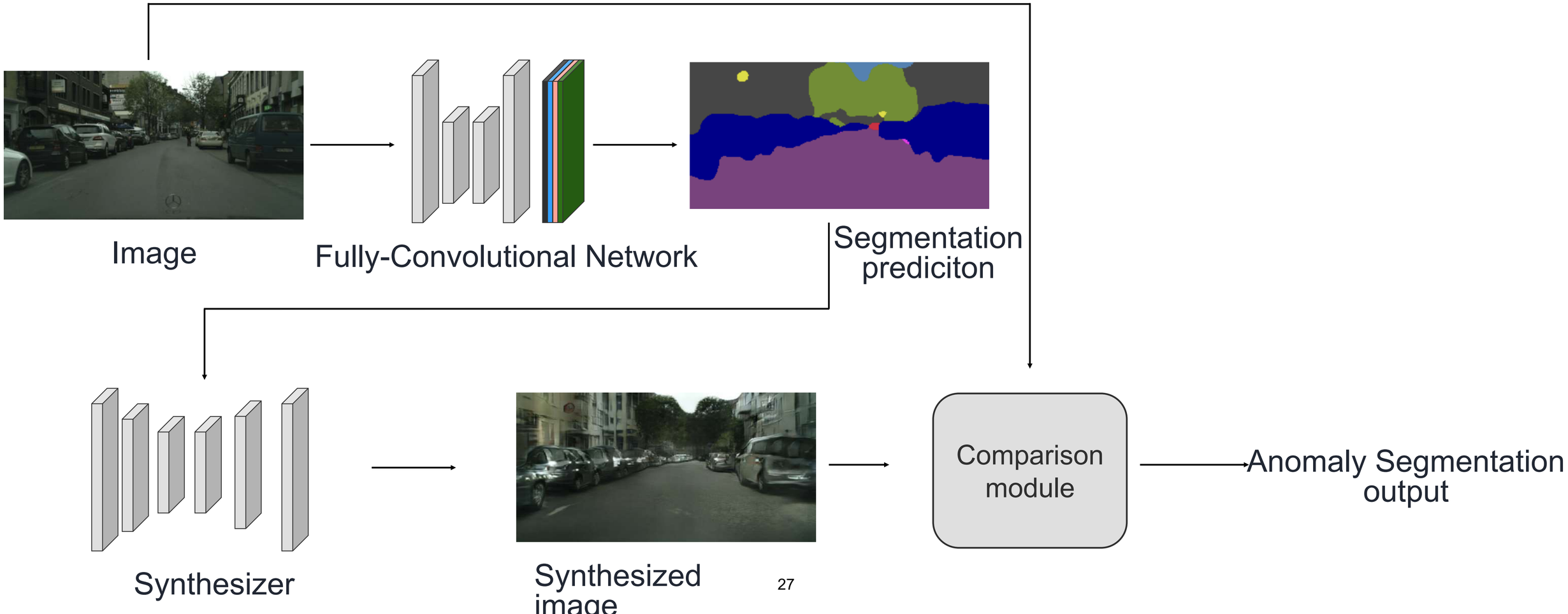
Traditional approaches



“PAnS: Prototopycal Anomaly Segmentation”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Barbara Caputo

Traditional approaches



“PAnS: Prototopycal Anomaly Segmentation”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Barbara Caputo

Traditional approaches

Due to semantic segmentation scenes complexity

Generative approaches tend to produce artifacts while synthesizing pixels of known classes

Reconstruction (top) of SPADE [1] on an image from StreetHazards dataset (middle) [2].

“PAnS: Prototopycal Anomaly Segmentation”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Barbara Caputo

Framework

Confidence values independent for each class

“PAnS: Prototopycal Anomaly Segmentation”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Barbara Caputo

Framework

Confidence values independent for each class

No additional computation required

“PAnS: Prototopycal Anomaly Segmentation”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Barbara Caputo

Framework

Confidence values independent for each class

No additional computation required

Scores bounded so that a threshold can be applied

“PAnS: Prototopycal Anomaly Segmentation”

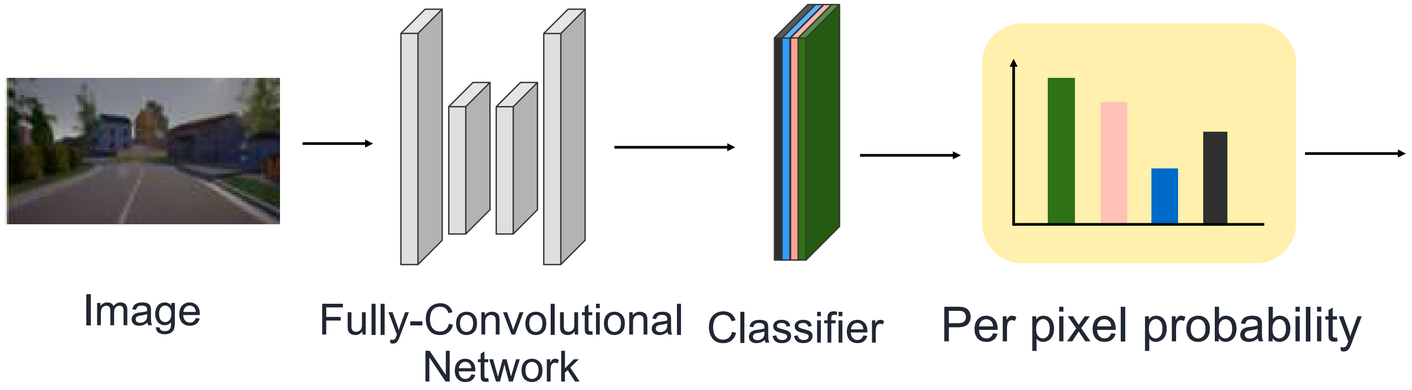
Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Barbara Caputo

Framework

Confidence values independent for each class

No additional computation required

Scores bounded so that a threshold can be applied



“PAnS: Prototopycal Anomaly Segmentation”

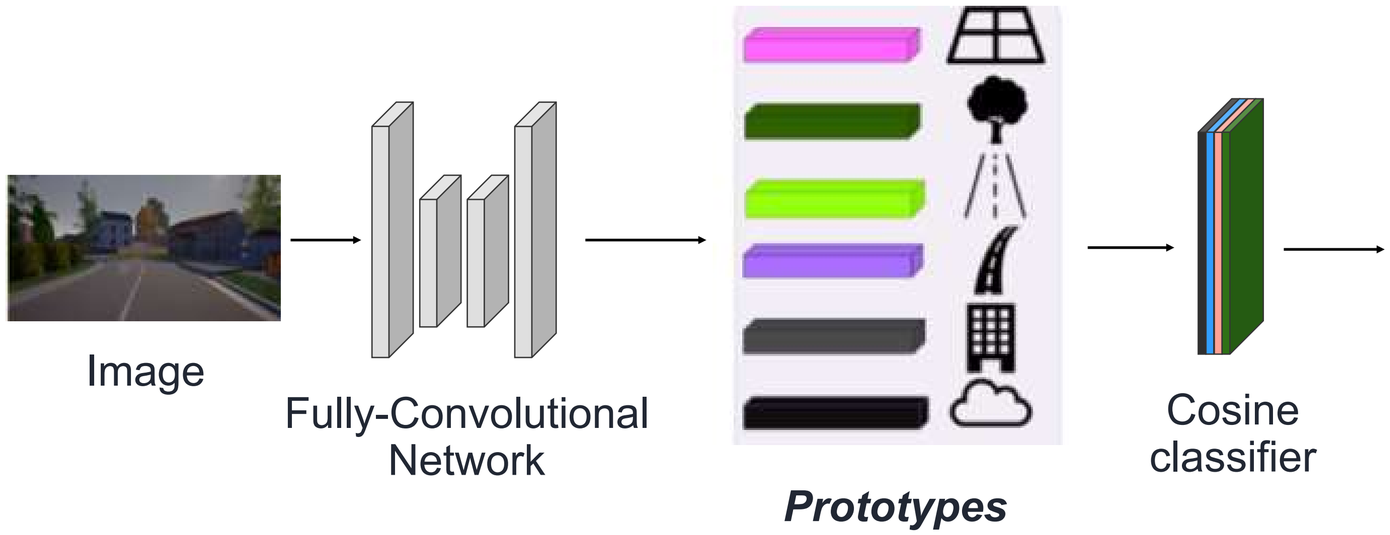
Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Barbara Caputo

Framework

Confidence values independent for each class

No additional computation required

Scores bounded so that a threshold can be applied

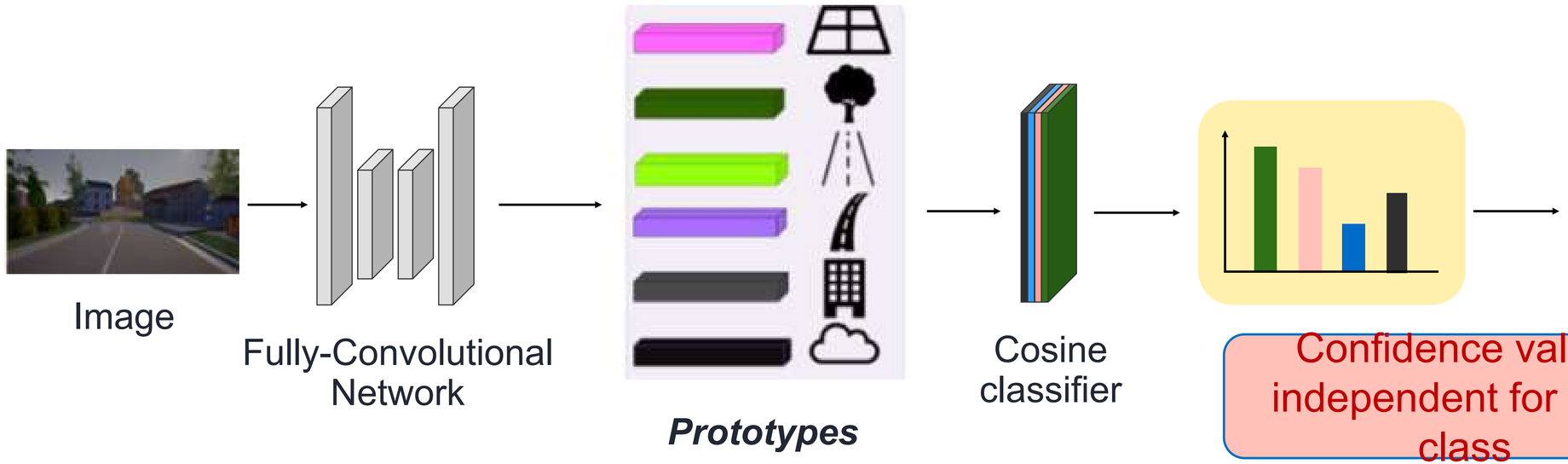


“PAnS: Prototopycal Anomaly Segmentation”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Barbara Caputo

Framework

Scores bounded so that a threshold can be applied



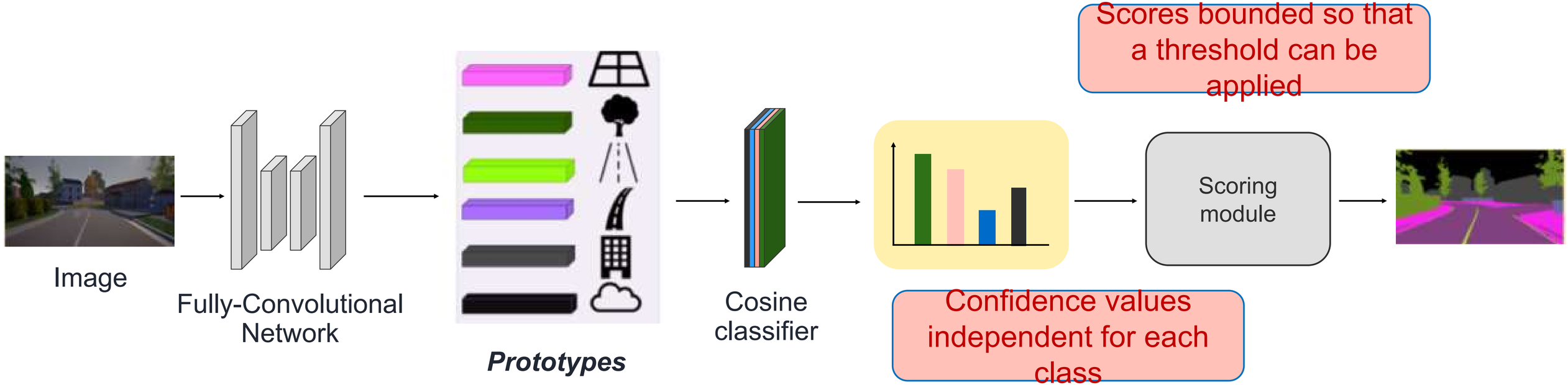
Confidence values independent for each class

No additional computation required

“PAnS: Prototopycal Anomaly Segmentation”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Barbara Caputo

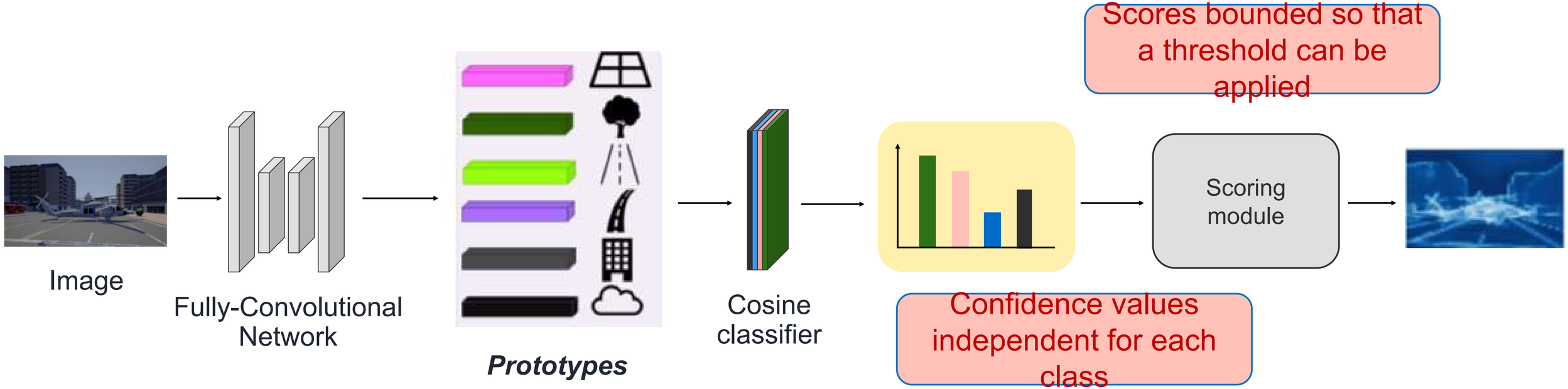
Framework



“PAnS: Prototopycal Anomaly Segmentation”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Barbara Caputo

Framework



“PAnS: Prototopycal Anomaly Segmentation”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Barbara Caputo

Experimental setting

AUPR

area under the Precision-Recall curve

AUROC

area under the True Positive Rate and False Positive Rate curve

FPR95

measures the False Positive Rate (FPR) at 95% of recall

Architecture

ResNet-50 architecture [67] as backbone and PSPNet [26] as head module.

“PAnS: Prototopycal Anomaly Segmentation”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Barbara Caputo

Results

StreetHazards

D.Hendrycs, et al. *A benchmark for Anomaly Segmentation (2019)*

“PAnS: Prototopycal Anomaly Segmentation”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Barbara Caputo

Results

StreetHazards

D.Hendrycs, et al. *A benchmark for Anomaly Segmentation (2019)*

Method	*AUPR* ↑	*AUROC* ↑	*FPR95* ↓
AE	2.2	66.1	91.7
Dropout	7.5	69.9	79.4
MSP	6.6	87.7	33.7
MSP + CRF	6.5	88.1	29.9
SynthCP	**9.3**	88.5	28.4
PAnS	8.8	**91.1**	**23.2**

“PAnS: Prototopycal Anomaly Segmentation”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Barbara Caputo

Results

StreetHazards

D.Hendrycs, et al. *A benchmark for Anomaly Segmentation (2019)*

Method	*AUPR* ↑	*AUROC* ↑	*FPR95* ↓
AE	2.2	66.1	91.7
Dropout	7.5	69.9	79.4
MSP	6.6	87.7	33.7
MSP + CRF	6.5	88.1	29.9
SynthCP	**9.3**	88.5	28.4
PAnS	8.8	**91.1**	**23.2**

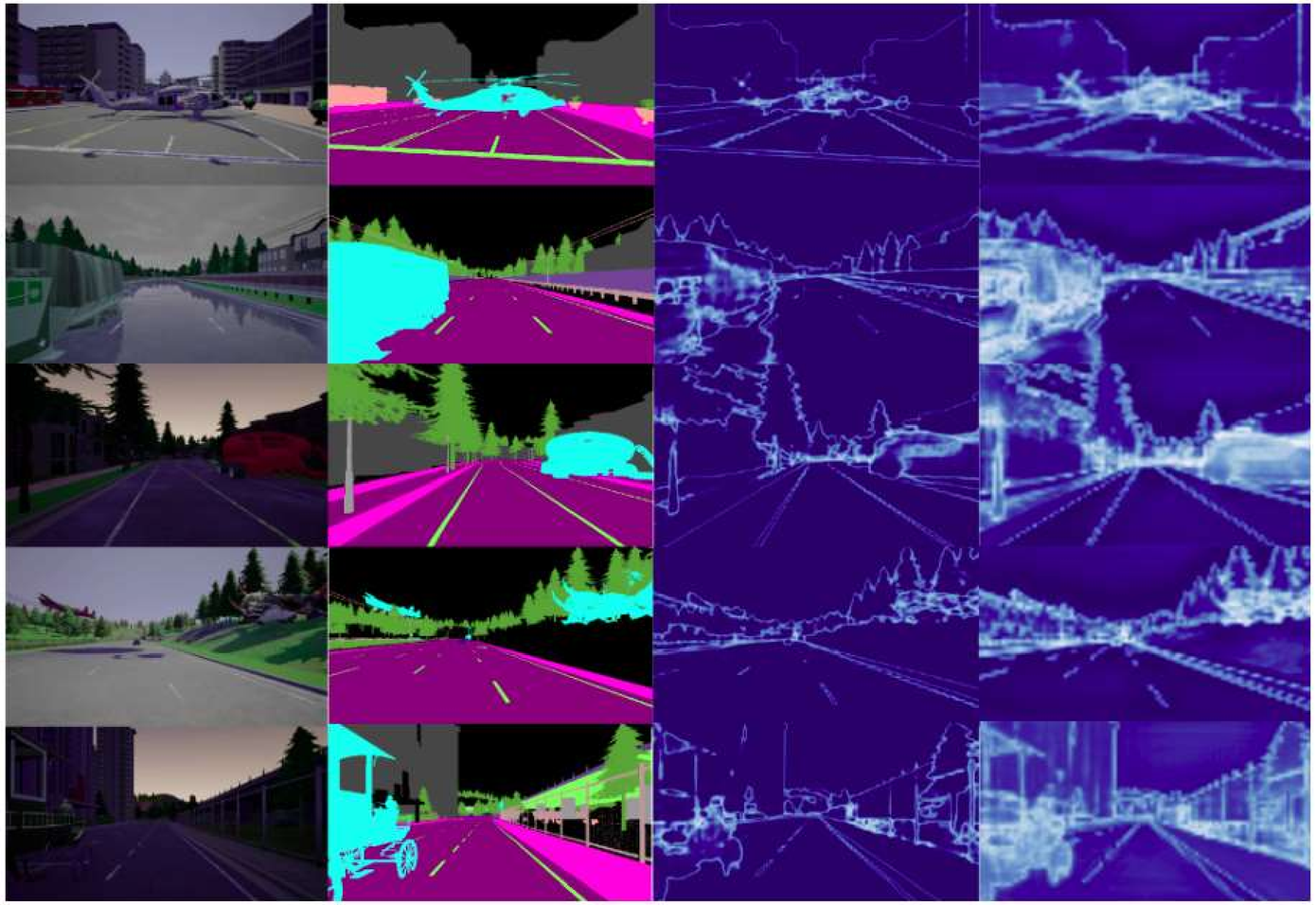
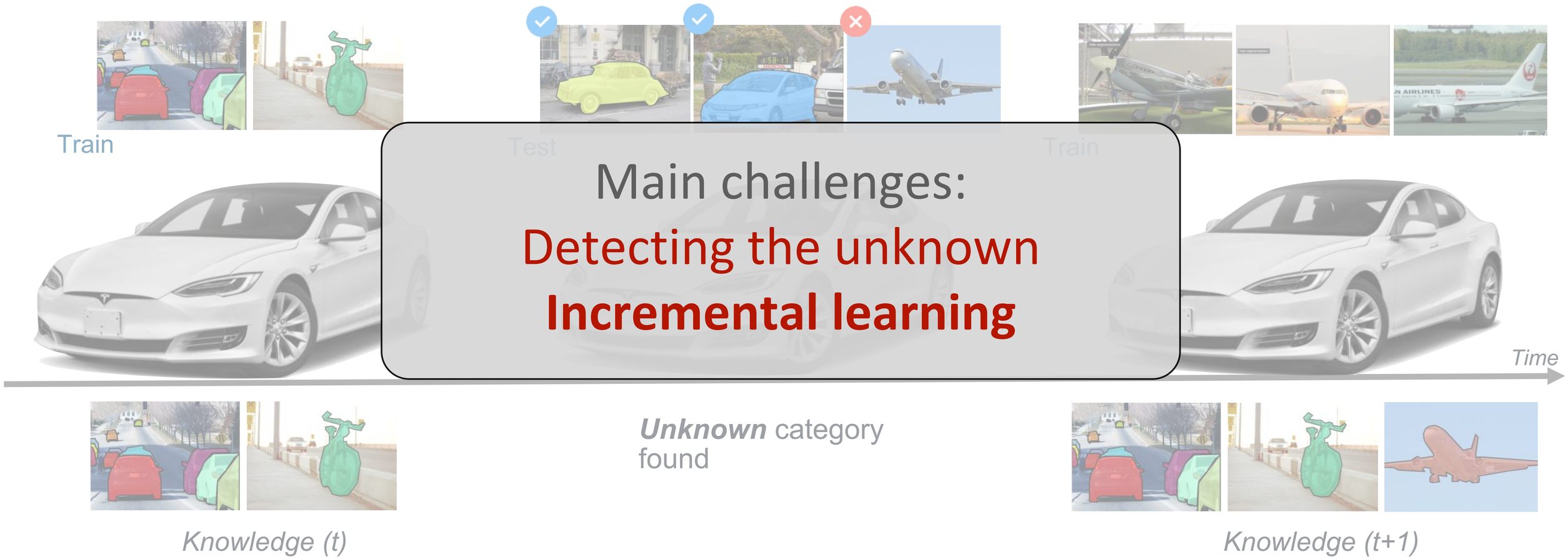


Image Semantic labels MSP PAnS

Preliminaries

Open world



Main challenges:

Detecting the unknown

Incremental learning

Unknown category found

“Weakly supervised Incremental Learning for Semantic Segmentation”

Fabio Cermelli, Dario Fontanel, Antonio Tavera, Marco Ciccone, Barbara Caputo

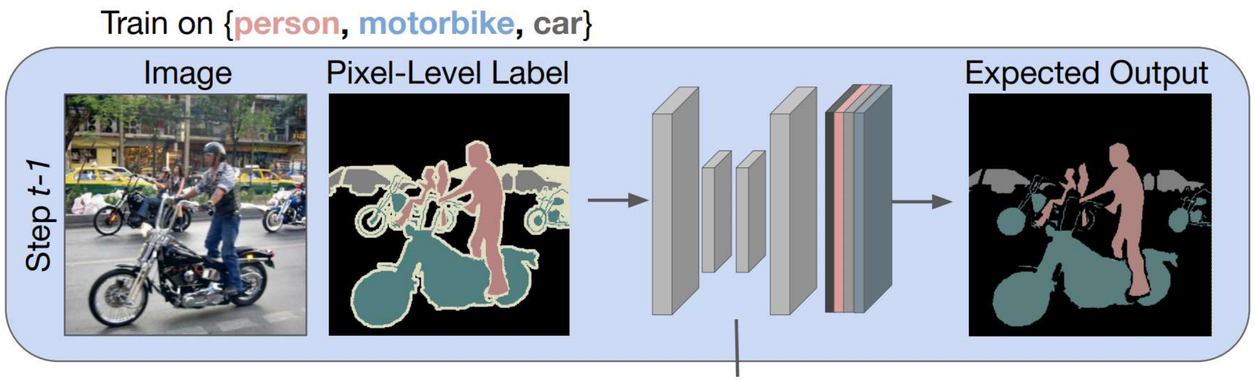
IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2022, New Orleans, USA

Weakly supervised Incremental Learning for Semantic Segmentation

Fabio Cermelli, Dario Fontanel, Antonio Tavera, Marco Ciccone, Barbara Caputo

Task overview

Suppose we are given a model pre-trained on a set of classes, using expensive pixel-level annotations



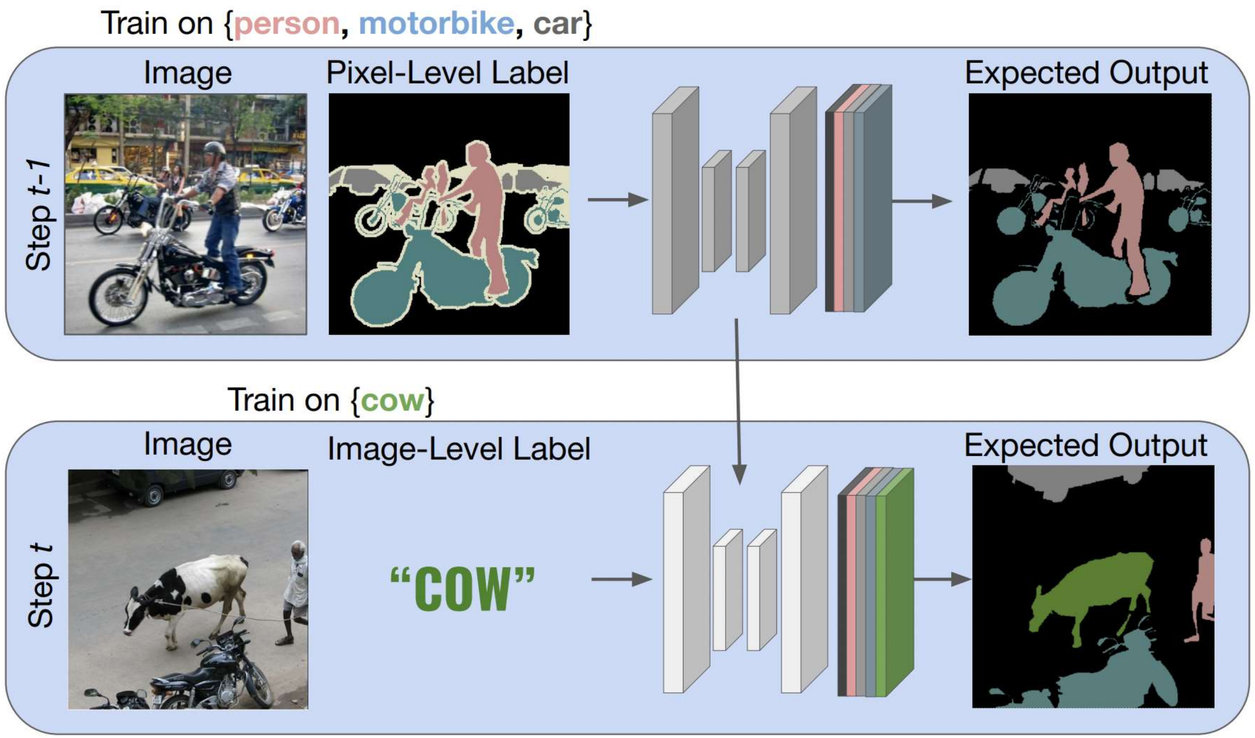
Weakly supervised Incremental Learning for Semantic Segmentation

Fabio Cermelli, Dario Fontanel, Antonio Tavera, Marco Ciccone, Barbara Caputo

Task overview

Suppose we are given a model pre-trained on a set of classes, using expensive pixel-level annotations

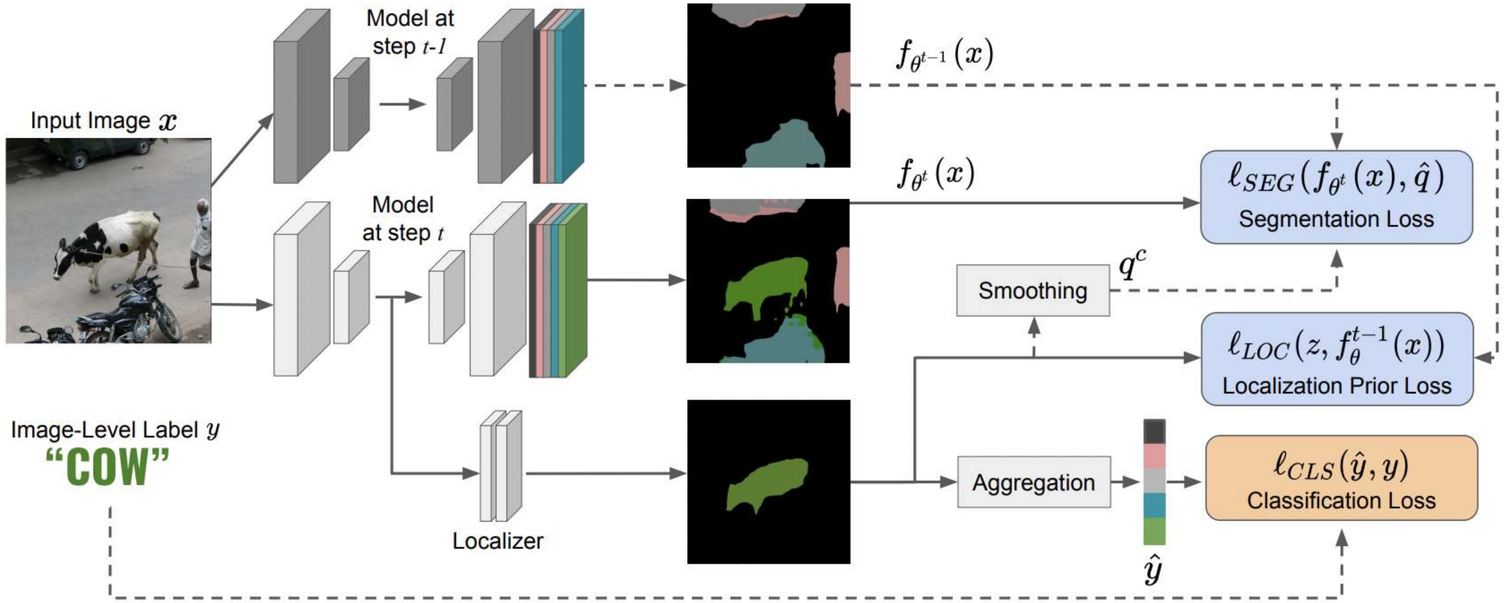
Our goal is to incrementally segment new classes over time using only image-level labels



Weakly supervised Incremental Learning for Semantic Segmentation

Fabio Cermelli, Dario Fontanel, Antonio Tavera, Marco Ciccone, Barbara Caputo

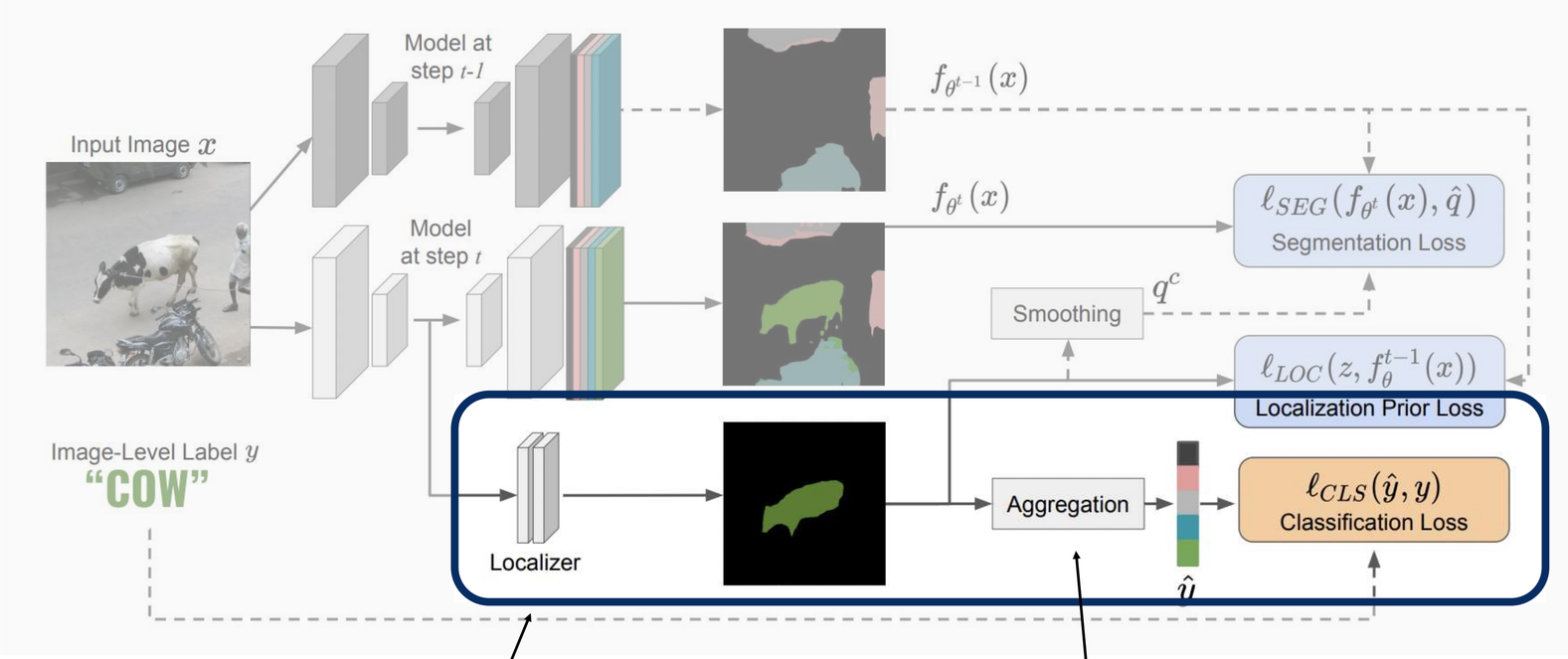
Method



Weakly supervised Incremental Learning for Semantic Segmentation

Fabio Cermelli, Dario Fontanel, Antonio Tavera, Marco Ciccone, Barbara Caputo

Method



Stack of convolutional layers

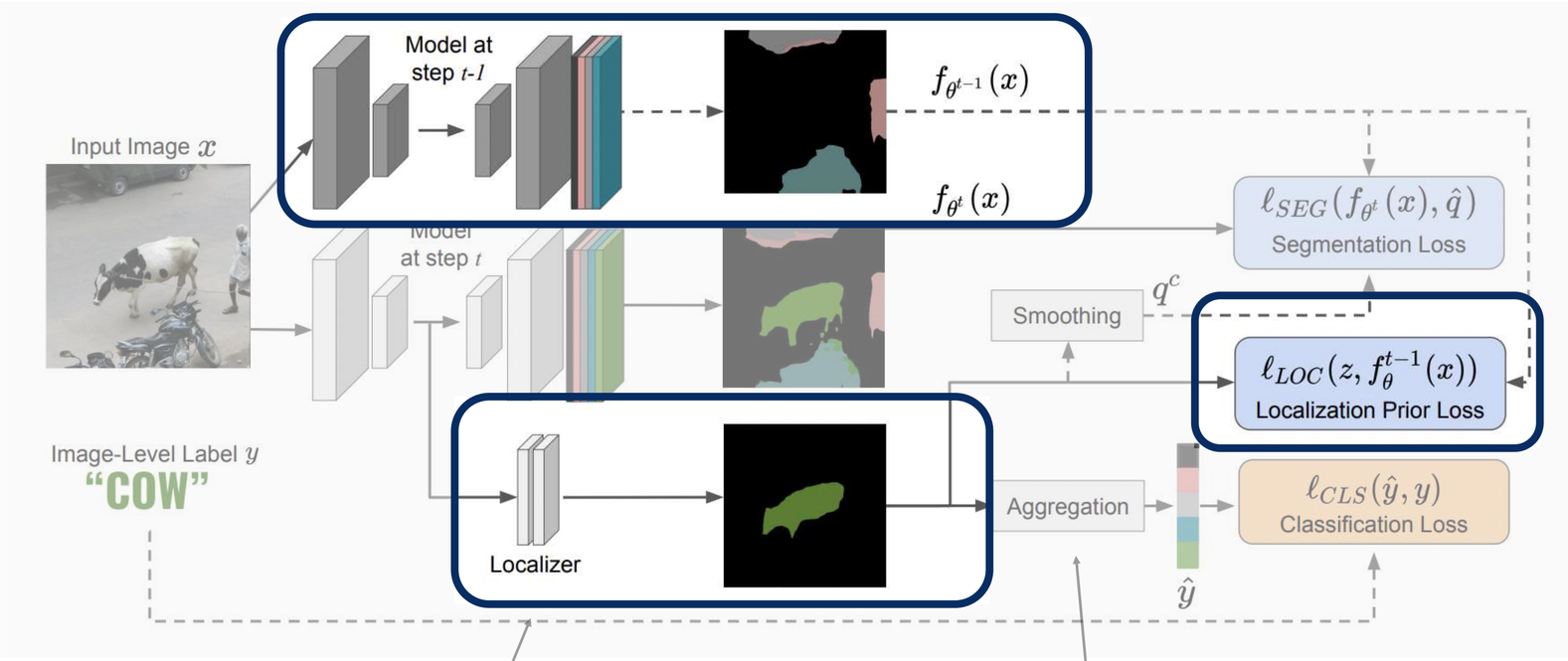
Spatial aggregation module transforms the output volume into a 1D tensor

Weakly supervised Incremental Learning for Semantic Segmentation

Fabio Cermelli, Dario Fontanel, Antonio Tavera, Marco Ciccone, Barbara Caputo

Method

The model at the previous time step prevents catastrofic forgetting and serves as a localization prior



Stack of convolutional layers

Spatial aggregation module transforms the output volume into a 1D tensor

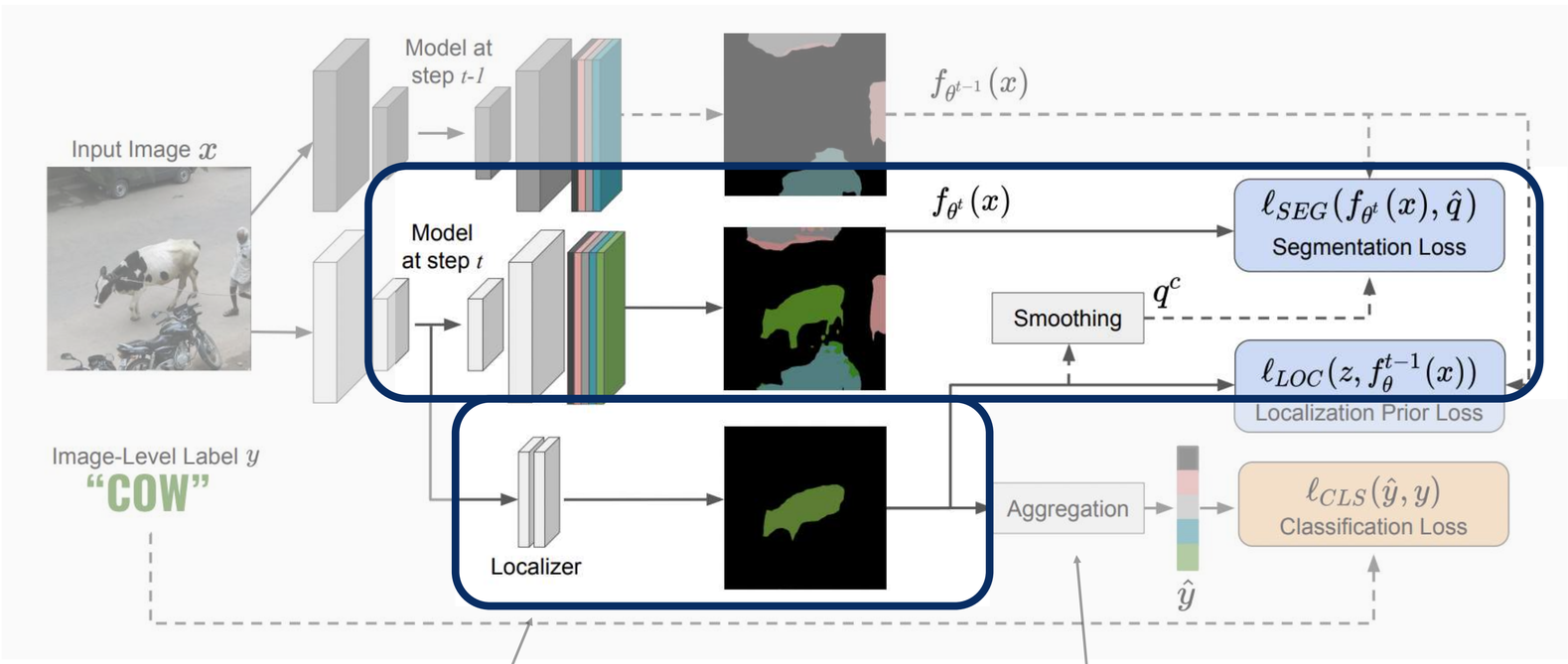
Weakly supervised Incremental Learning for Semantic Segmentation

Fabio Cermelli, Dario Fontanel, Antonio Tavera, Marco Ciccone, Barbara Caputo

Method

The model at the previous time step prevents catastrofic forgetting and serves as a localization prior

Traditional semantic segmentation pipeline



Stack of convolutional layers

Spatial aggregation module transforms the output volume into a 1D tensor

Weakly supervised Incremental Learning for Semantic Segmentation

Fabio Cermelli, Dario Fontanel, Antonio Tavera, Marco Ciccone, Barbara Caputo

Experimental setting

Comparisons

Fully supervised IL methods

- Standard fully supervised incremental learning approaches serving as upper bounds

Weakly supervised Incremental Learning for Semantic Segmentation

Fabio Cermelli, Dario Fontanel, Antonio Tavera, Marco Ciccone, Barbara Caputo

Experimental setting

Comparisons

Fully supervised IL methods

- Standard fully supervised incremental learning approaches serving as upper bounds

Weakly-supervised methods

- Used to generate pseudo-labels offline and train the same IL model

Weakly supervised Incremental Learning for Semantic Segmentation

Fabio Cermelli, Dario Fontanel, Antonio Tavera, Marco Ciccone, Barbara Caputo

Experimental setting

Comparisons

Fully supervised IL methods

Standard fully supervised incremental learning approaches serving as upper bounds

Weakly-supervised methods

Used to generate pseudo-labels offline and train the same IL model

Architecture

ResNet-101 architecture [67] as backbone and Deeplav v3 [26] as head module.

Weakly supervised Incremental Learning for Semantic Segmentation

Fabio Cermelli, Dario Fontanel, Antonio Tavera, Marco Ciccone, Barbara Caputo

Experimental setting

Comparisons

Fully supervised IL methods

Standard fully supervised incremental learning approaches serving as upper bounds

Weakly-supervised methods

Used to generate pseudo-labels offline and train same IL model
the

Architecture

ResNet-101 architecture [67] as backbone and Deeplav v3 head module.

[26] as

Pascal VOC 2012

20 semantic classes + background

10k images

15-5

15 classes first, 5 classes added

10-10

10 classes first, 10 classes added

The PASCAL Visual Object Classes (VOC) Challenge. Everingham, M., Van Gool, L., Williams, C. K. I., Winn, J. and Zisserman, A, IJCV10

Weakly supervised Incremental Learning for Semantic Segmentation

Fabio Cermelli, Dario Fontanel, Antonio Tavera, Marco Ciccone, Barbara Caputo

Experimental setting

Comparisons

Fully supervised IL methods

Standard fully supervised incremental learning approaches serving as upper bounds

Weakly-supervised methods

Used to generate pseudo-labels offline and train same IL model
the

Architecture

ResNet-101 architecture [67] as backbone and Deeplav v3 head module.

[26] as

Pascal VOC 2012

20 semantic classes + background
10k images

15-5

15 classes first,

5 classes added

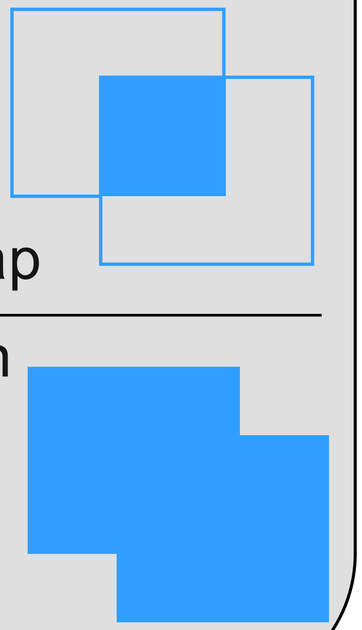
10-10

10 classes first,

10 classes added

mIoU

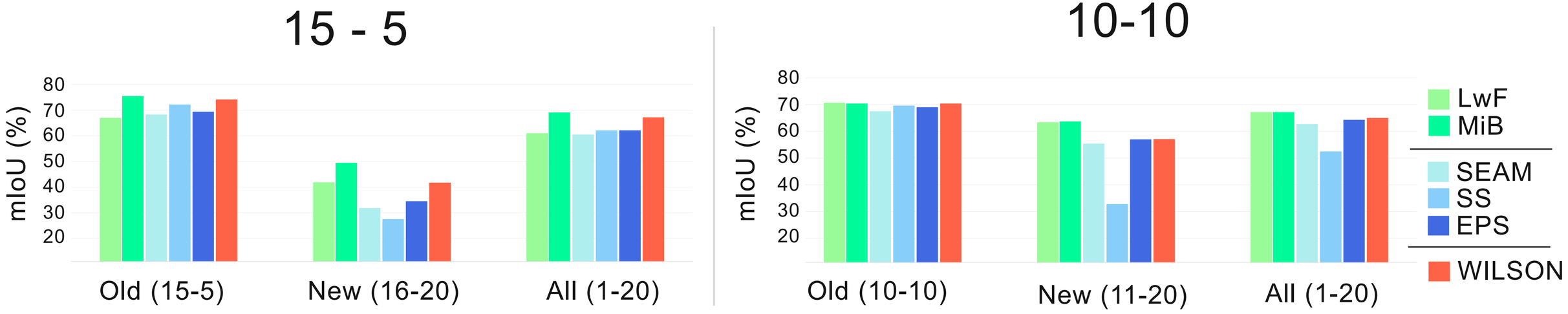
Intersection over Union (IoU) = $\frac{\text{area of overlap}}{\text{area of union}}$

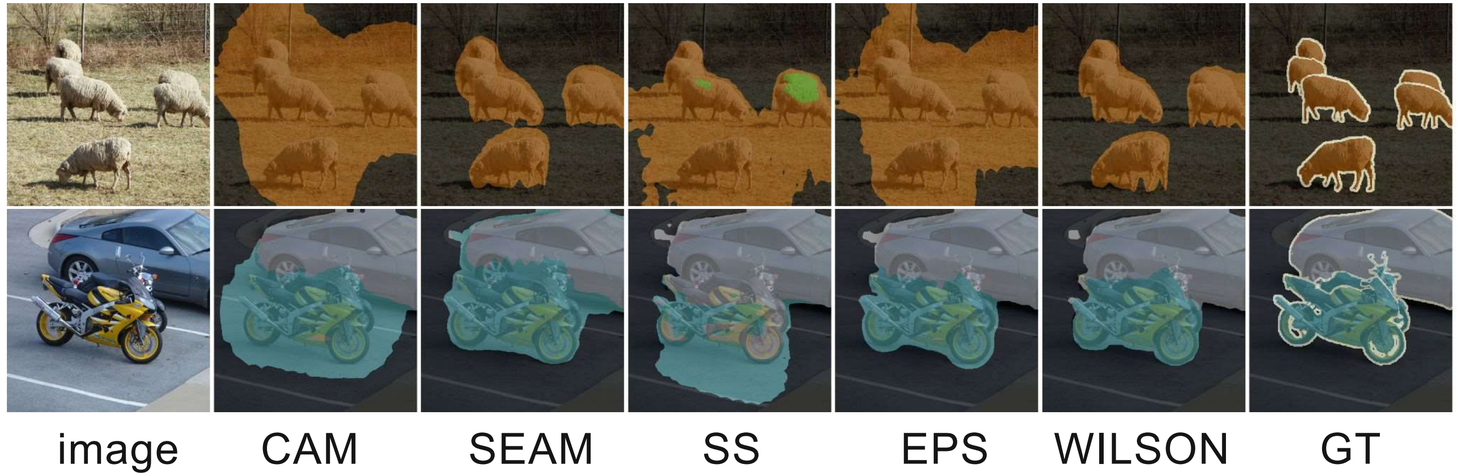


Weakly supervised Incremental Learning for Semantic Segmentation

Fabio Cermelli, Dario Fontanel, Antonio Tavera, Marco Ciccone, Barbara Caputo

Results





Old classes: car

New classes: sheep, motorbike

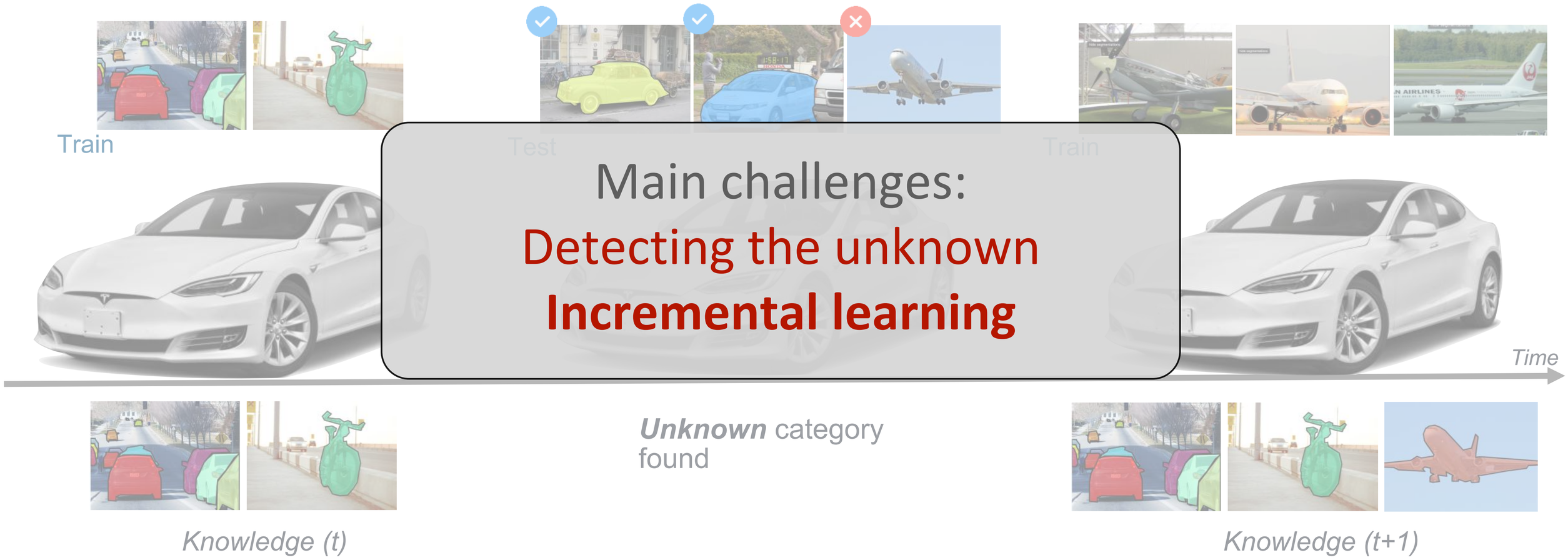
“Boosting Deep Open World Recognition by Clustering”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Samuel Rota Bulo’, Elisa Ricci, Barbara Caputo

IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) 2020, Las Vegas, USA
IEEE Robotics and Automation Letters 2020

Preliminaries

Open world



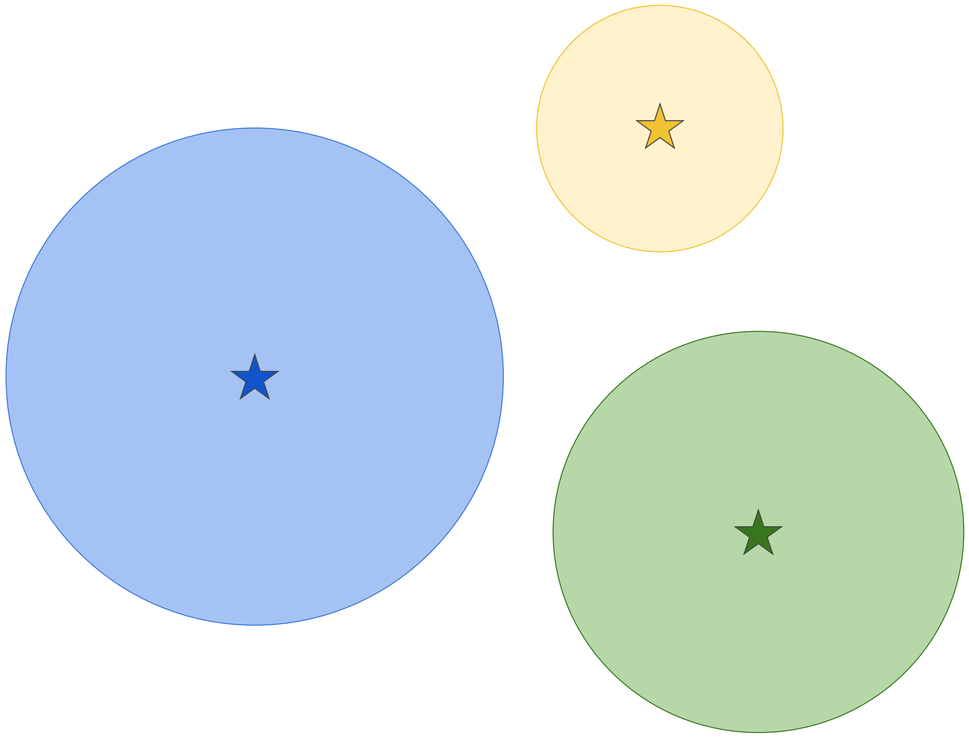
Main challenges:

Detecting the unknown

Incremental learning

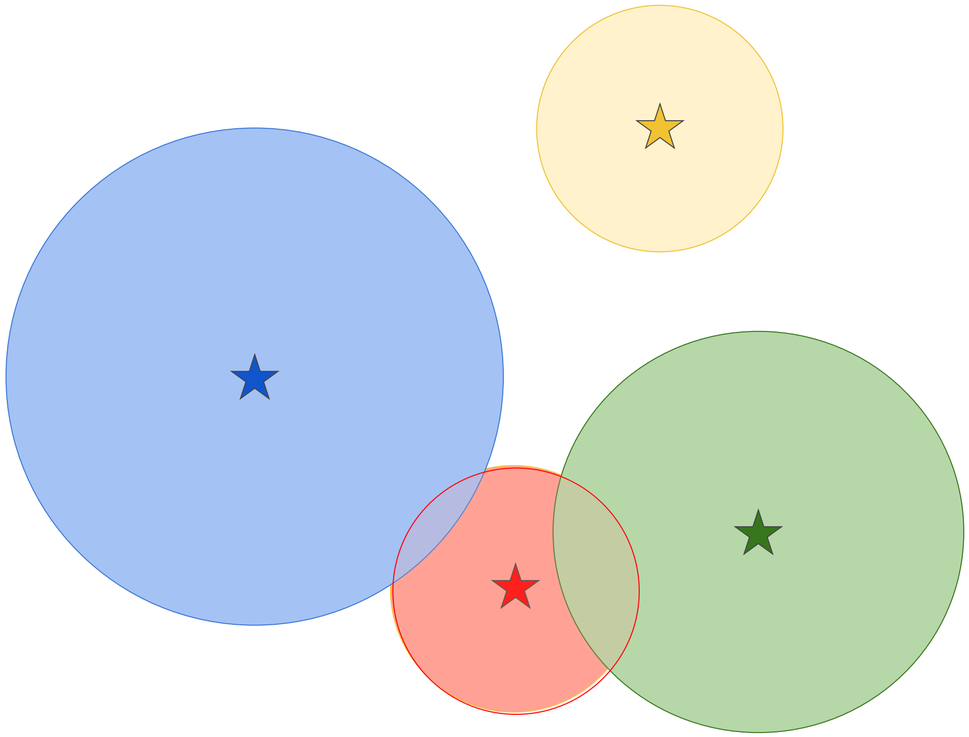
“Boosting Deep Open World Recognition by Clustering (B-DOC)”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Samuel Rota Bulo’, Elisa Ricci, Barbara Caputo



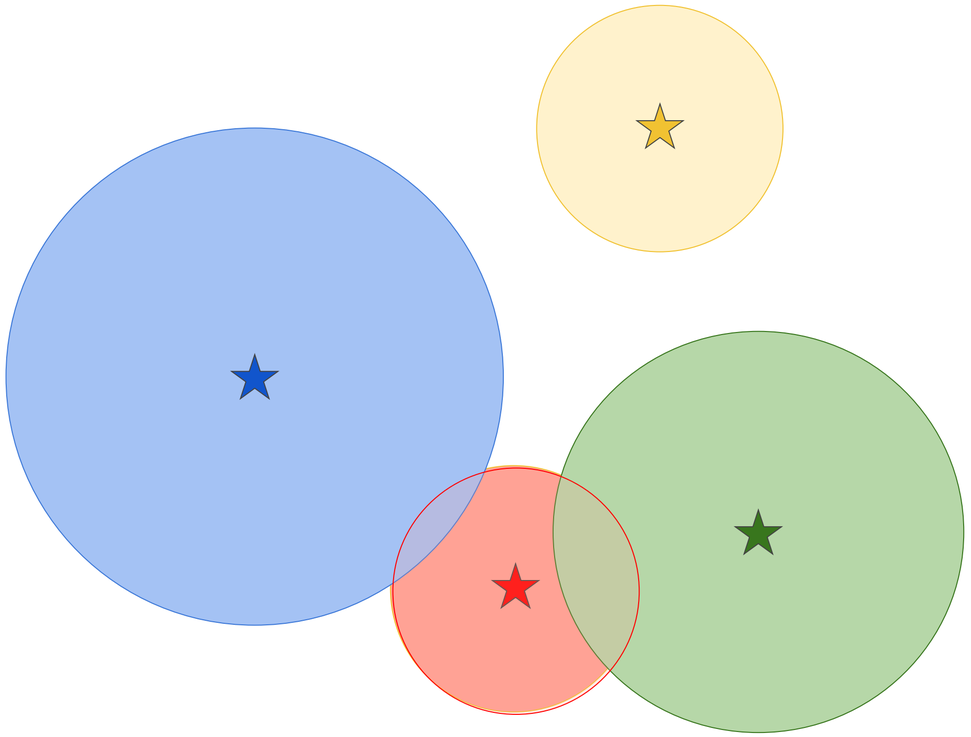
“Boosting Deep Open World Recognition by Clustering (B-DOC)”

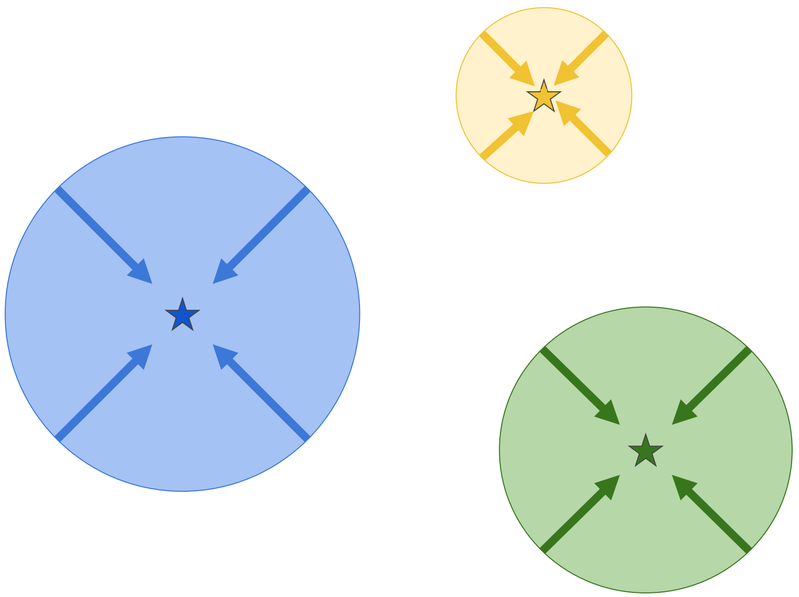
Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Samuel Rota Bulo’, Elisa Ricci, Barbara Caputo



“Boosting Deep Open World Recognition by Clustering (B-DOC)”

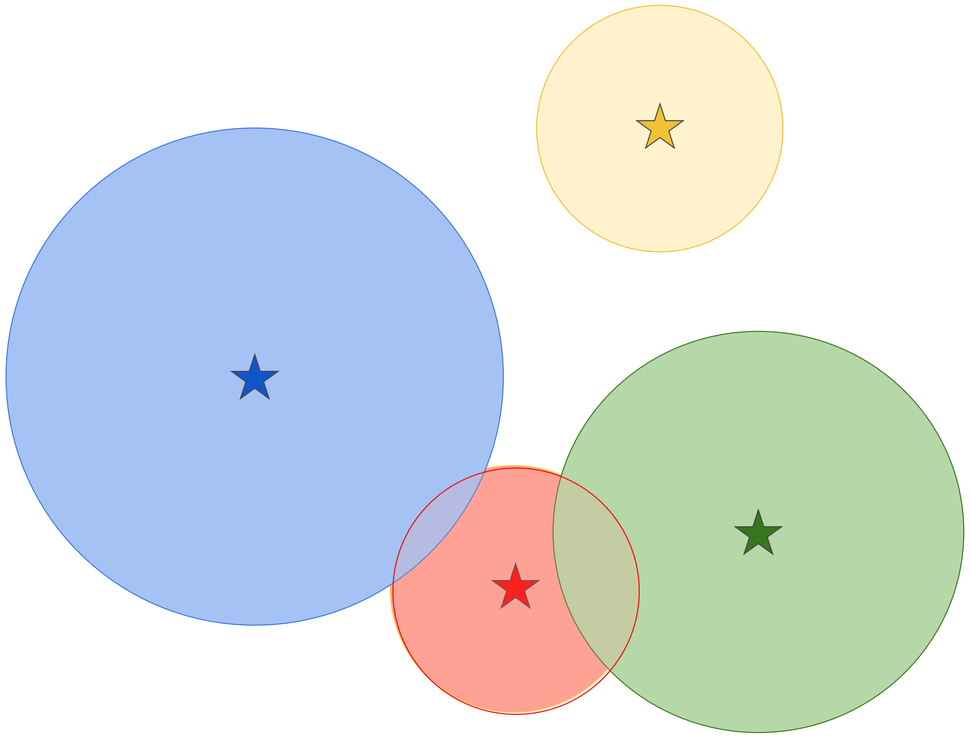
Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Samuel Rota Bulo’, Elisa Ricci, Barbara Caputo

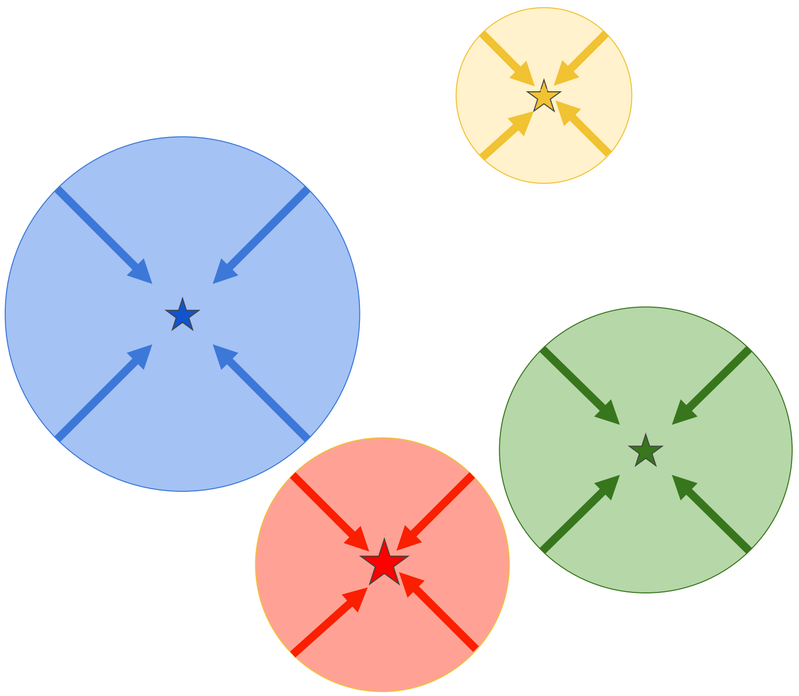




“Boosting Deep Open World Recognition by Clustering (B-DOC)”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Samuel Rota Bulo’, Elisa Ricci, Barbara Caputo





“Boosting Deep Open World Recognition by Clustering (B-DOC)”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Samuel Rota Bulo’, Elisa Ricci, Barbara Caputo

Global clustering loss

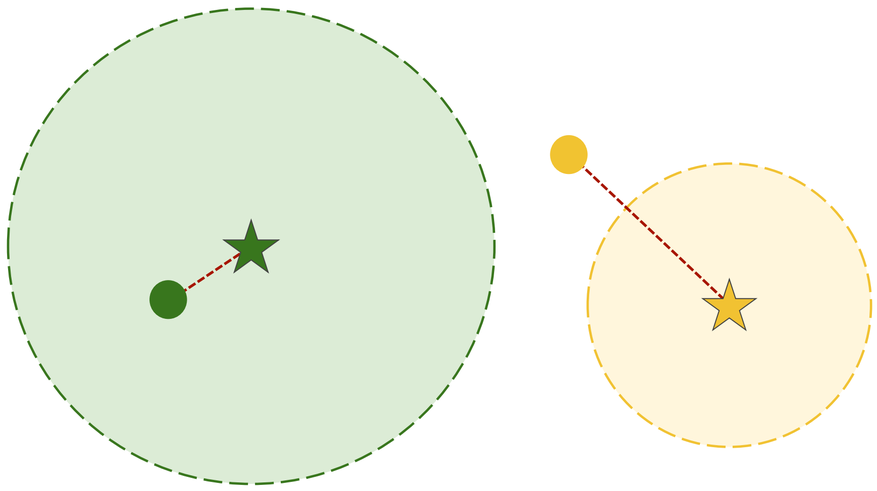
→ Scores based on ***distances*** between samples **features** and **respective class centroids**

“Boosting Deep Open World Recognition by Clustering (B-DOC)”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Samuel Rota Bulo’, Elisa Ricci, Barbara Caputo

Global clustering loss

Scores based on ***distances*** between samples **features** and **respective class centroids**



high score
minor update

low score
major update

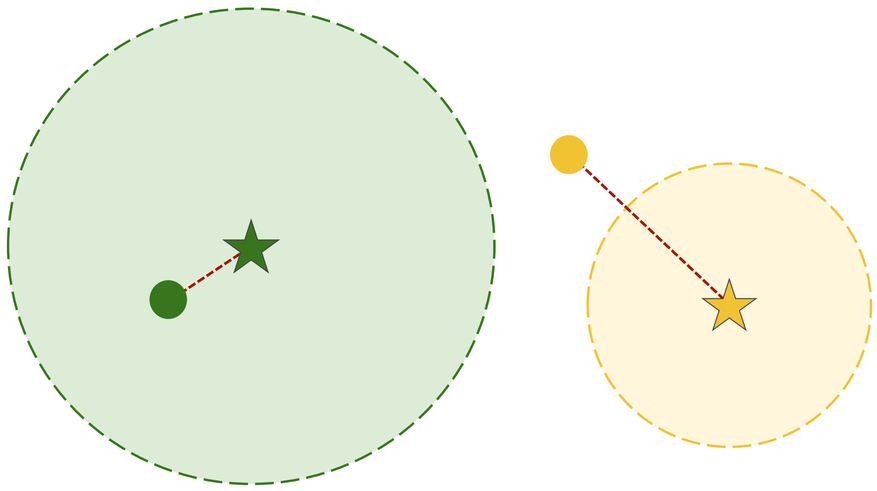
“Boosting Deep Open World Recognition by Clustering (B-DOC)”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Samuel Rota Bulo’, Elisa Ricci, Barbara Caputo

Global clustering loss

Scores based on ***distances*** between samples **features** and **respective class centroids**

The network learns **to keep representations** of images **as close as possible** to the **centroids** of the **respective class**



high score
minor update

low score
major update

“Boosting Deep Open World Recognition by Clustering (B-DOC)”

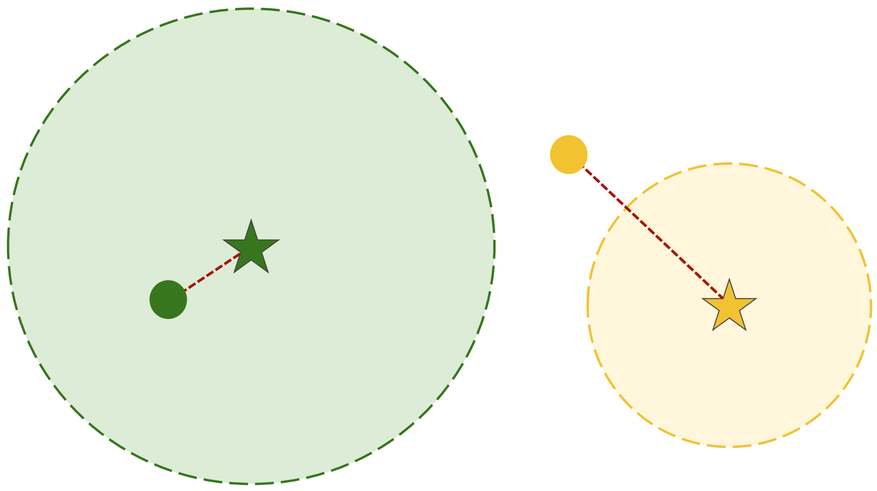
Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Samuel Rota Bulo’, Elisa Ricci, Barbara Caputo

Global clustering loss

Scores based on ***distances*** between samples **features** and **respective class centroids**

Cross entropy loss

The network learns **to keep representations** of images **as close as possible** to the **centroids** of the **respective class**



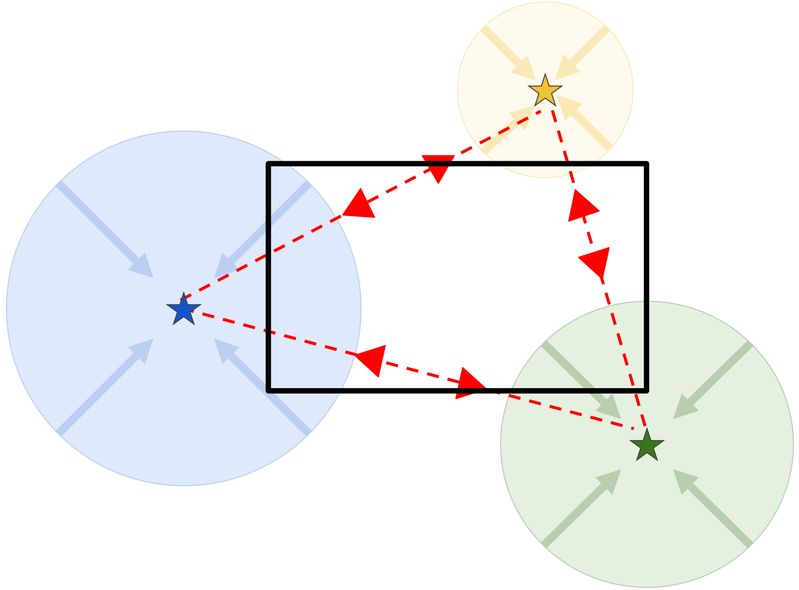
high score
minor update

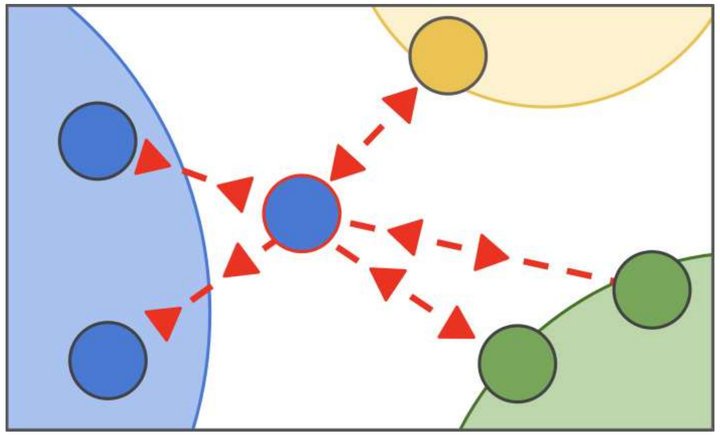
low score
major update

“Boosting Deep Open World Recognition by Clustering (B-DOC)”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Samuel Rota Bulo’, Elisa Ricci, Barbara Caputo

Local clustering loss



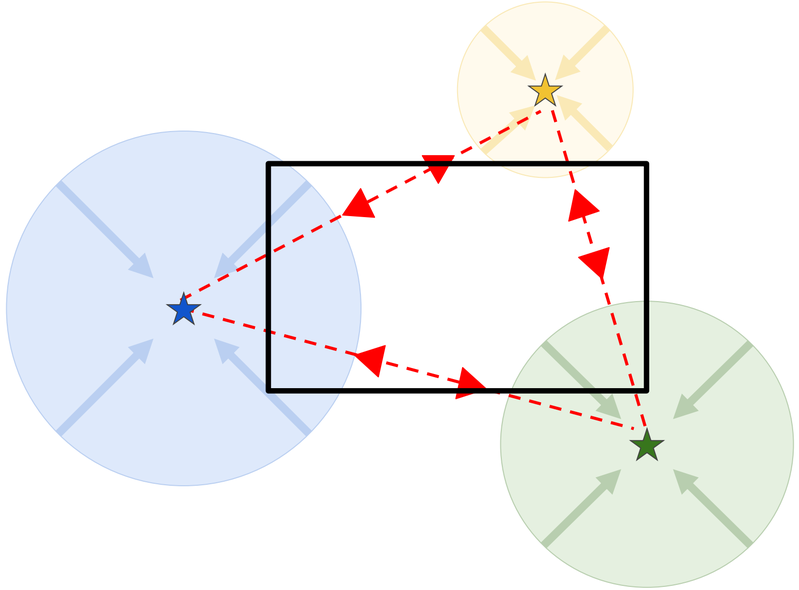


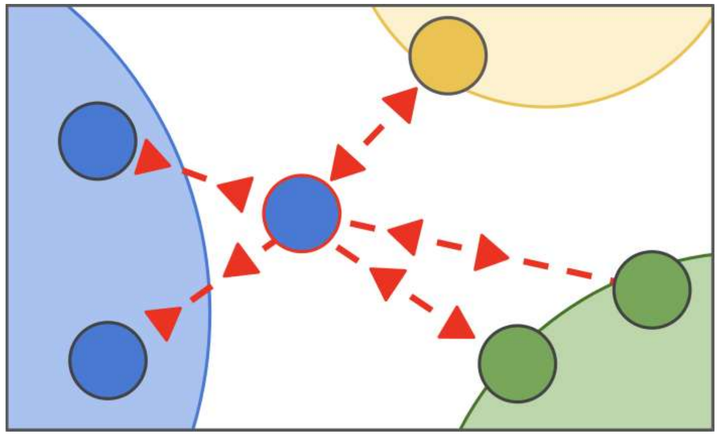
“Boosting Deep Open World Recognition by Clustering (B-DOC)”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Samuel Rota Bulo’, Elisa Ricci, Barbara Caputo

Local clustering loss

Scores based on ***distances*** between **features** of samples of **different classes**

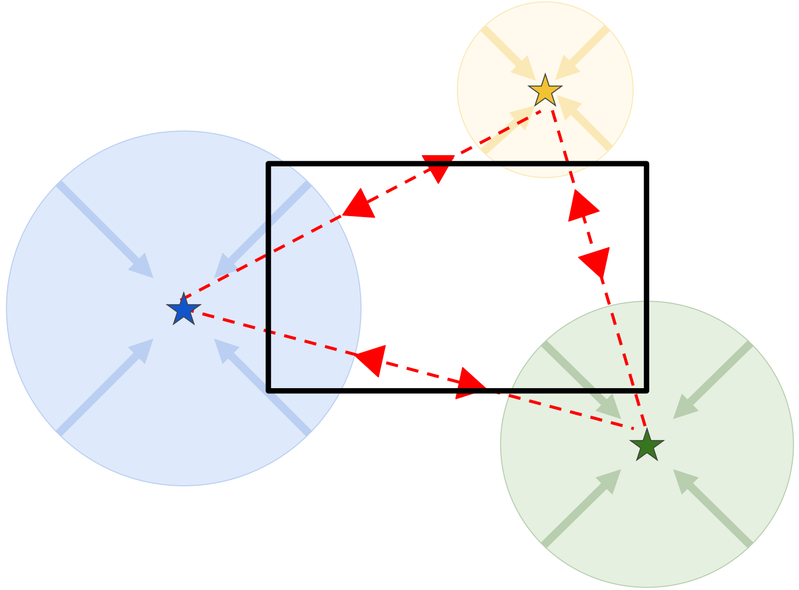




“Boosting Deep Open World Recognition by Clustering (B-DOC)”

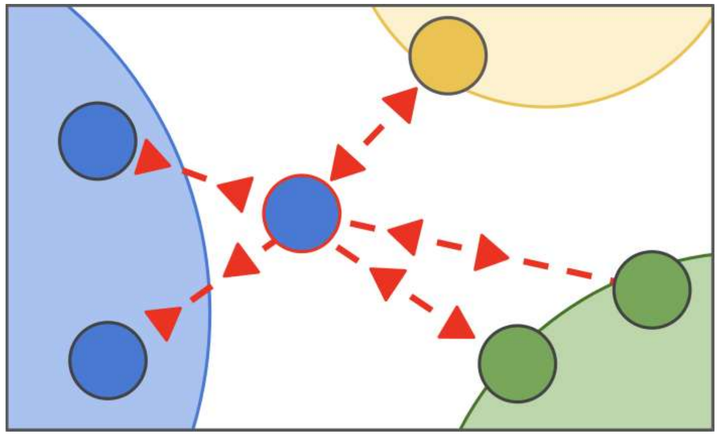
Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Samuel Rota Bulo’, Elisa Ricci, Barbara Caputo

Local clustering loss



Scores based on ***distances*** between **features** of samples of **different classes**

The network learns to keep samples of the **same class closer**, while **pushing away** samples of **different classes**



“Boosting Deep Open World Recognition by Clustering (B-DOC)”

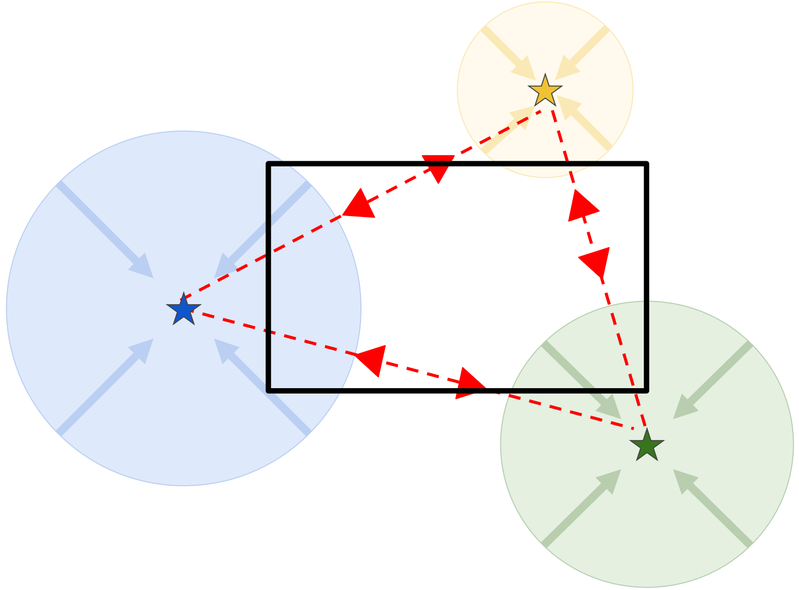
Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Samuel Rota Bulo’, Elisa Ricci, Barbara Caputo

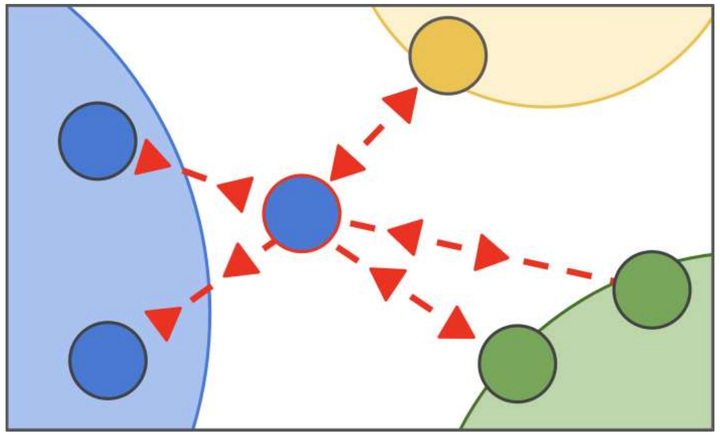
Local clustering loss

Scores based on ***distances*** between **features** of samples of **different classes**

Local clustering loss

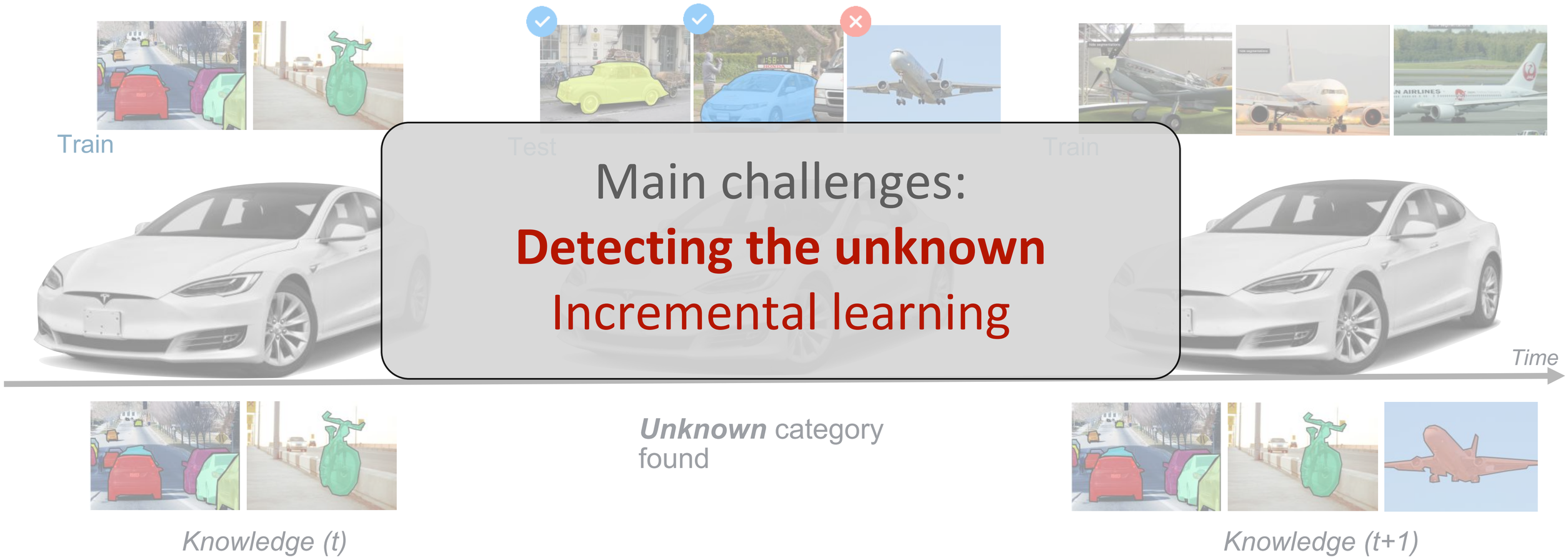
The network learns to keep samples of the **same class closer**, while **pushing away** samples of **different classes**





Preliminaries

Open world



Main challenges:

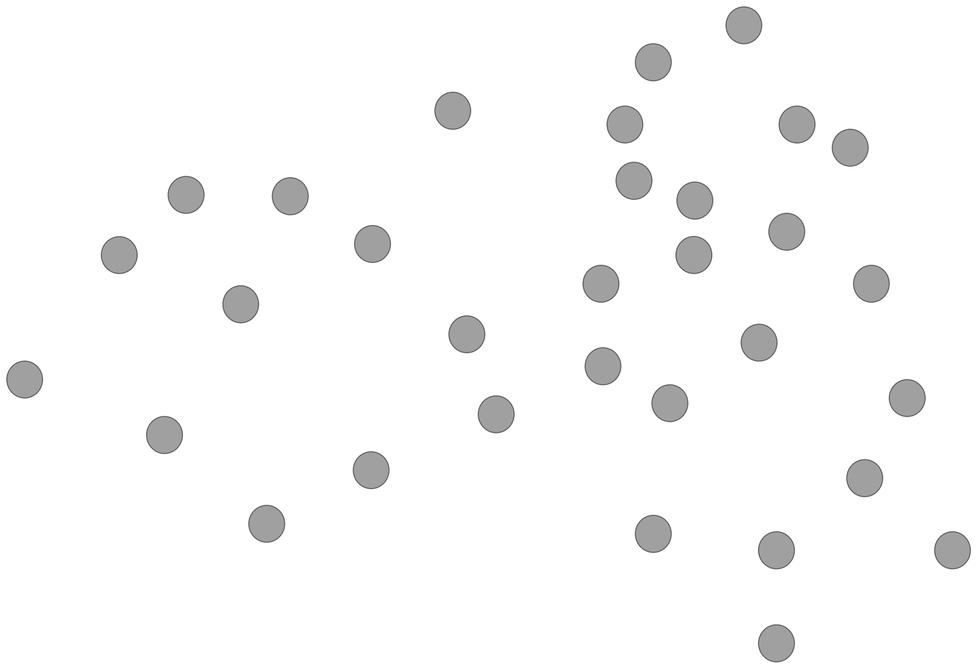
Detecting the unknown

Incremental learning

“Boosting Deep Open World Recognition by Clustering (B-DOC)”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Samuel Rota Bulo’, Elisa Ricci, Barbara Caputo

Class-specific rejection thresholds

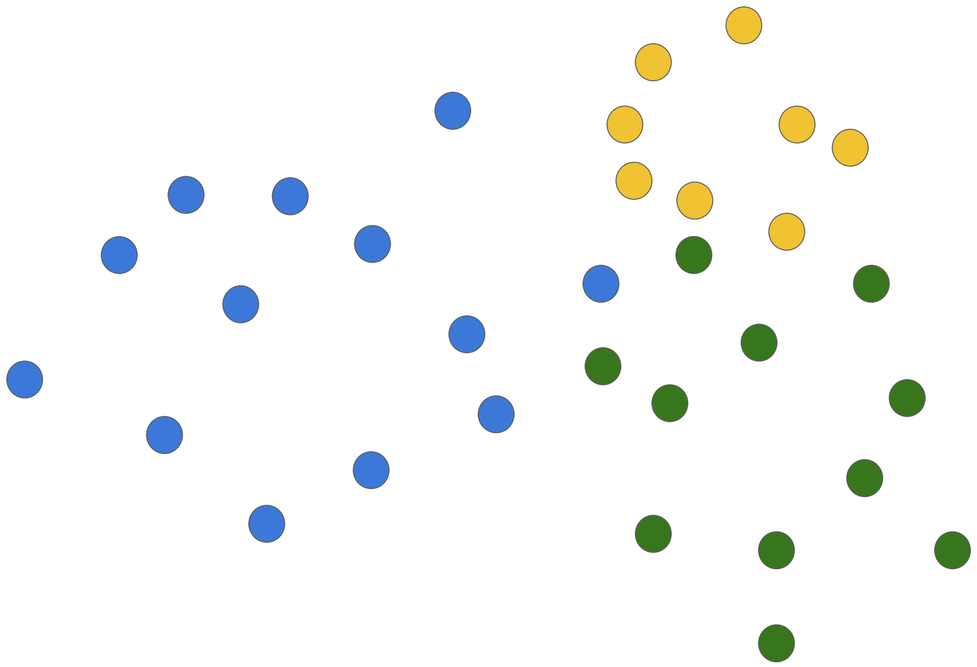


To exploit the clustered feature space, we introduce class-specific rejection thresholds Δ based on distances, *learned* on an held out set

“Boosting Deep Open World Recognition by Clustering (B-DOC)”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Samuel Rota Bulo’, Elisa Ricci, Barbara Caputo

Class-specific rejection thresholds

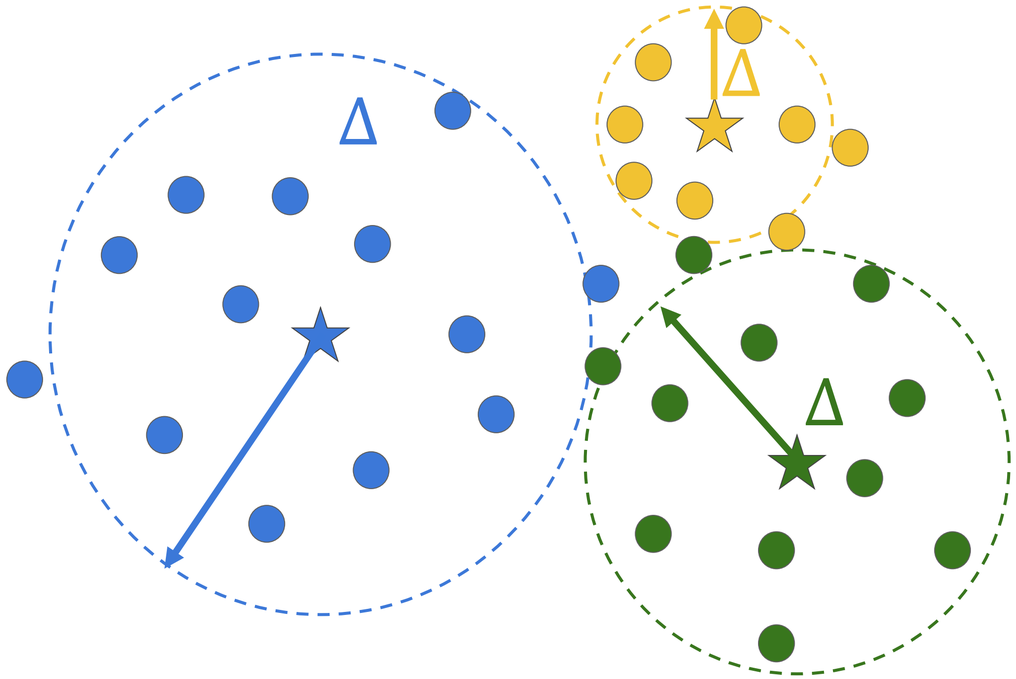


To exploit the clustered feature space, we introduce class-specific rejection thresholds Δ based on distances, *learned* on an held out set

“Boosting Deep Open World Recognition by Clustering (B-DOC)”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Samuel Rota Bulo’, Elisa Ricci, Barbara Caputo

Class-specific rejection thresholds

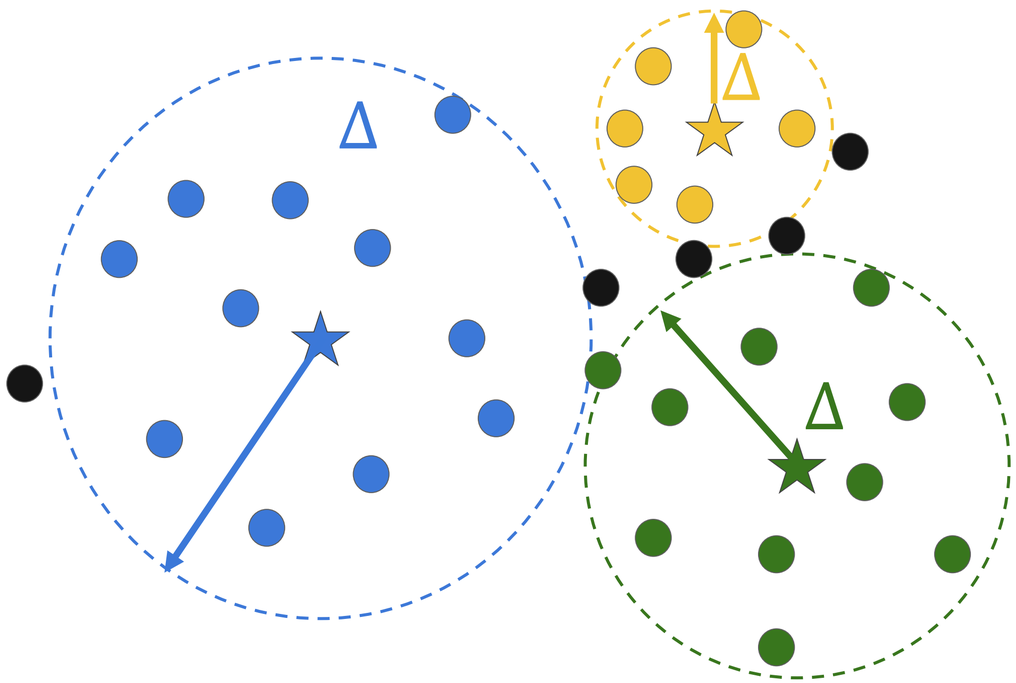


To exploit the clustered feature space, we introduce class-specific rejection thresholds Δ based on distances, *learned* on an held out set

“Boosting Deep Open World Recognition by Clustering (B-DOC)”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Samuel Rota Bulo’, Elisa Ricci, Barbara Caputo

Class-specific rejection thresholds

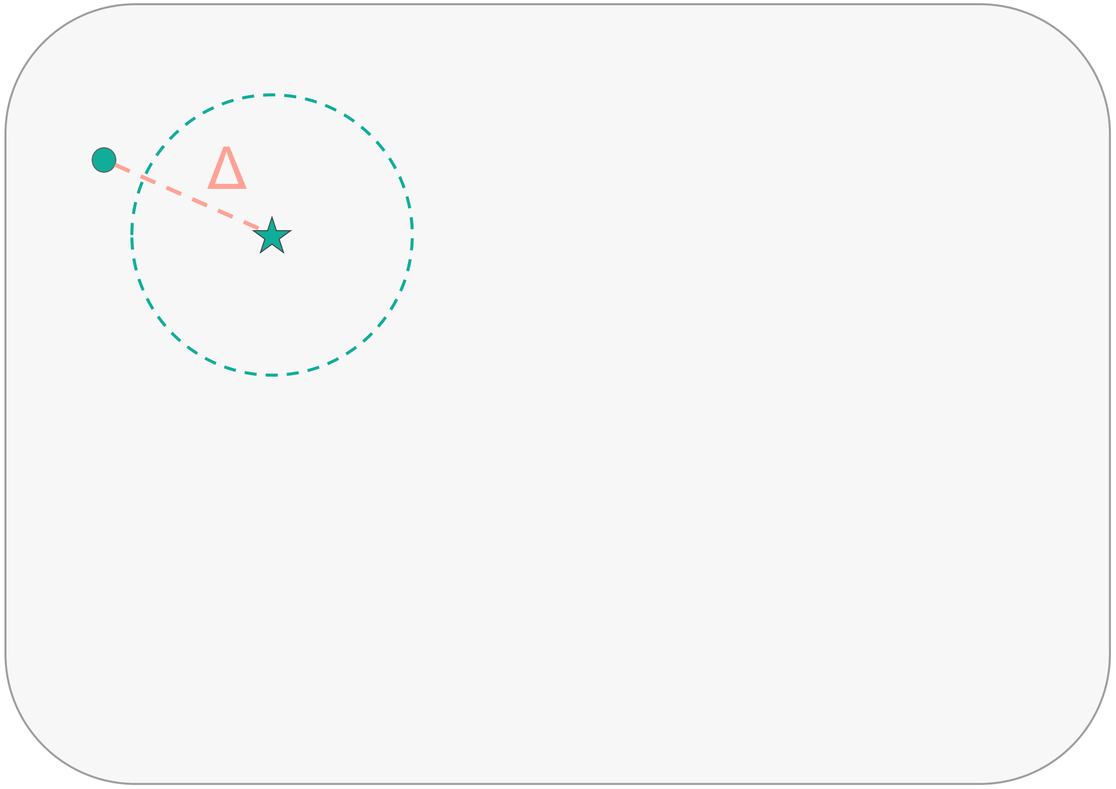


To exploit the clustered feature space, we introduce class-specific rejection thresholds Δ based on distances, *learned* on an held out set

“Boosting Deep Open World Recognition by Clustering (B-DOC)”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Samuel Rota Bulo’, Elisa Ricci, Barbara Caputo

Class-specific rejection thresholds



To exploit the clustered feature space, we introduce class-specific rejection thresholds Δ based on distances, *learned* on an held out set

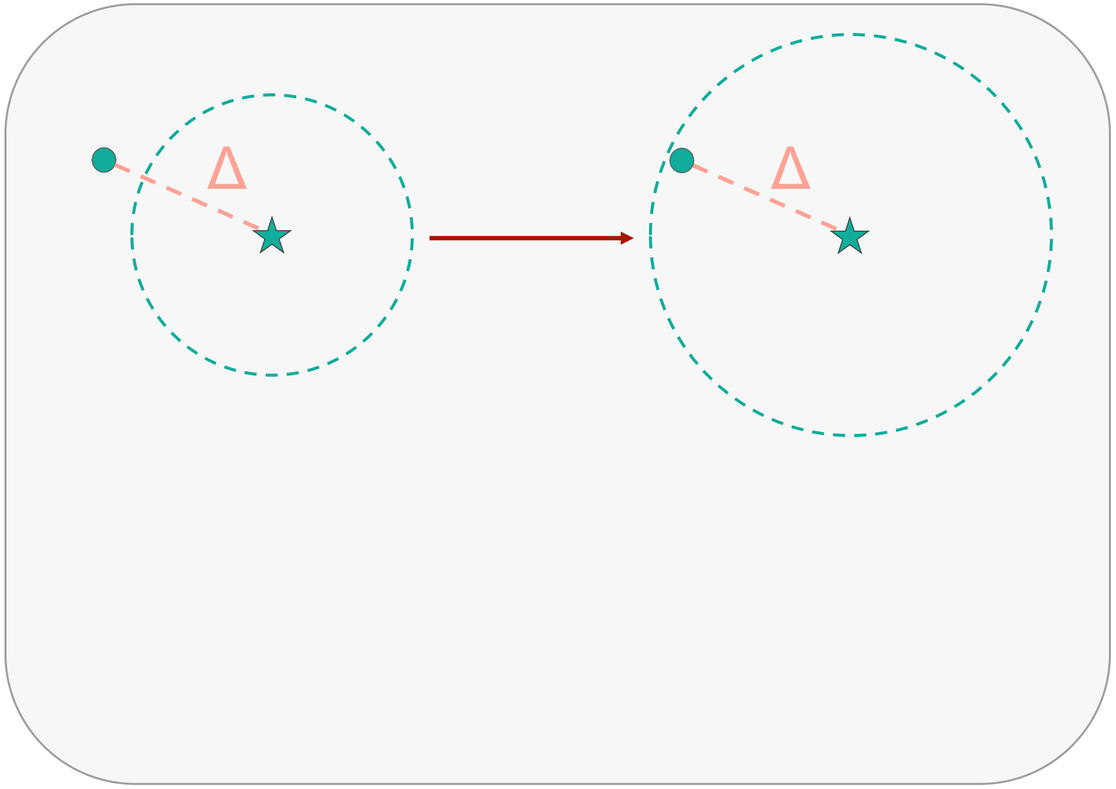
→ ***Maximum distance loss***

If the **distance** between a sample of class y and its class centroid is **greater** than Δ_y, the loss **increases** Δ_y.

“Boosting Deep Open World Recognition by Clustering (B-DOC)”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Samuel Rota Bulo’, Elisa Ricci, Barbara Caputo

Class-specific rejection thresholds



To exploit the clustered feature space, we introduce class-specific rejection thresholds Δ based on distances, *learned* on an held out set

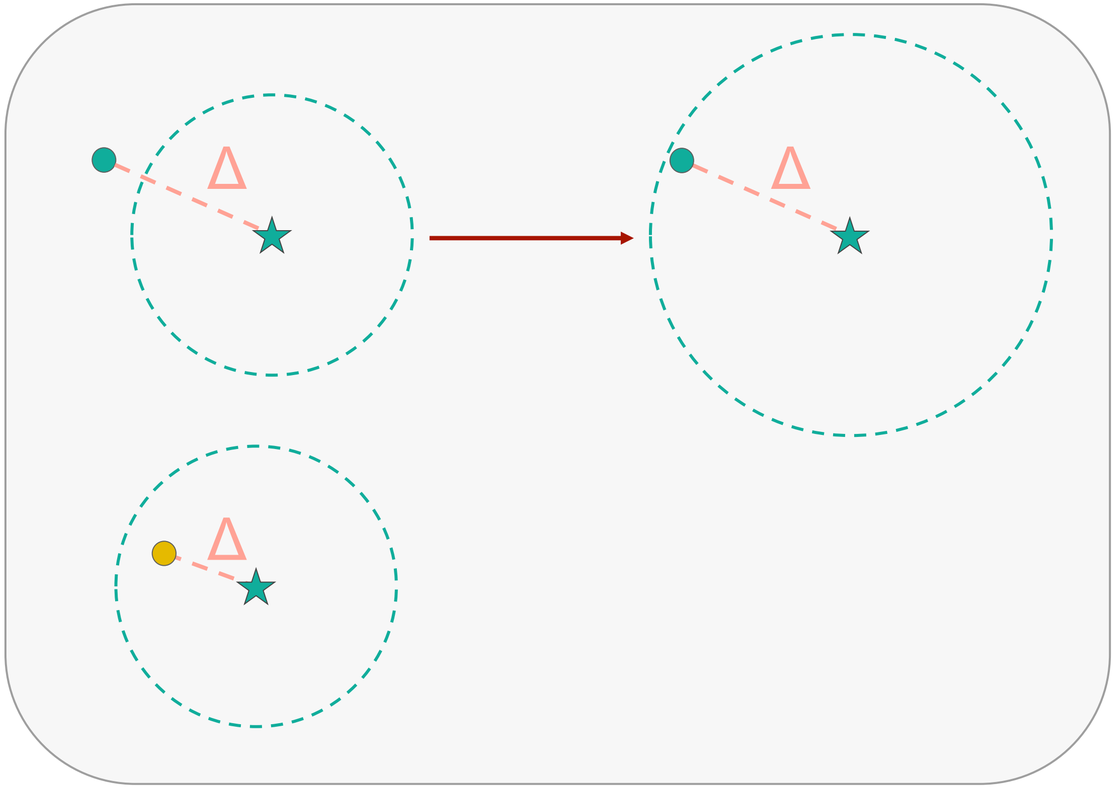
→ ***Maximum distance loss***

If the **distance** between a sample of class y and its class centroid is **greater** than Δ_y, the loss **increases** Δ_y.

“Boosting Deep Open World Recognition by Clustering (B-DOC)”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Samuel Rota Bulo’, Elisa Ricci, Barbara Caputo

Class-specific rejection thresholds



To exploit the clustered feature space, we introduce class-specific rejection thresholds Δ based on distances, *learned* on an held out set

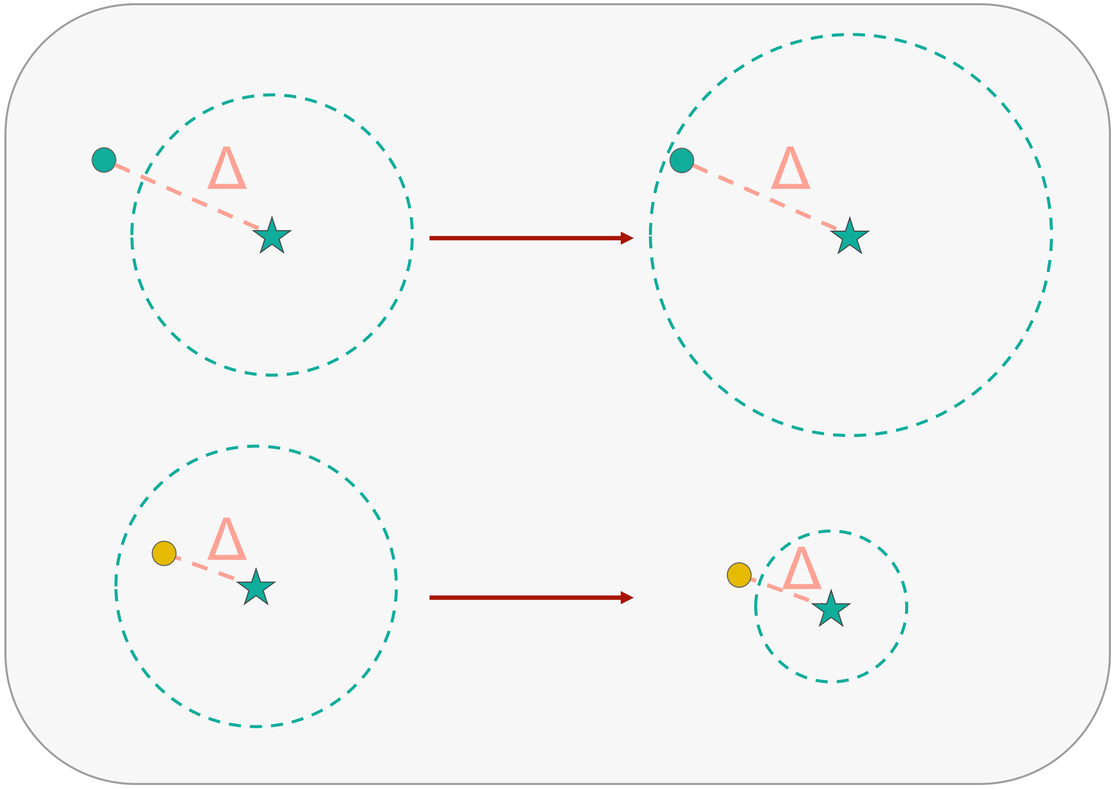
→ ***Maximum distance loss***

If the **distance** between a sample of class y and its class centroid is **greater** than Δ_y, the loss **increases** Δ_y.
If the **distance** between a sample *not* in class y and y class centroid is **less** than Δ_y, the loss **decreases** Δ_y.

“Boosting Deep Open World Recognition by Clustering (B-DOC)”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Samuel Rota Bulo’, Elisa Ricci, Barbara Caputo

Class-specific rejection thresholds



To exploit the clustered feature space, we introduce class-specific rejection thresholds Δ based on distances, *learned* on an held out set

→ ***Maximum distance loss***

If the **distance** between a sample of class y and its class centroid is **greater** than Δ_y, the loss **increases** Δ_y.
If the **distance** between a sample *not* in class y and y class centroid is **less** than Δ_y, the loss **decreases** Δ_y.

“Boosting Deep Open World Recognition by Clustering (B-DOC)”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Samuel Rota Bulo’, Elisa Ricci, Barbara Caputo

Experimental settings

Comparisons

Incremental learning methods

- LwF
- iCaRL

Open World Recognition methods

- NNO
- DeepNNO
- B-DOC

“Boosting Deep Open World Recognition by Clustering (B-DOC)”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Samuel Rota Bulo’, Elisa Ricci, Barbara Caputo

Experimental settings

Without Unknown (no rejection)

Test set does **not** contain **unknown** classes

Comparisons

Incremental learning methods

- LwF
- iCaRL

Open World Recognition methods

- NNO
- DeepNNO
- B-DOC

“Boosting Deep Open World Recognition by Clustering (B-DOC)”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Samuel Rota Bulo’, Elisa Ricci, Barbara Caputo

Experimental settings

Comparisons

Incremental learning methods

- LwF
- iCaRL

Open World Recognition methods

- NNO
- DeepNNO
- B-DOC

Without Unknown (no rejection)

Test set does **not** contain **unknown** classes

Without Unknown (with rejection)

Test set does **not** contain **unknown** classes but the model has the possibility to **reject**

“Boosting Deep Open World Recognition by Clustering (B-DOC)”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Samuel Rota Bulo’, Elisa Ricci, Barbara Caputo

Experimental settings

Comparisons

Incremental learning methods

LwF
iCaRL

Open World Recognition methods

NNO
DeepNNO
B-DOC

Without Unknown (no rejection)

Test set does **not** contain **unknown** classes

Without Unknown (with rejection)

Test set does **not** contain **unknown** classes but the model has the possibility to **reject**

Open World Recognition (Avg)

Average between Without Unknown (**with rejection**) and With Unknown (**with rejection**)

“Boosting Deep Open World Recognition by Clustering (B-DOC)”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Samuel Rota Bulo’, Elisa Ricci, Barbara Caputo

Experimental settings

Comparisons

Incremental learning methods

LwF
iCaRL

Open World Recognition methods

NNO
DeepNNO
B-DOC

Without Unknown (no rejection)

Test set does **not** contain **unknown** classes

Without Unknown (with rejection)

Test set does **not** contain **unknown** classes but the model has the possibility to **reject**

Open World Recognition (Avg)

Average between Without Unknown (**with rejection**) and With Unknown (**with rejection**)

Open World Recognition (Harm)

Harmonic mean between Without Unknown (**with rejection**) and With Unknown (**with rejection**)

“Boosting Deep Open World Recognition by Clustering (B-DOC)”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Samuel Rota Bulo’, Elisa Ricci, Barbara Caputo

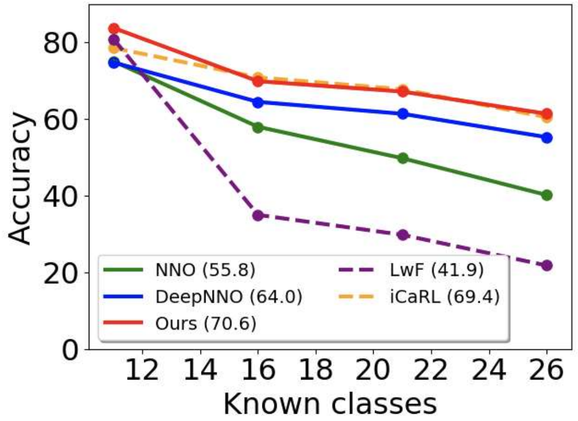
Results

RGB-D Object

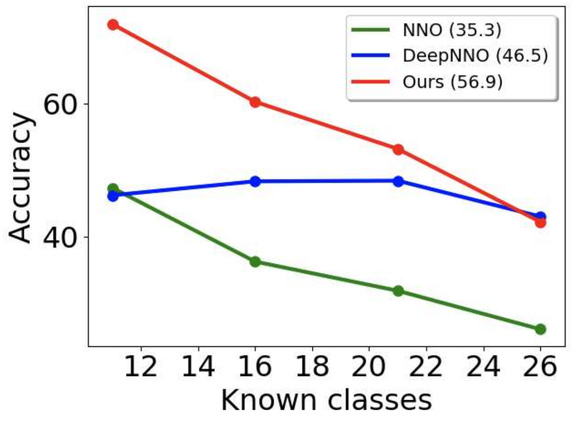
Categories: 26 known + 25 unknown
Incremental steps: 11 first + 5 each

K. Lai, L. Bo, X. Ren, D. Fox,
A large-scale hierarchical multi-view rgb-d object dataset, ICRA-11

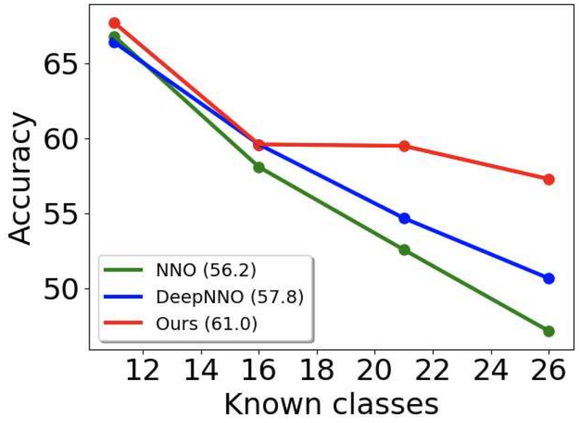
Without Unknown (no rejection)



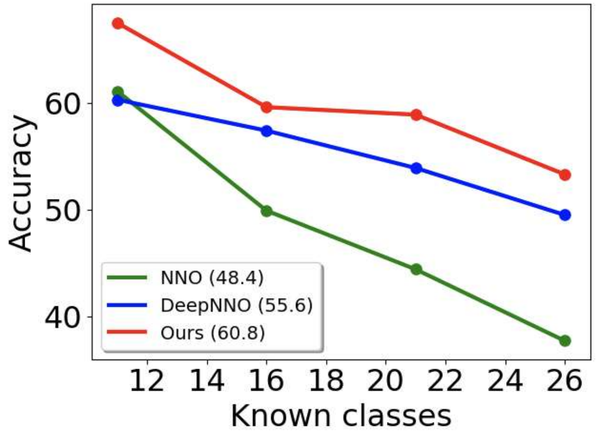
Without Unknown (with rejection)



Open World Recognition (Avg)



Open World Recognition (Harm)



“Boosting Deep Open World Recognition by Clustering (B-DOC)”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Samuel Rota Bulo’, Elisa Ricci, Barbara Caputo

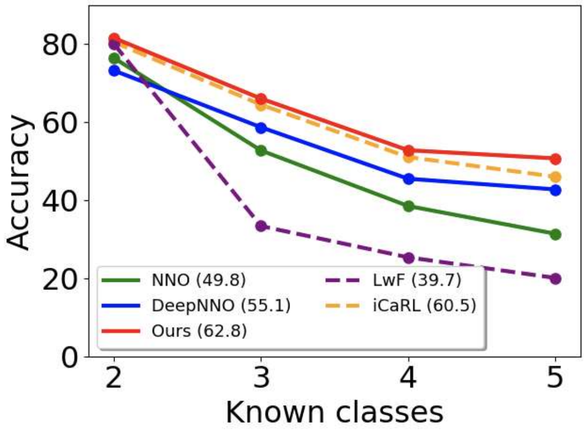
Results

CORE-50

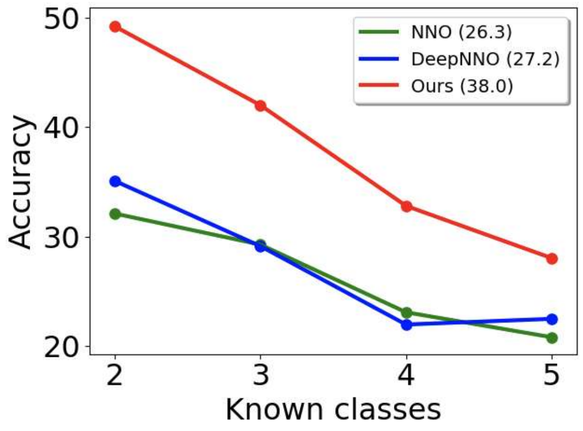
Categories: 5 known + 5 unknown
Incremental steps: 2 first + 1 each

V. Lomonaco and D. Maltoni,
Core50: a new dataset and benchmark for continuous object recognition, CoRL-17

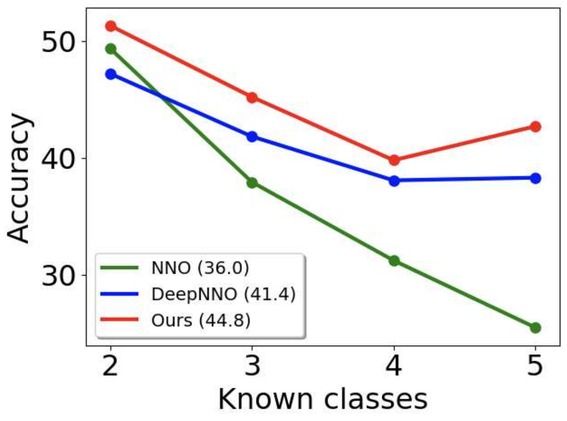
Without Unknown (no rejection)



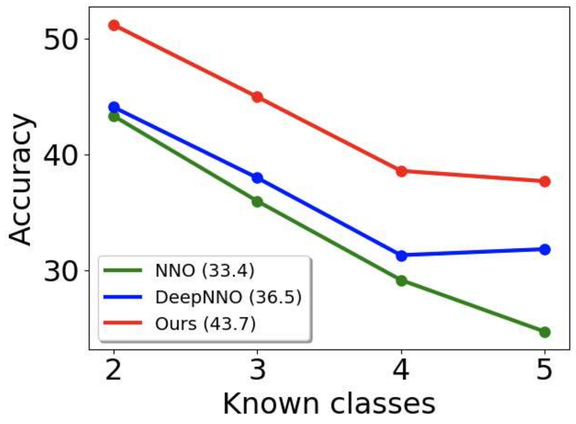
Without Unknown (with rejection)



Open World Recognition (Avg)



Open World Recognition (Harm)



“On the Challenges of Open World Recognition under Shifting Visual Domains”

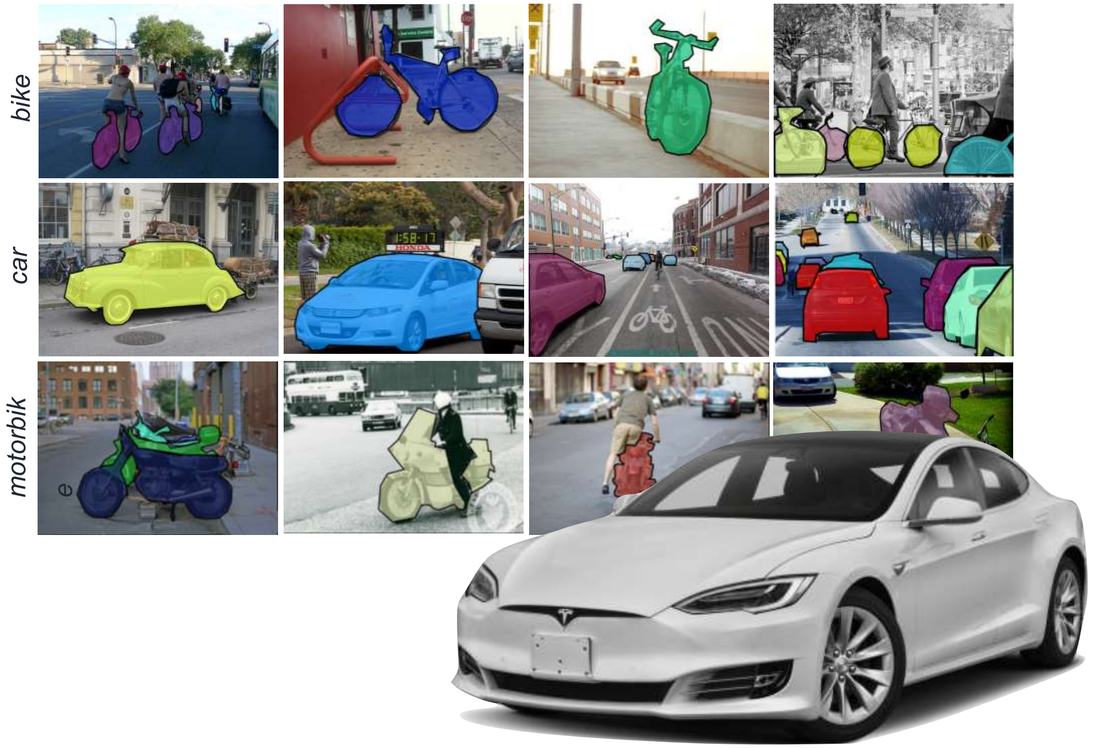
Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Barbara Caputo

IEEE International Conference on Robotics and Automation (ICRA) 2021, Xi’an, China
IEEE Robotics and Automation Letters 2021

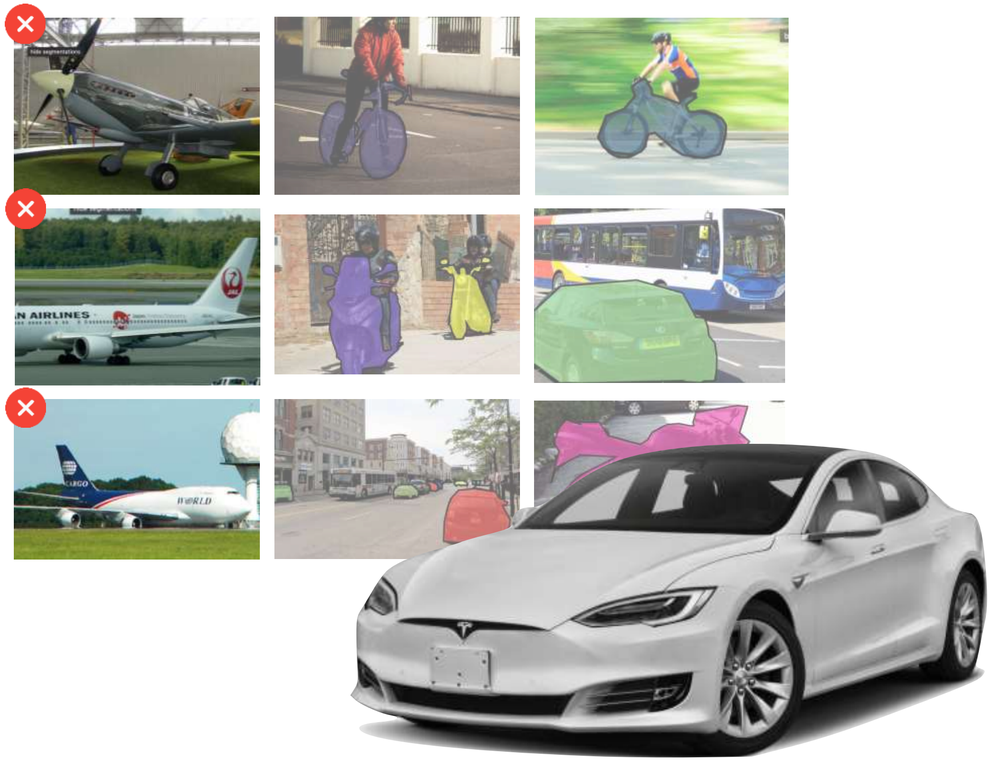
Preliminaries

Closed domain assumption

Training



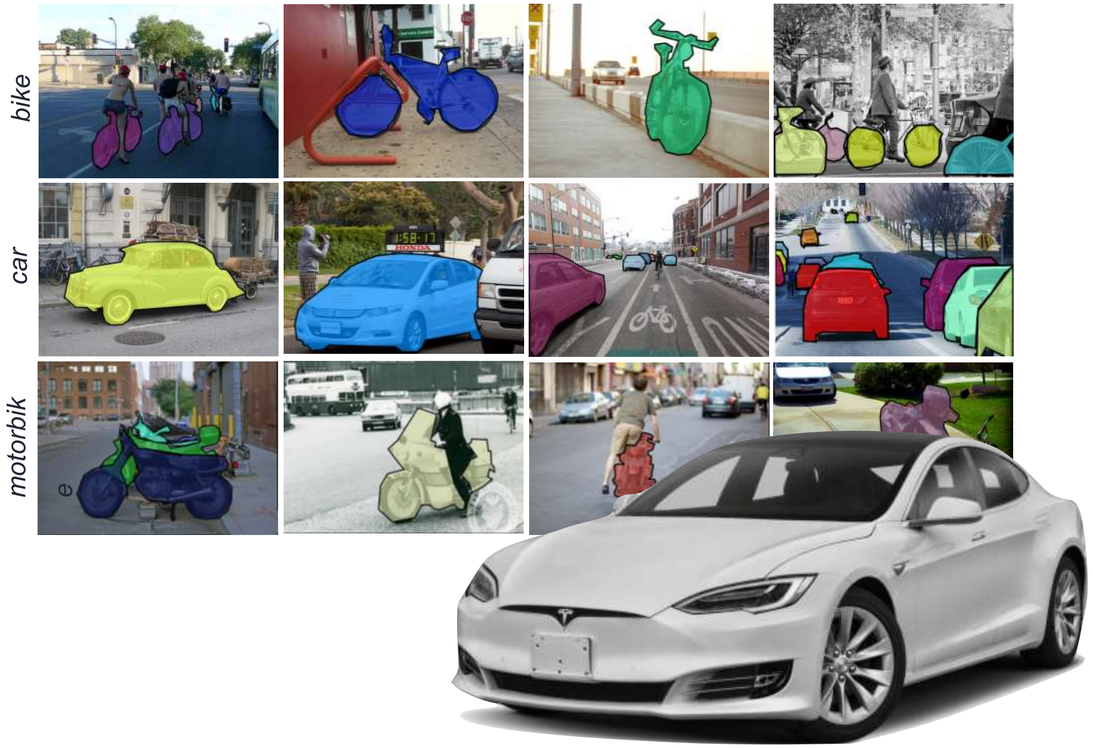
Deployment



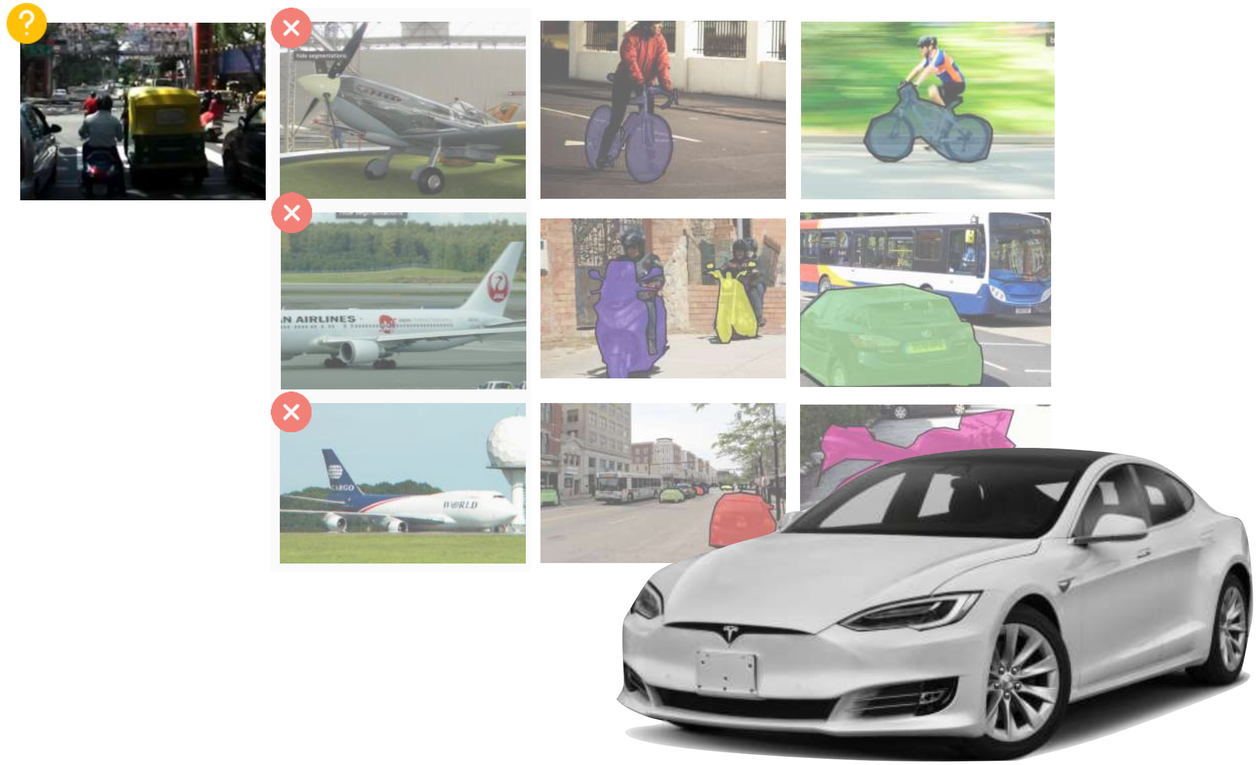
Preliminaries

Closed domain assumption

Training



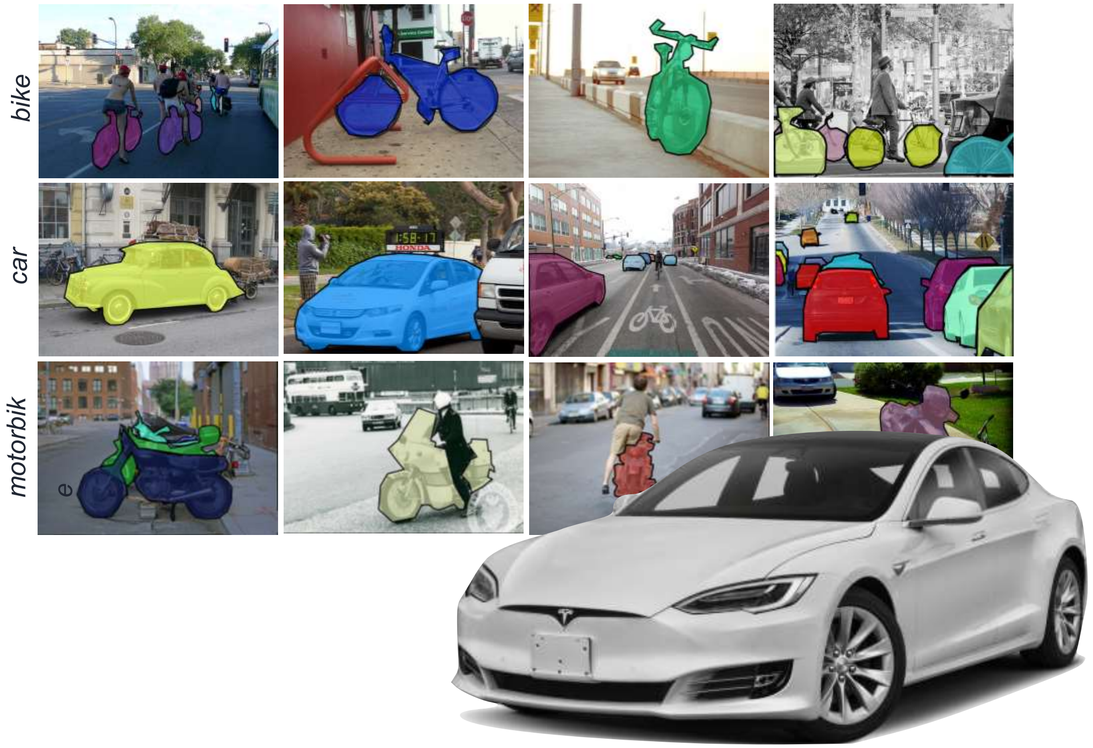
Deployment



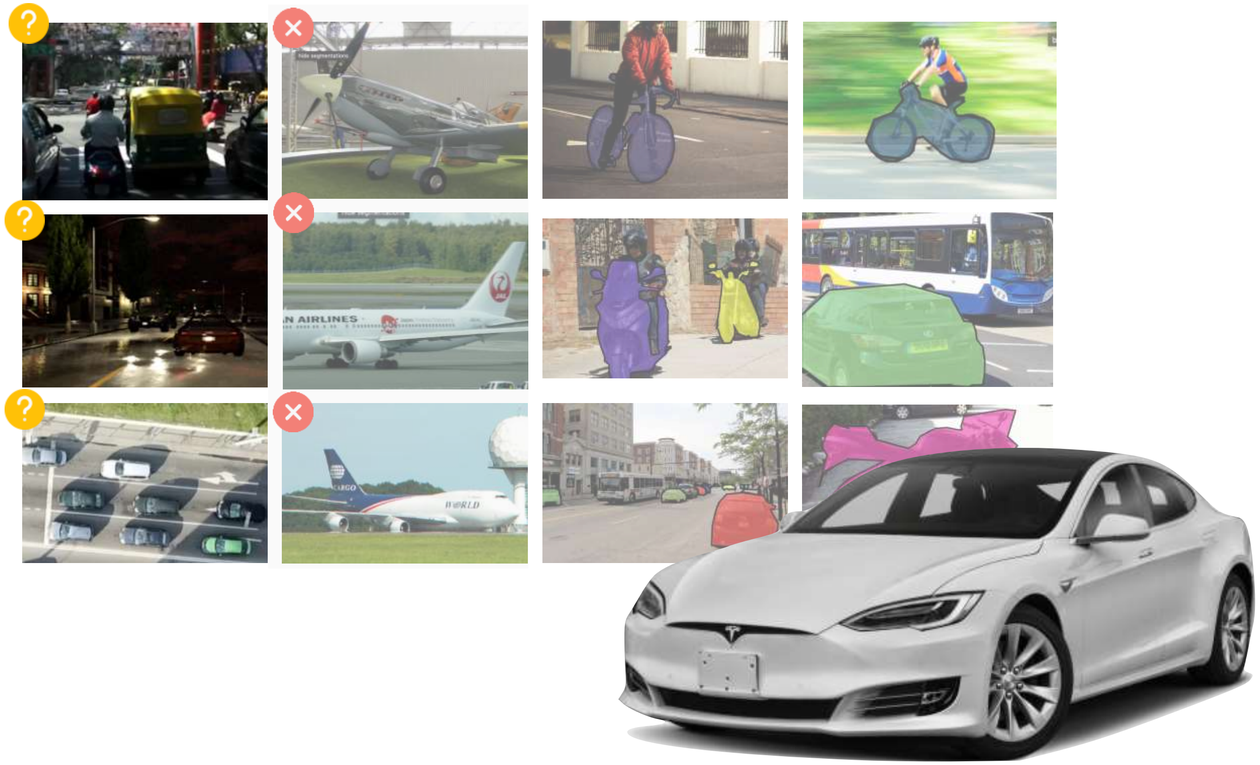
Preliminaries

Closed domain assumption

Training

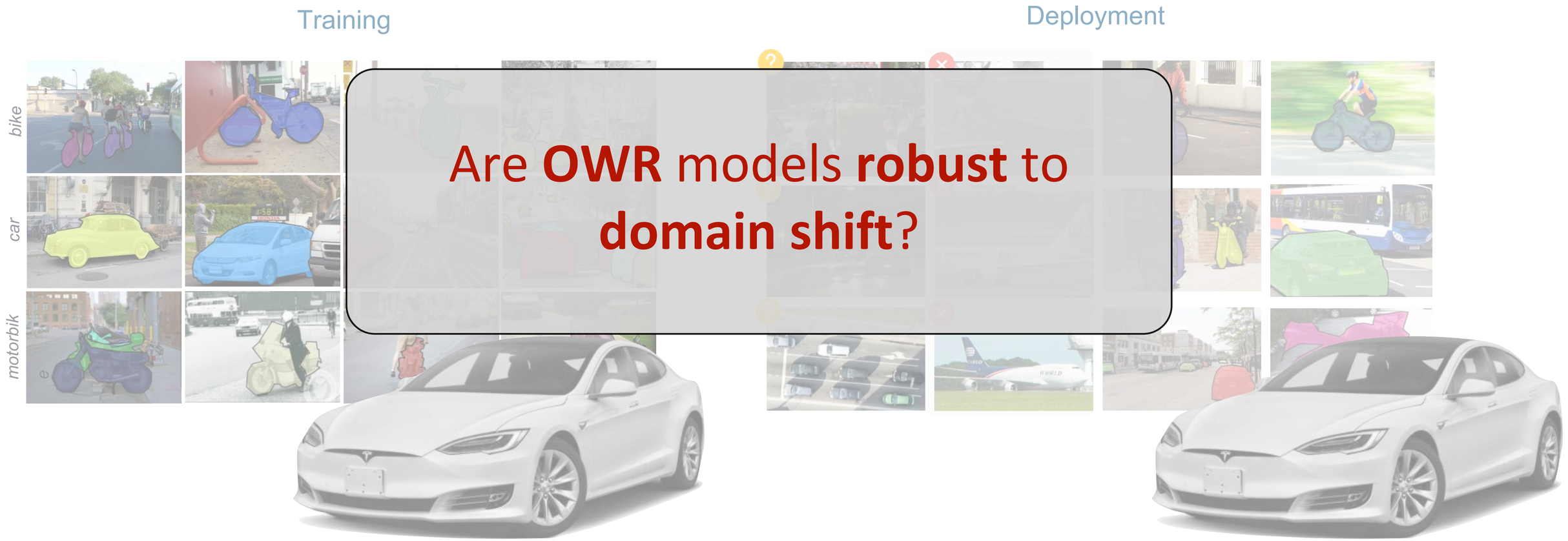


Deployment



Preliminaries

Closed domain assumption



Are **OWR** models **robust** to **domain shift**?

“On the Challenges of Open World Recognition under Shifting Visual Domains”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Barbara Caputo

Benchmark

- ***Three*** *datasets containing the same* ***51 daily-life objects***
- ***Different*** *acquisition* ***conditions***
- ***Standard protocol: 26 (known) + 25 (unknown)***
  11 (initial step) + 5 (each incremental step)

synROD **ROD** **ARID**

“On the Challenges of Open World Recognition under Shifting Visual Domains”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Barbara Caputo

Benchmark

synROD

Synthetic dataset generated utilizing publicly freely available 3D models.

Reproduction of realistic lighting.

Unsupervised Domain Adaptation through Inter-modal Rotation for RGB-D Object Recognition,
Loghmani Mohammad Reza, Robbiano Luca, Planamente Mirco, Park Kiru, Caputo Barbara and Vincze Markus, RAL 20

“On the Challenges of Open World Recognition under Shifting Visual Domains”

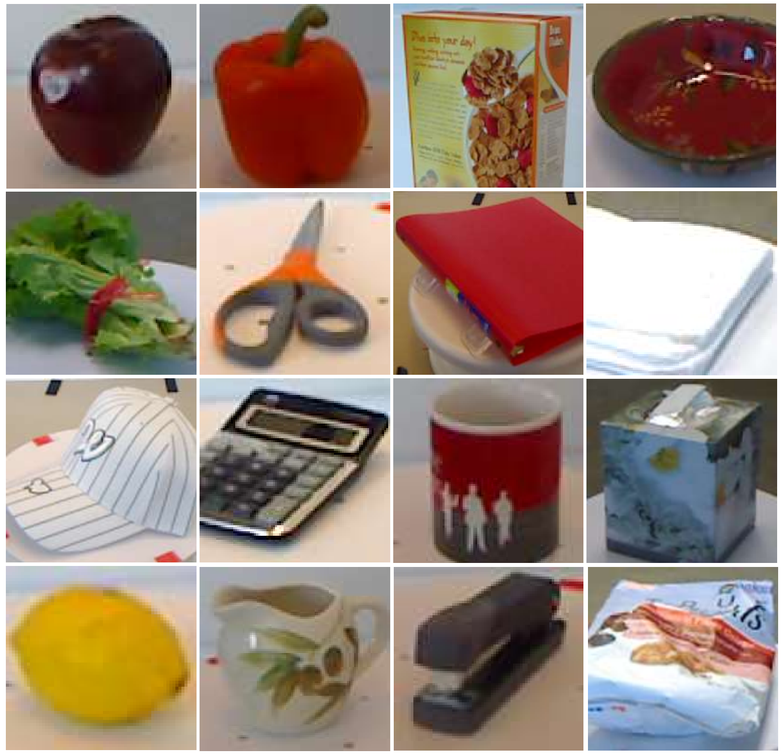
Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Barbara Caputo

Benchmark

ROD

Objects captured in a controlled scenario.

No clutter or changes in illumination or background, only varying camera angles.



“On the Challenges of Open World Recognition under Shifting Visual Domains”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Barbara Caputo

Benchmark

ARID

More challenging environment with objects depicted against various backgrounds, scales, views, lighting, and occlusions.

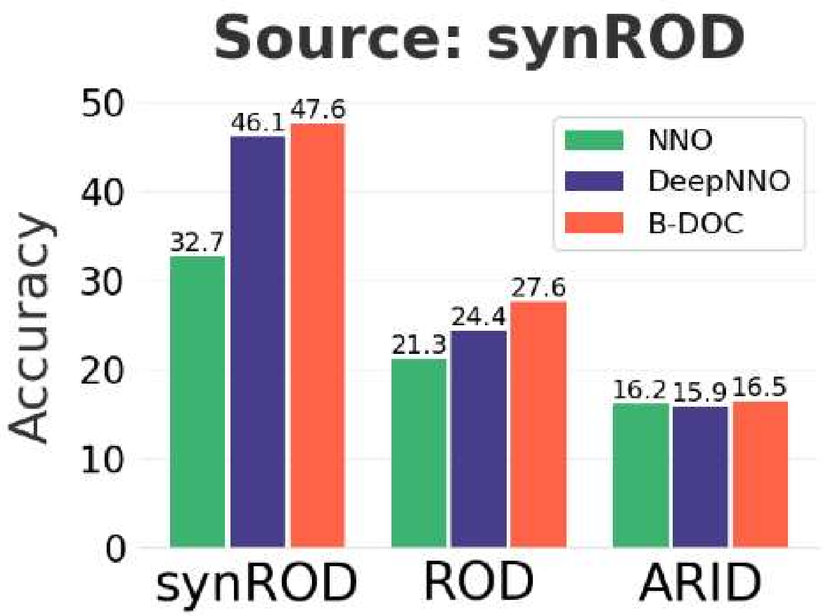
Recognizing Objects In-the-wild: Where Do We Stand?, Loghmani Mohammad Reza, Caputo Barbara and Vincze Markus, ICRA 18

“On the Challenges of Open World Recognition under Shifting Visual Domains”

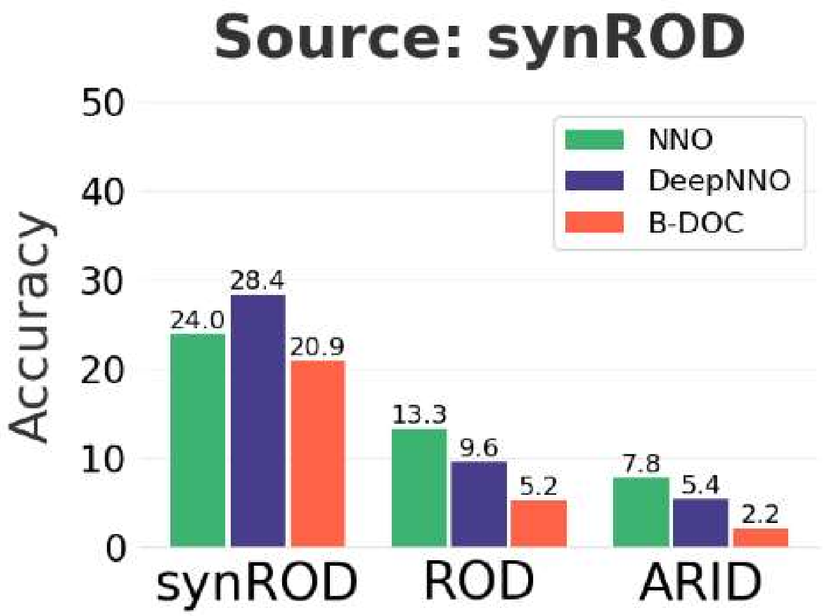
Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Barbara Caputo

Are OWR models Robust to Domain Shift?

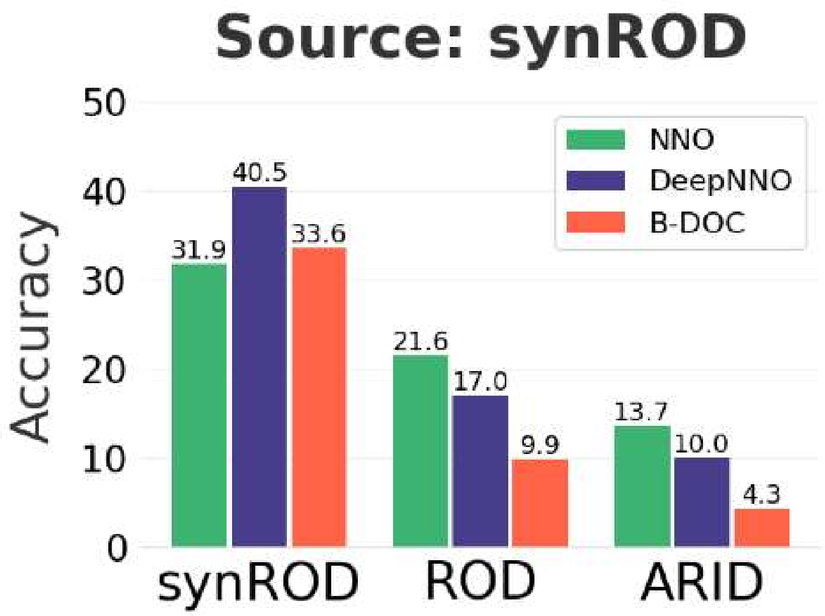
Synthetic-to-Real



Without Unknown (no rejection)



Without Unknown (with rejection)



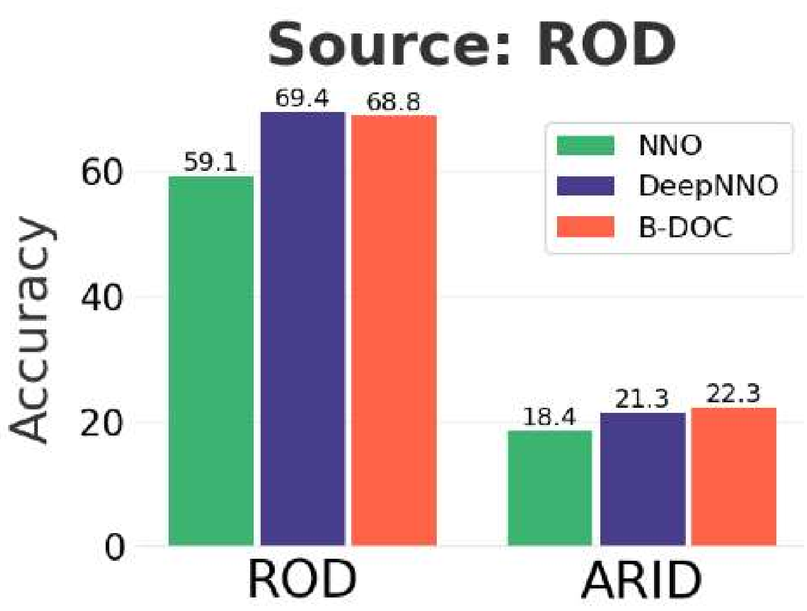
Open World Recognition (Harm)

“On the Challenges of Open World Recognition under Shifting Visual Domains”

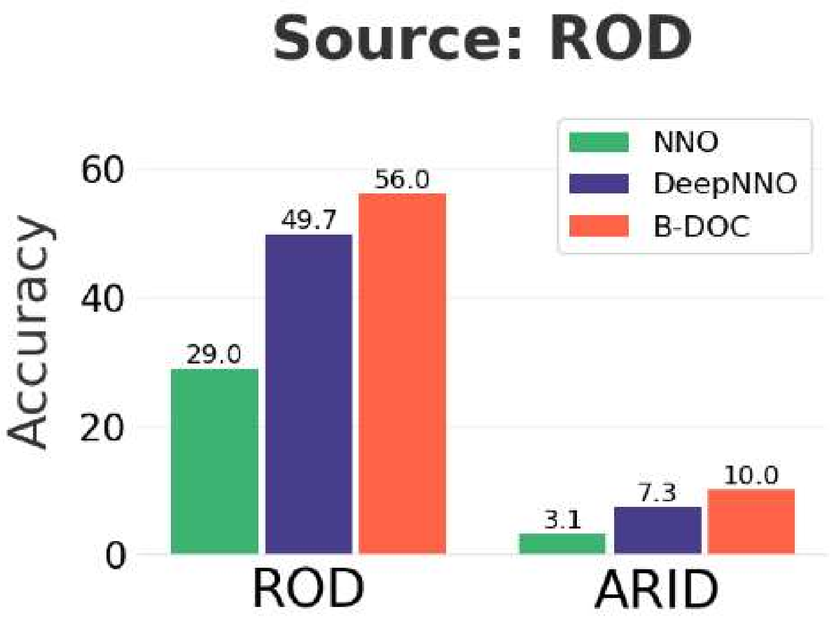
Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Barbara Caputo

Are OWR models Robust to Domain Shift?

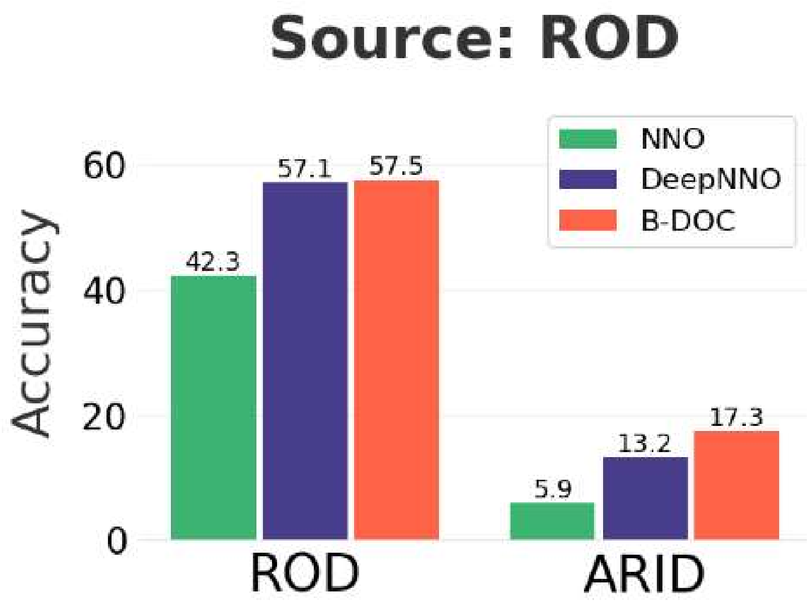
Constrained-to-Unconstrained



Without Unknown (no rejection)



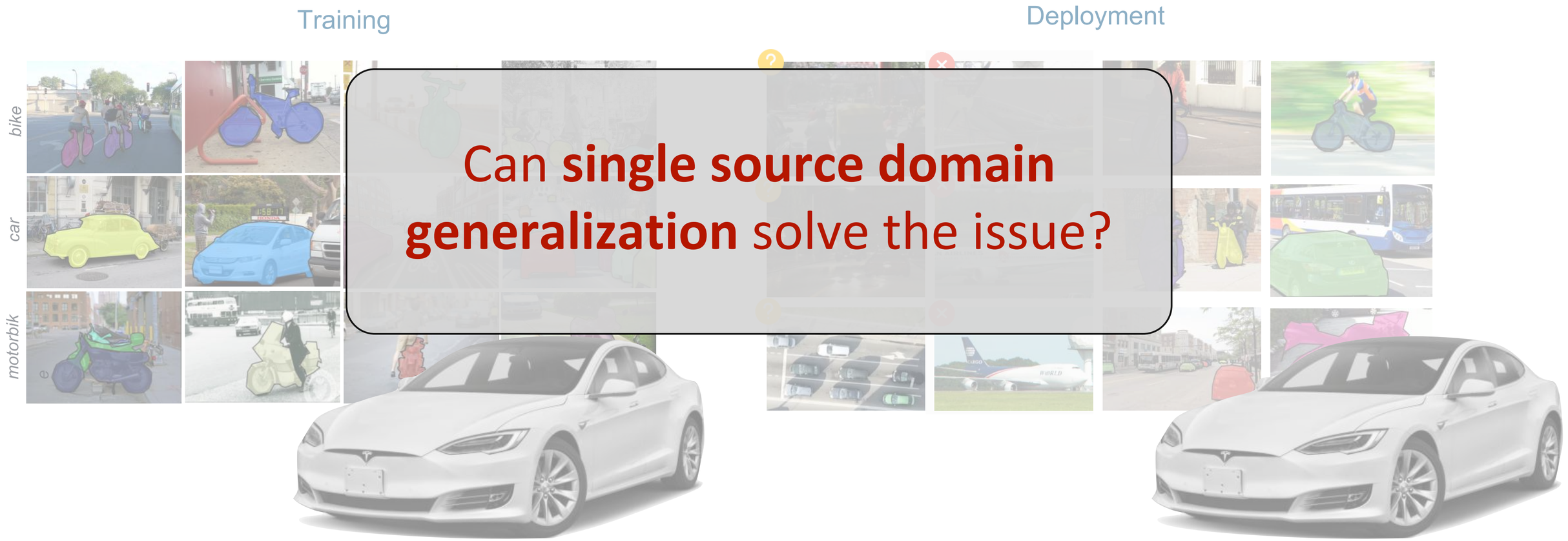
Without Unknown (with rejection)



Open World Recognition (Harm)

Preliminaries

Closed domain assumption



Can **single source domain generalization** solve the issue?

Preliminaries

Single source domain generalization

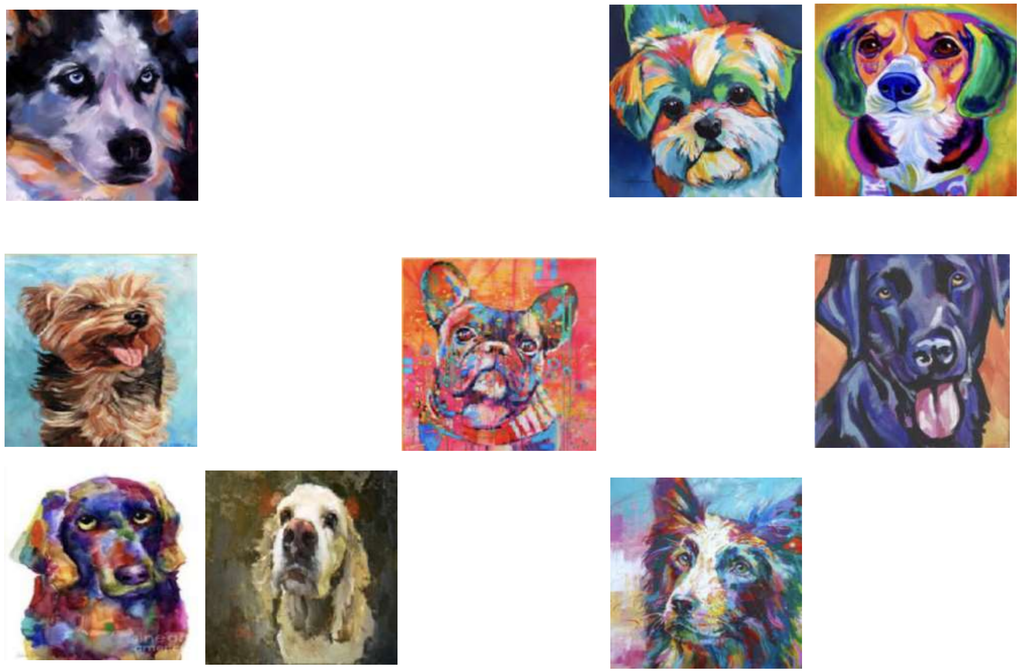
Source domain

Preliminaries

Single source domain generalization

Source domain

Target domain

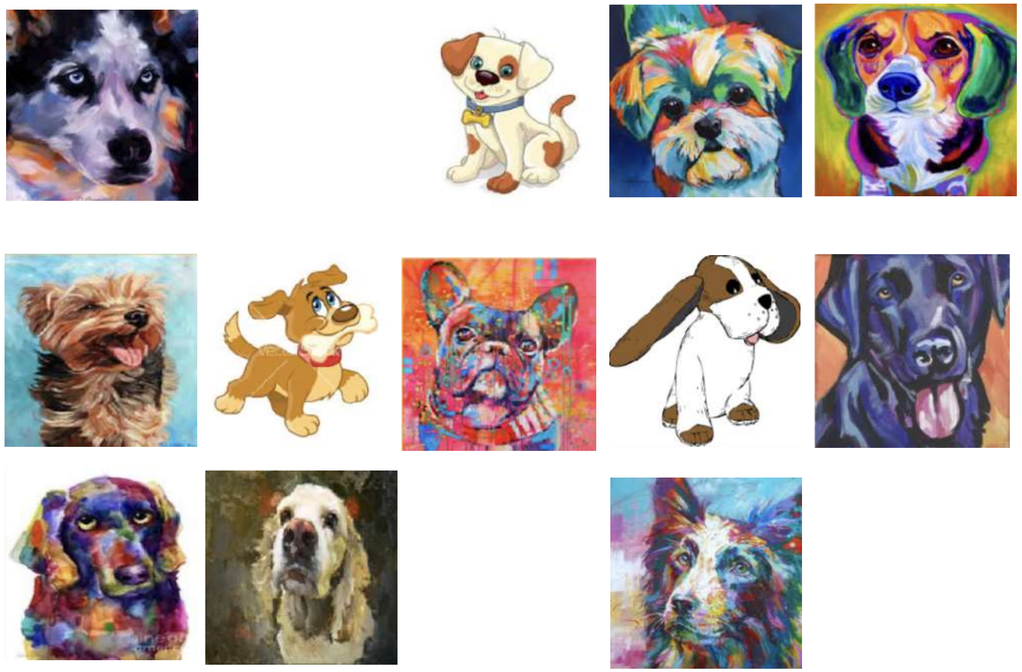


Preliminaries

Single source domain generalization

Source domain

Target domain s

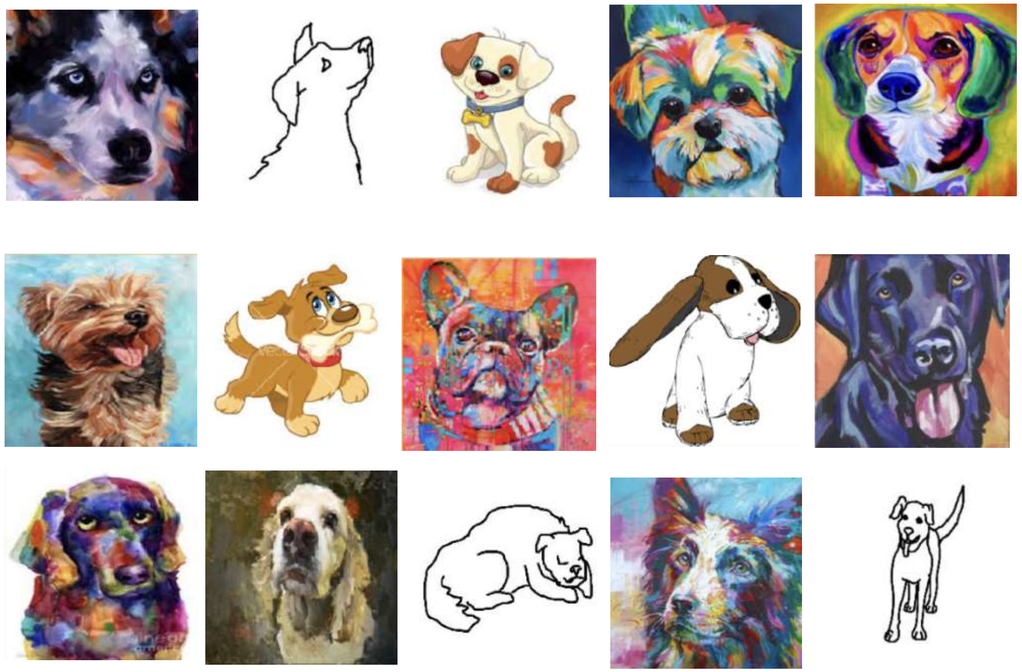


Preliminaries

Single source domain generalization

Source domain

Target domains



“On the Challenges of Open World Recognition under Shifting Visual Domains”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Barbara Caputo

Benchmark

Data augmentation with transformation sets (RSDA)

An evolutionary-based algorithm select the set of transformations that results in the worst model performance and randomly applies a subset of them.

Visual content taken from *Volpi, R., & Murino, V.. Addressing model vulnerability to distributional shifts over image transformation sets. (ICCV 19)*

“On the Challenges of Open World Recognition under Shifting Visual Domains”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Barbara Caputo

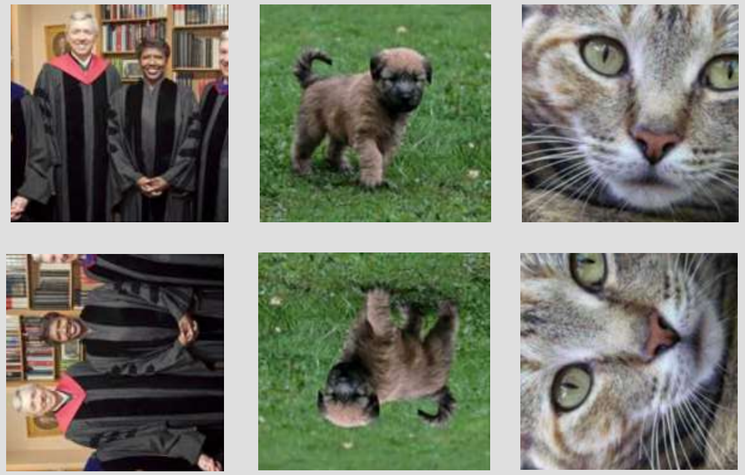
Benchmark

Data augmentation with transformation sets (RSDA)

An evolutionary-based algorithm select the set of transformations that results in the worst model performance and randomly applies a subset of them.

Self-supervised learning with relative rotations (RR)

By employing an auxiliary self-supervised task, the model concentrates on discriminative invariances and regularities, enhancing generalization to new domains



Visual content taken from *Gidaris, S., Singh, P., & Komodakis, N.. Unsupervised representation learning by predicting image rotations. (ICLR18)*

“On the Challenges of Open World Recognition under Shifting Visual Domains”

Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Barbara Caputo

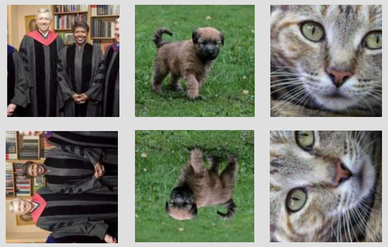
Benchmark

Data augmentation with transformation sets (RSDA)

An evolutionary-based algorithm select the set of transformations that results in the worst model performance and randomly applies a subset of them.

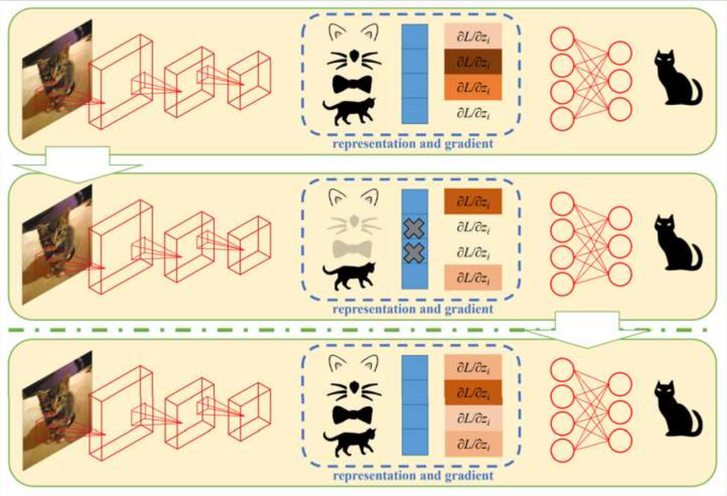
Self-supervised learning with relative rotations (RR)

By employing an auxiliary self-supervised task, the model concentrates on discriminative invariances and regularities, enhancing generalization to new domains.



Regularization through self-challenging (SC)

The model learns to classify features that have been corrupted by removing the elements that mostly contributed to a correct classification of the corresponding sample



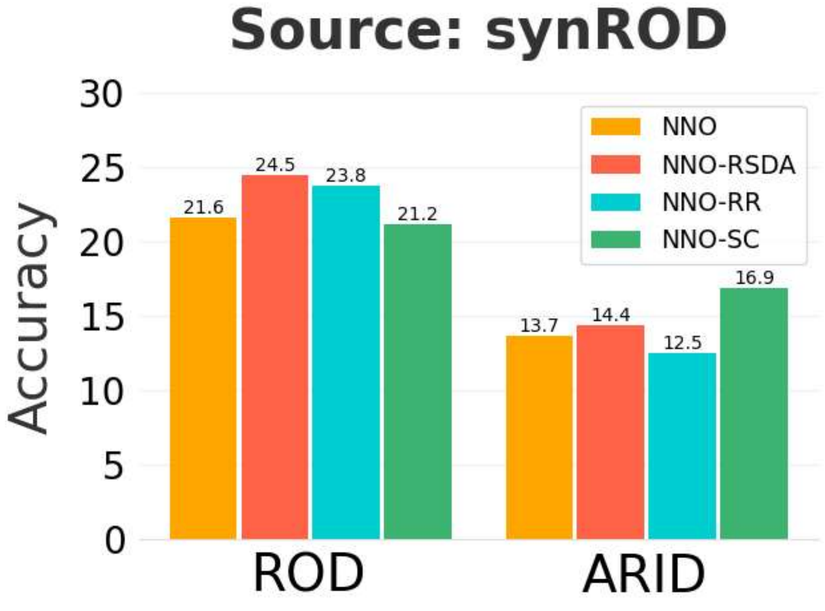
Visual content taken from Huang, Z., Wang, H., Xing, E. P., & Huang, D. Self-Challenging Improves Cross-Domain Generalization. (ECCV20)

“On the Challenges of Open World Recognition under Shifting Visual Domains”

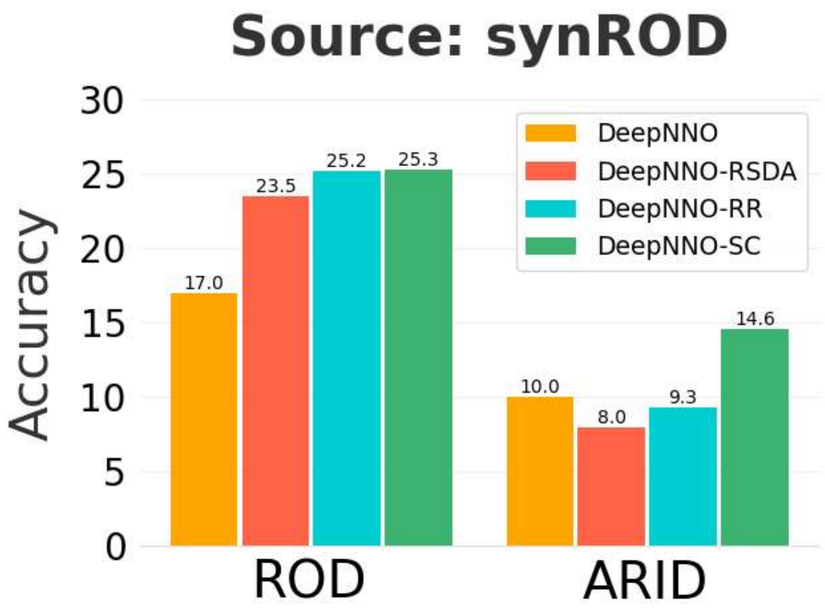
Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Barbara Caputo

Can DG methods address the problem?

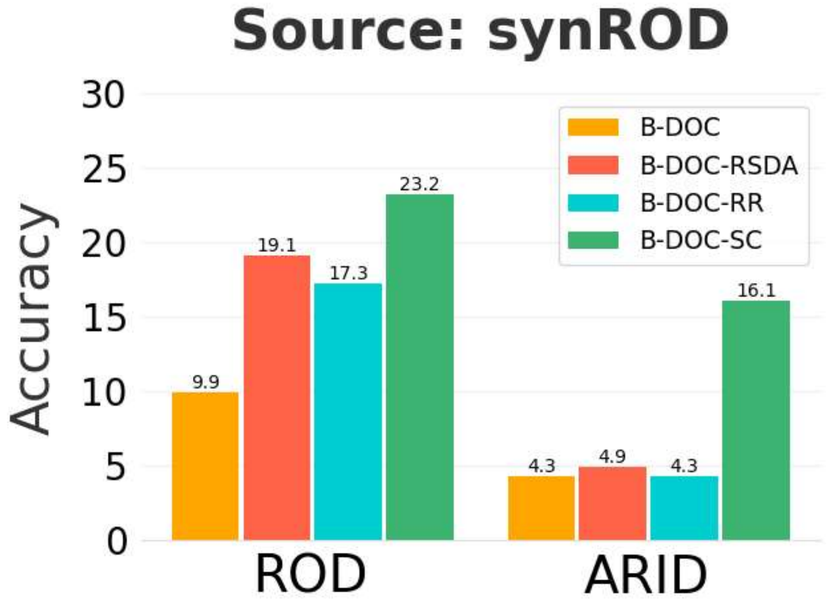
Synthetic-to-Real



Without Unknown (No rejection)



Without Unknown (with rejection)



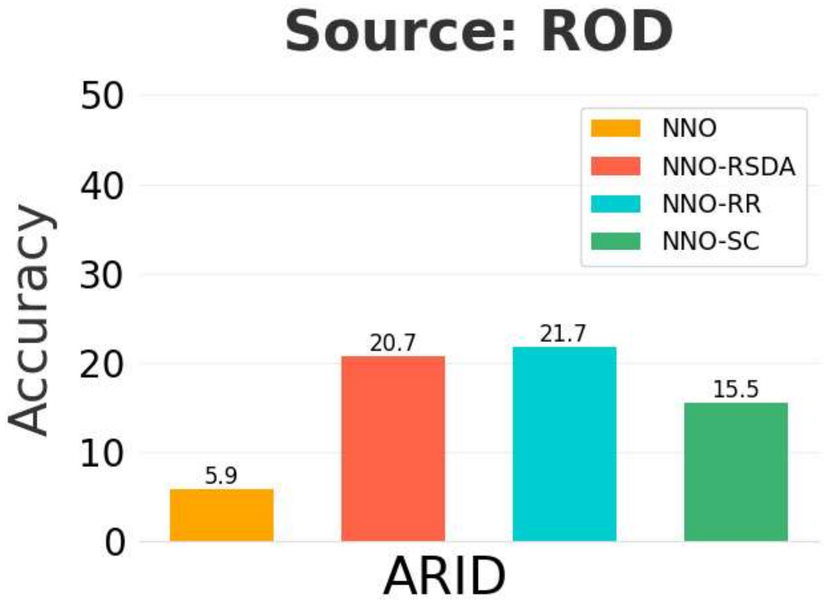
Open World Recognition (Harm)

“On the Challenges of Open World Recognition under Shifting Visual Domains”

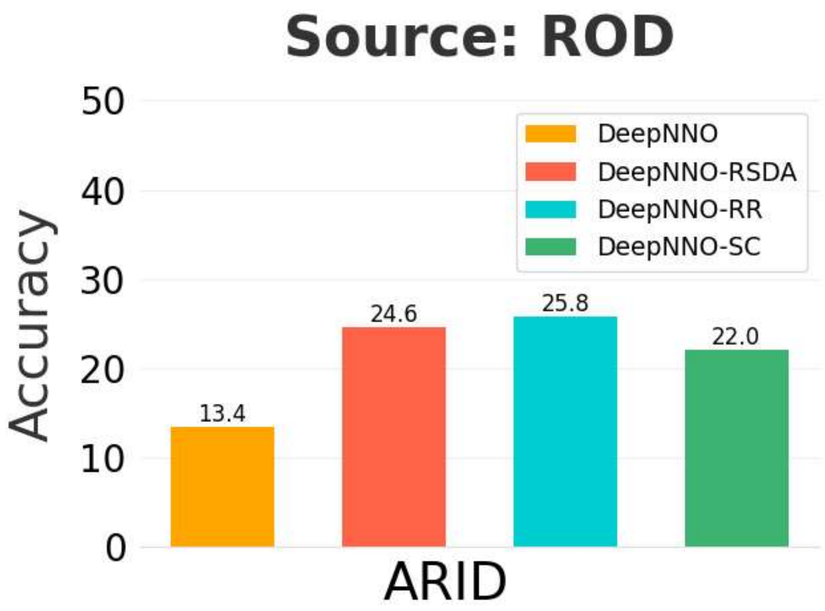
Dario Fontanel, Fabio Cermelli, Massimiliano Mancini, Barbara Caputo

Can DG methods address the problem?

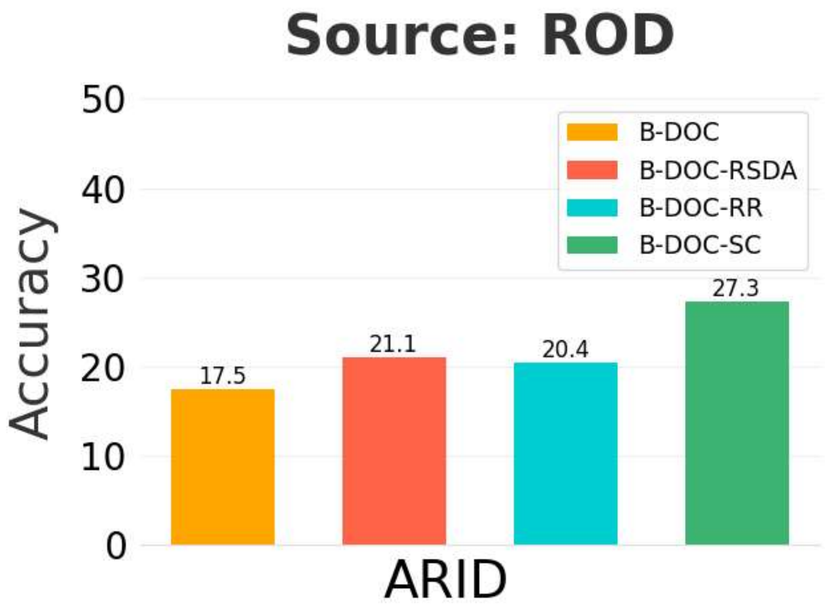
Constrained-to-Unconstrained



Without Unknown (No rejection)



Without Unknown (with rejection)



Open World Recognition (Harm)

Conclusion

Recognizing New Visual Categories in Real Open World Environments

Detection of the unknown

We introduced PAnS that learns class-specific prototypes overcoming the limitations of the softmax operation in anomaly segmentation.

Incremental learning

We introduced WILSON that learns to generate pseudo labels to train a segmentation network from image level labels only, reducing annotation costs.

Unified framework

We introduced B-DOC that adopts a global-to-clustering training loss objective, and a learnable rejection threshold per each class.

Impact of domain shift

We introduced the first benchmark to assess OWR methods under shifting visual domains and laid the foundations for further research.

Future works

Recognizing New Visual Categories in Real Open World Environments

Open World Recognition methods

OWR approaches need to collect and label a new set of images every time a new class is discovered, usually involving a human in the loop.

Future works

Recognizing New Visual Categories in Real Open World Environments

Open World Recognition methods

OWR approaches need to collect and label a new set of images every time a new class is discovered, usually involving a human in the loop.

- *Adopt an active learning pipeline to identify the most meaningful data to be labelled*

Future works

Recognizing New Visual Categories in Real Open World Environments

Open World Recognition methods

OWR approaches need to collect and label a new set of images every time a new class is discovered, usually involving a human in the loop.

- *Adopt an active learning pipeline to identify the most meaningful data to be labelled*
- *Investigate how to generate pseudo-labels to avoid the labelling process*

Future works

Recognizing New Visual Categories in Real Open World Environments

Open World Recognition methods

OWR approaches need to collect and label a new set of images every time a new class is discovered, usually involving a human in the loop.

- *Adopt an active learning pipeline to identify the most meaningful data to be labelled*
- *Investigate how to generate pseudo-labels to avoid the labelling process*
- *Adopt a few shot approach that requires limited amounts of labelled data*

Politecnico di Torino

ellis
European Laboratory for Learning and Intelligent Systems

Thank you for the attention!