EPFL

Unsupervised Test-Time Domain adaptation

Ismail Nejjar

March 2022

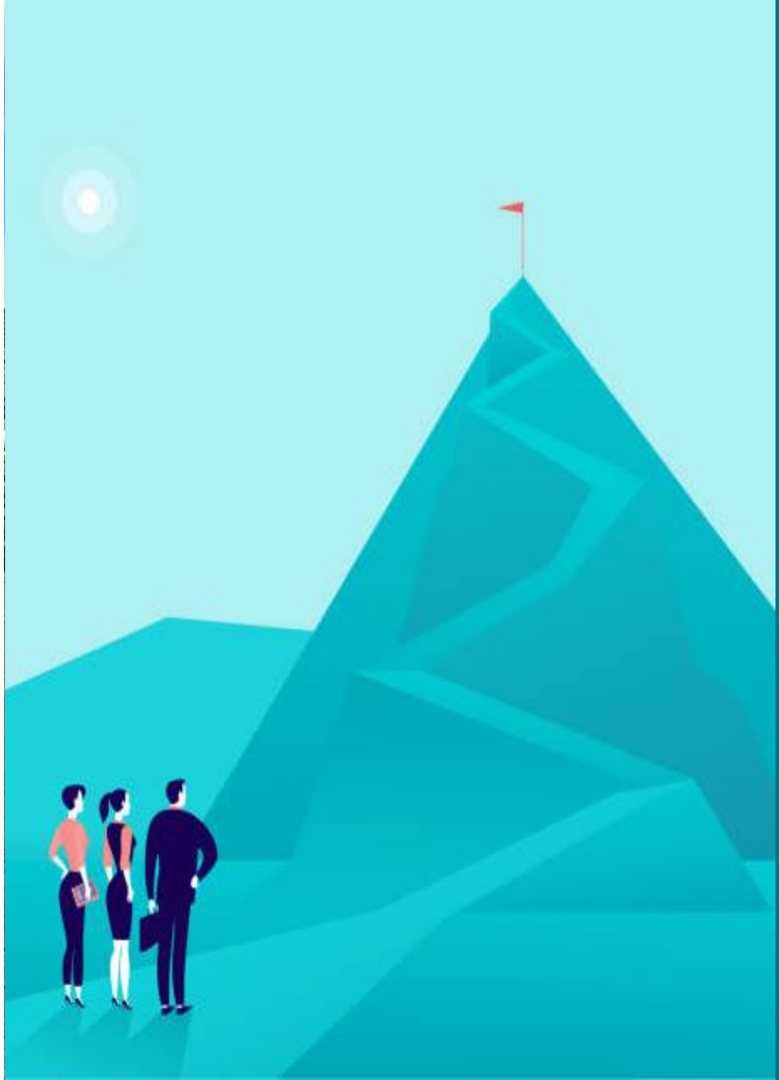
École polytechnique fédérale de Lausanne

Unsupervised Test-Time Domain adaptation

Different from the traditional domain adaptation.

Test-time adaptation aims to adapt a model to **changing conditions**

by updating to new and different data during testing without altering training or requiring more supervision.



OUTLINE

The Purpose

The Need

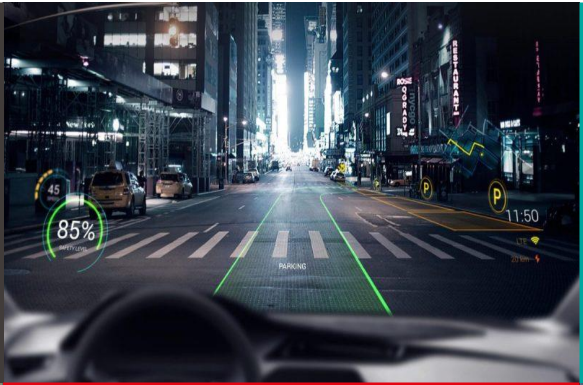
Use case Road crack detection

Ismail Nejjar

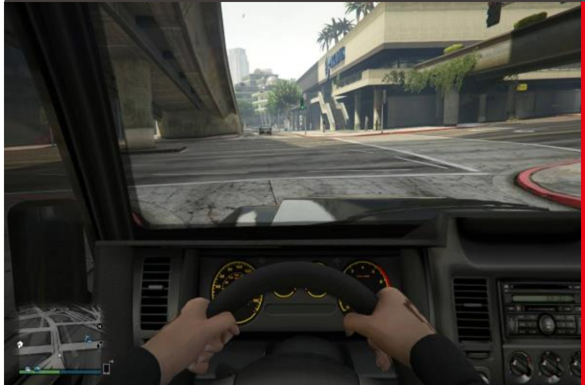
Test-Time Adaptation : The Purpose (1)

Adaptation to Different Shifts : How to reduce generalization error on simulation-to-real discrepancies

Game



Real world



Simulator

Unsupervised Test-Time Domain adaptation

Test-Time Adaptation : The Purpose (2)

Adaptation to Different Shifts : How to reduce generalization error on simulation-to-real discrepancies, different operating conditions

Single

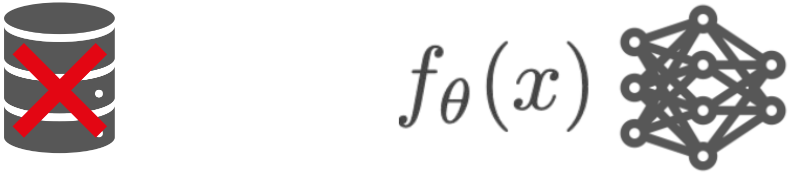
Off Shore

On Shore

Test-Time Adaptation : The Purpose (3)

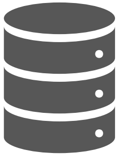
Adaptation to Different Shifts : How to reduce generalization error on simulation-to-real discrepancies, different operating conditions and other shifts ?

Test-Time Adaptation : How to improve during testing without relying on training data, using soly the pre-trained model and target data.



Training data
‘source’

Model



Test data
‘target’

OUTLINE

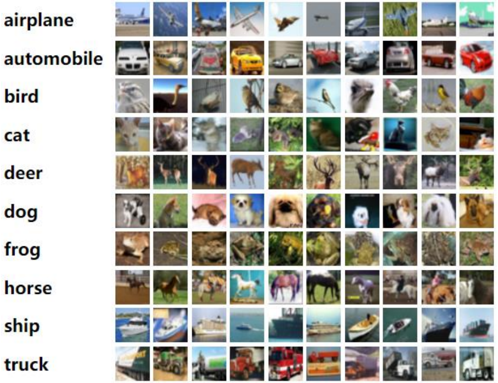
The Purpose

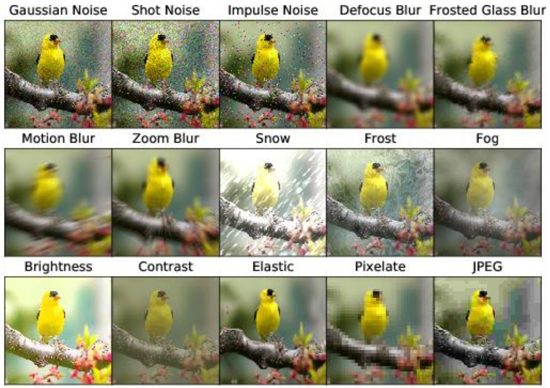
The Need

Use case Road crack detection

Test-Time Adaptation : The Need (1)

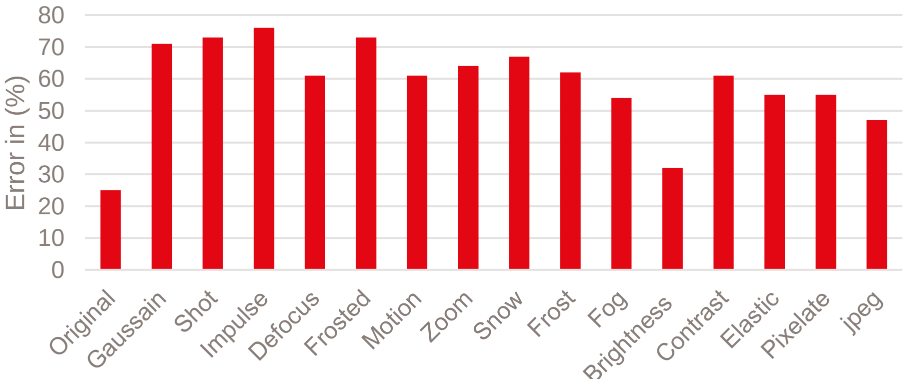
How to reduce generalization error on new and different example at test time ?





Original Cifar 10

Cifar 10-C



Test-Time Adaptation : The Need (2)

How to reduce generalization error on new and different example at test time ?

Model Performance: A model could perform poorly without adaption

Unavailability: the source data are unavailable for privacy reasons or limited bandwidth.

Efficiency: It may not be practical to process source data during testing.

OUTLINE

The Purpose

The Need

Use case Road Crack detection

- **Road Damage Dataset**
- Previous Method
- Current Approach
- Results

Road Damage Dataset 2020 (1)

Road Damage Dataset 2020 (RDD2020) is a real-world dataset of road damage detection collected using a smartphone **under different conditions** in **3 countries** :

- Japan
- India
- Czech

State Highways

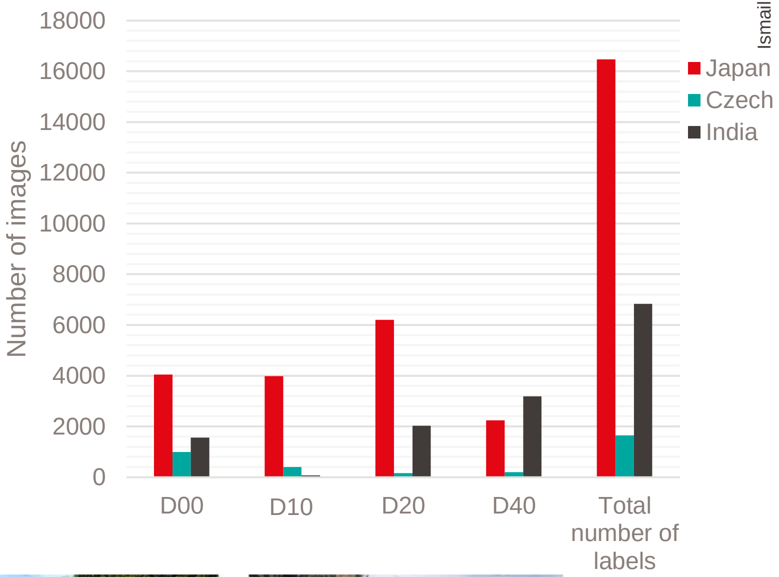
Local Road

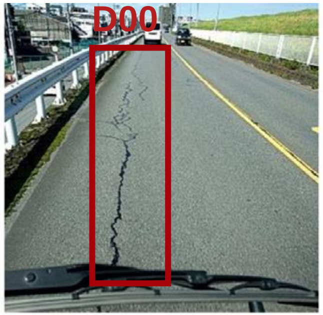
Urban Road

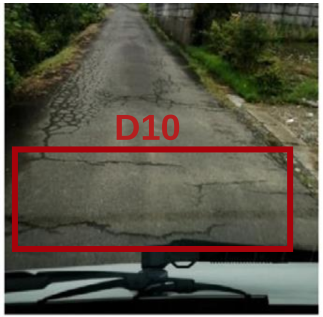
Road Damage Dataset 2020 (2)

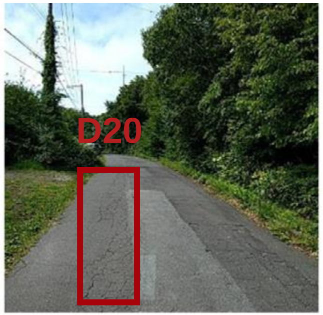
There are 4 types of Road Damage Types denoted as :

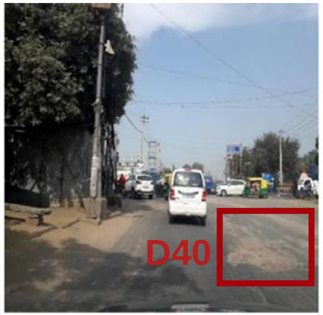
- D00 : Longitudinal linear crack
- D10 : Lateral linear crack
- D20 : Alligator crack
- D40 : Other corruption (ex. bump, pothole)









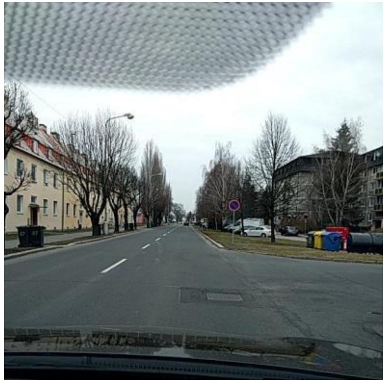


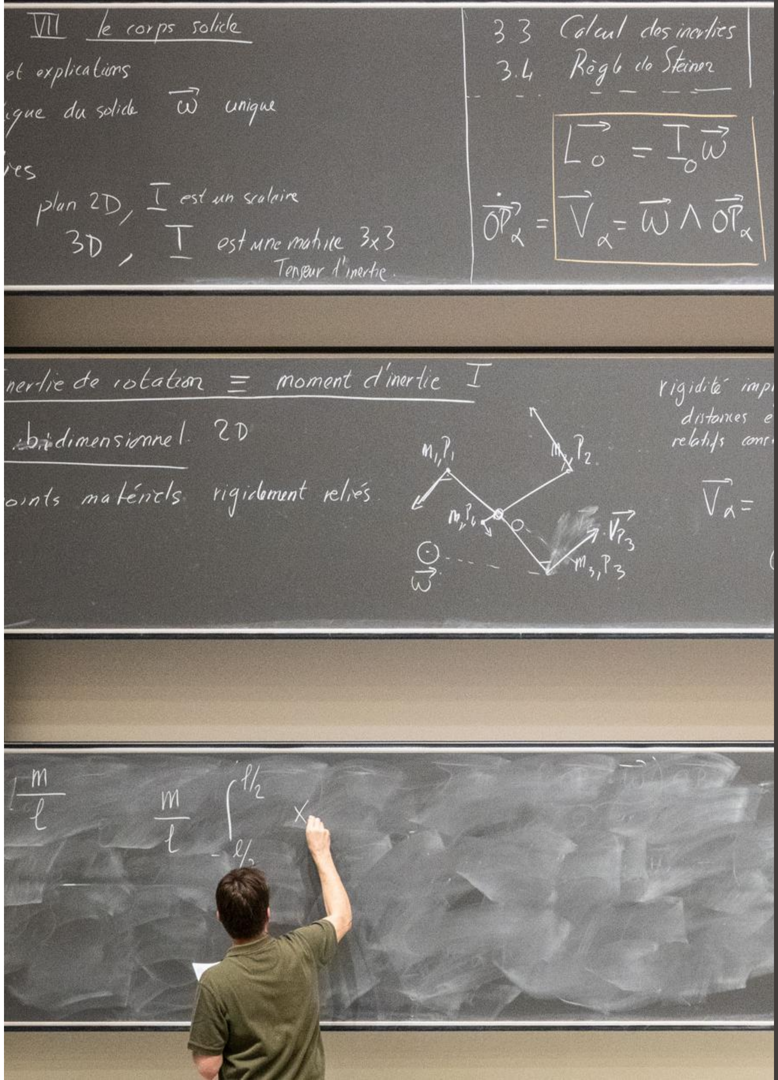
Road Damage Dataset 2020 (3)

Japan

India

Czech Republic




VII le corps solide
et explications
3.3 Calcul des inerties
3.4 Règle de Steiner
plan 2D, I est un scalaire
3D, I est une matrice 3x3
Tenseur d'inertie
$\vec{L}_O = I_O \vec{\omega}$
$\vec{V}_\alpha = \vec{\omega} \wedge \vec{OP}_\alpha$
inertie de rotation ≡ moment d'inertie I
bidimensionnel 2D
points matériels rigidement reliés
rigidité
distances
relatifs

O U T L I N E

The Purpose

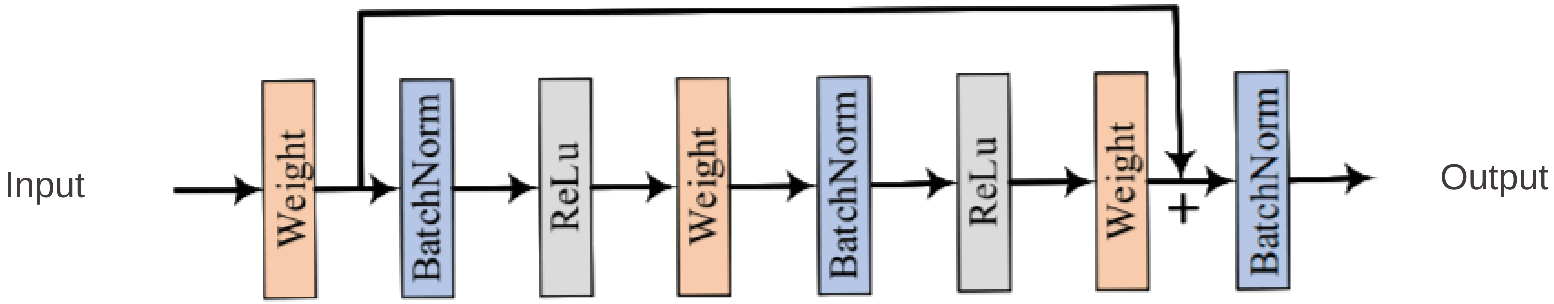
The Need

Use case Road Crack detection

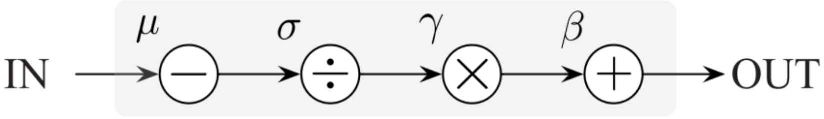
- Road Damage Dataset
- **Previous Method**
- Current Approach
- Results

Prerequisite

The classical Resnet architecture is structured as follow :



The batch normalization layers are illustrated below:



Defined by :

- Standardization part : $\mu \leftarrow \mathbb{E}[x\ ], \sigma^2 \leftarrow \mathbb{E}[(\mu - x\ )^2]$
- Linear transformation: γ and β are learnable parameters

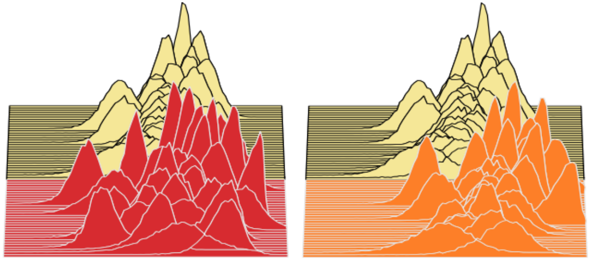
Adaptive batch normalization (AdaBN)

AdaBN assumes that model performance **deterioration** on **the target distribution** is caused by a **distribution gap** on the **intermediate layers.**

It proposes to replace the batch normalization statistics of the source data with those of the target data :

$$\mu \leftarrow \mathbb{E}[x_t], \sigma^2 \leftarrow \mathbb{E}[(\mu - x_t)^2]$$

The distribution gap should be reduced in each layer:



(a) Source (b) BN

Advantages : AdaBN doesn’t require any additional parameters or further training. Only the batch normalization statistics are refined.

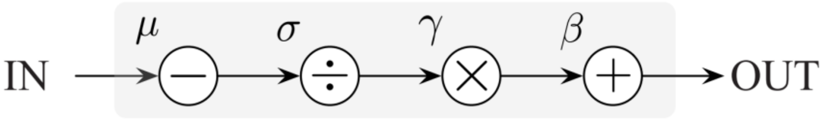
Test entropy minimization (TENT)

Tent assumes :

- Model performance **deterioration** on **the target distribution** is caused by a **distribution gap.**

- **Entropy** can serve as an estimate of the degree of shift.

It proposes to **replace the batch normalization statistics** of the source data with those of the target data, and **minimize entropy** of model predictions on the target dataset:



$$\text{normalization} \quad \mu \leftarrow \mathbb{E}[x_t], \sigma^2 \leftarrow \mathbb{E}[(\mu - x_t)^2]$$

$$\text{transformation} \quad \gamma \leftarrow \gamma + \partial H / \partial \gamma, \beta \leftarrow \beta + \partial H / \partial \beta$$

Advantages: The number of updated parameters remains relatively small compared to the total number of parameters.

Source HypOthesis Transfer (SHOT)

SHOT aims to learns a **domain specific feature encoding**, with respect to a fixed source classifier.

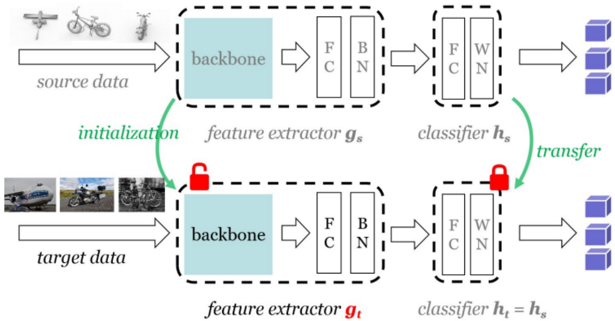
The intuition behind : is the neural network is adapted to a **source-like representations** for **target data** ⇒ that the classification outputs from the source classifier (hypothesis) for target data should be similar to those of source data.

The learning objective aims to minimize the entropy while maximizing the diversity of the predictions:

$$L_{ent}(x_t) = -\sum_c p(\hat{y}_c)\log p(\hat{y}_c)$$

$$L_{div}(x_t) = \sum_c \tilde{p}(\hat{y}_c)\log \tilde{p}(\hat{y}_c)$$

where $\tilde{p}(\hat{y}_c) = \frac{1}{B}\sum_b p(\hat{y}_c)$ and B the batch-size.



Disadvantages: The number of trained parameters is drastically bigger than the previous methods requiring more computational power at test-time.

OUTLINE

The Purpose

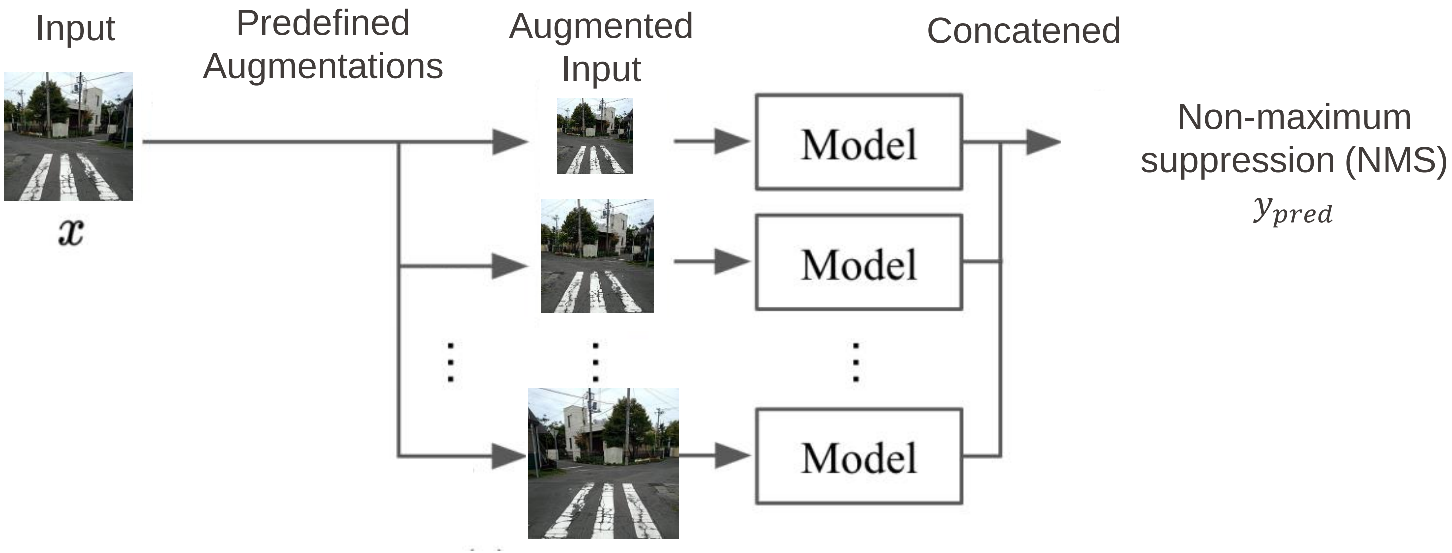
The Need

Use case Road Crack detection

- Road Damage Dataset
- Previous Method
- **Current Approach**
- Results

Proposed Method : Test-time augmentation (TTA)

Given a test image from the **target domain**, test-time augmentation first creates an augmented set of images for the given test image.



Proposed Method : Test-time augmentation (TTA)

The final test-time augmented predictions can be utilized in **2 different ways** :

1. In the inference stage, directly use the prediction for evaluation on a given pre-trained model

2. Given the predictions and their corresponding scores

   - Define a threshold to create a confident list of **pseudo-labels** :

$$y_t = \operatorname*{argmin}_{c \in C} \mathbb{1}[p(\hat{y}_c) > \theta]$$

   - The pseudo labels can be used to fine-tune the pre-trained model using the **cross-entropy loss**

$$L_{pseudo}(x_t, y_t) = -\sum_c q_c \log \delta_c(f(x_t))$$

where $\delta_c(x) = \frac{exp(x_c)}{\sum_i exp(x_i)}$ denotes the c-th element of the softmax function

OUTLINE

The Purpose

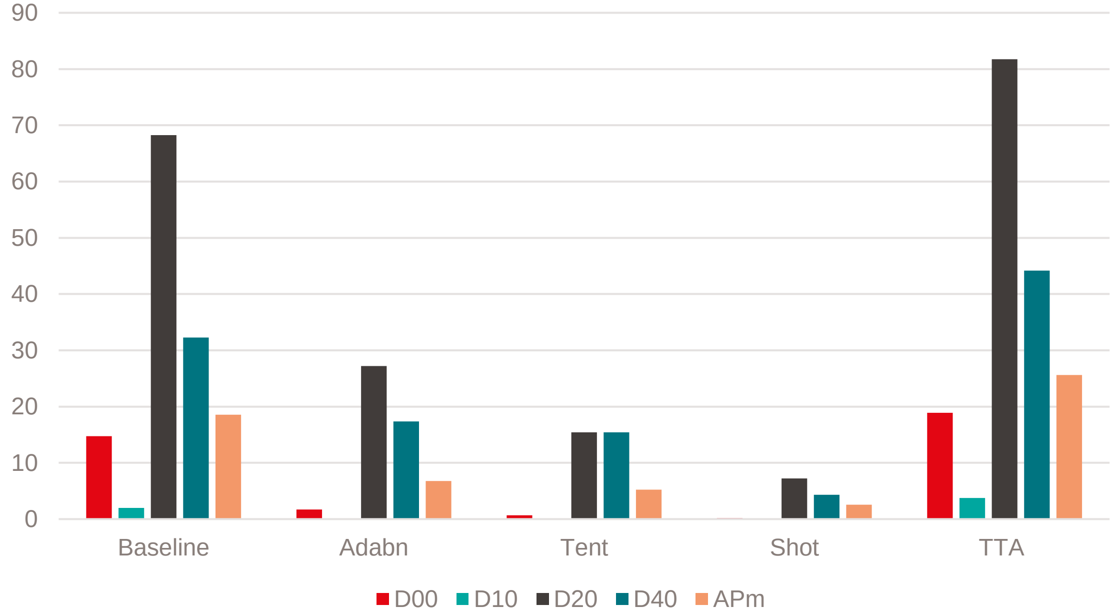
The Need

Use case Road Crack detection

- Road Damage Dataset
- Previous Method
- Current Approach
- Results

Results

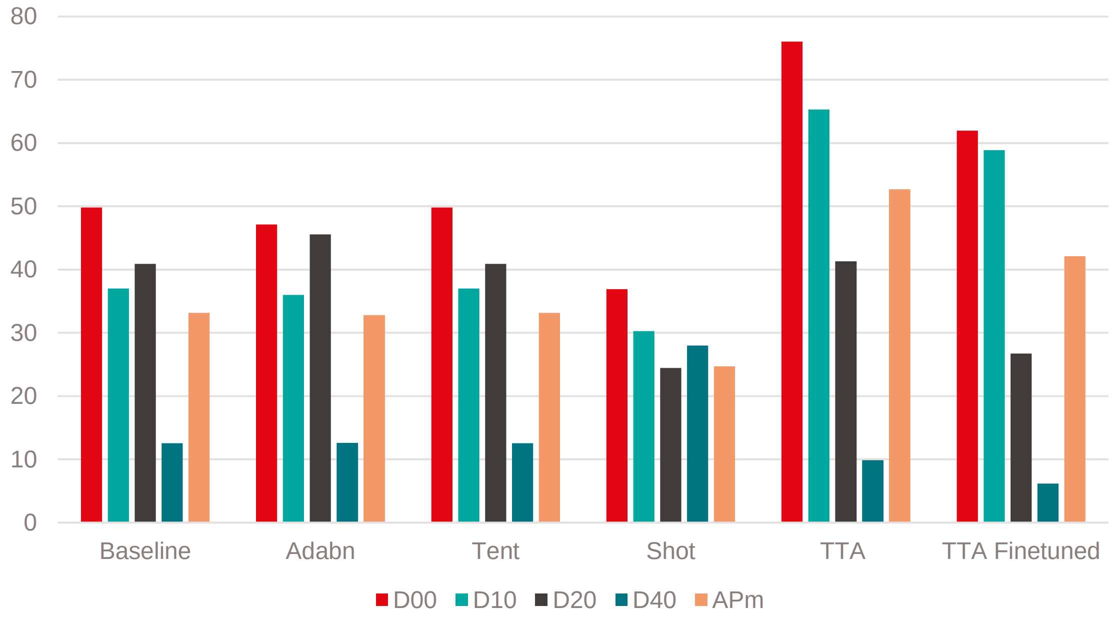
Japan ⟶ India



We report the Average Precision (AP) for each of the four damage types as well as the Mean Average Precision (mAP) in ‰

Results

Japan ⟶ Czech Republic

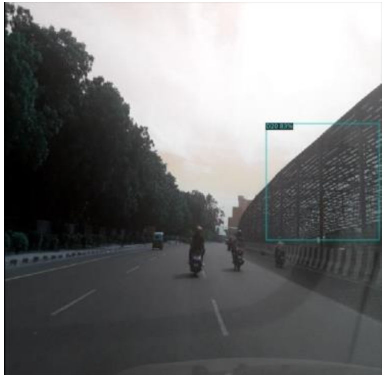


We report the Average Precision (AP) for each of the four damage types as well as the Mean Average Precision (mAP) in ‰

Possible improvement

Input

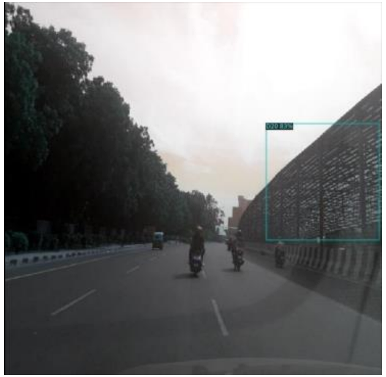
Prediction



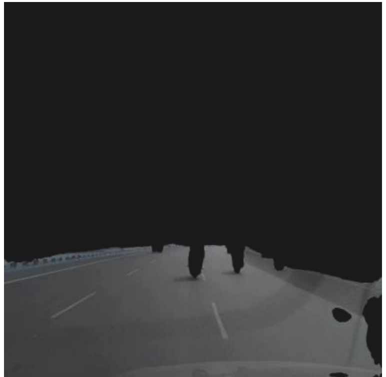
Possible improvement

Input

Prediction



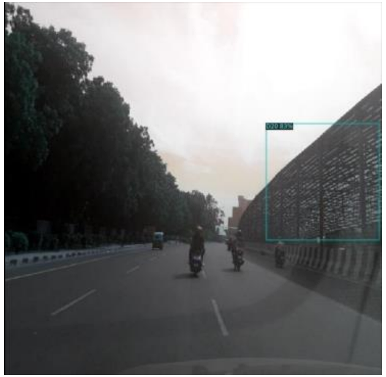
Segmented image of the road



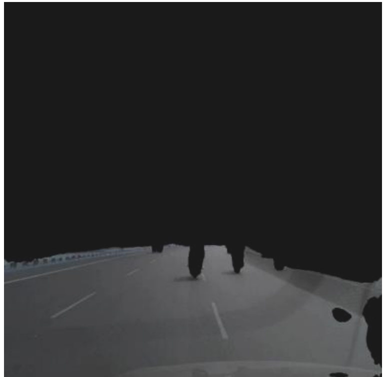
Possible improvement

Input

Prediction



Segmented image of the road



Unsupervised Test-Time Domain adaptation

Conclusion

- Test-time adaptation is a special setting of unsupervised domain adaptation where a trained model on the source domain has to adapt to the target domain without accessing source data.

- We studied the behavior of **recent test-time adaptation algorithms** in the presence of several domain shifts for crack detection

- Data augmentation can be used both for boosting the results of your model at **training** time but also at **testing** time.

- There is still room for improvement !

Merci

Ismail Nejjar

École
polytechnique
fédérale
de Lausanne