# Analysis of satellite image time series for classification and change detection

Elliot Vincent - May 27th, 2025

Committee:
Sébastien LEFEVRE (reviewer, Univ. Bretagne Sud)
Jan Dirk WEGNER (reviewer, Univ. of Zurich)
Pauline LUC (Google DeepMind)
Charlotte PELLETIER (Univ. Bretagne Sud)
Gabriele FACCIOLO (ENS Paris-Saclay)
Mathieu AUBRY (advisor, ENPC)
Jean PONCE (co-advisor, ENS-PSL/NYU)

ÉCOLE NATIONALE DES PONTS ET CHAUSSÉES
IP PARIS
Gaspard - Monge Laboratoire d'informatique
Inria

# Satellite image time series

Why do we care?

# Satellite image time series

Why do we care?

Nice to look at?

# Satellite image time series

Why do we care?

Nice to look at?

Acquired in enormous quantity every day?

# Satellite image time series

Why do we care?

Nice to look at?

Acquired in enormous quantity every day?

**Allow to describe land cover and detect changes over time**

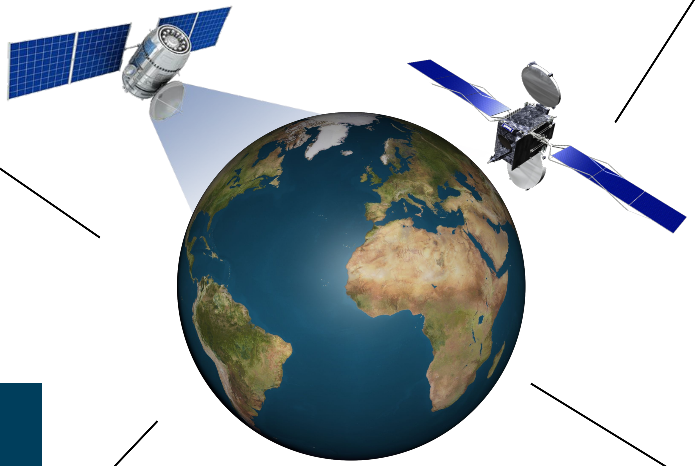
# Motivations

**Public policies**

- deforestation
- urban growth

**Agriculture**

- yield prediction
- monitoring



**Archaeology**

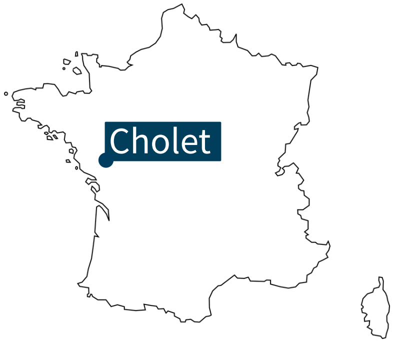
- discoveries
- monitoring

**Natural disasters**

- risk assessment
- monitoring

# Satellite image time series

A toy example




PlanetScope time series
5 images between April 2021 and August 2023
Each image ~1.2 $km^2$

# Satellite image time series

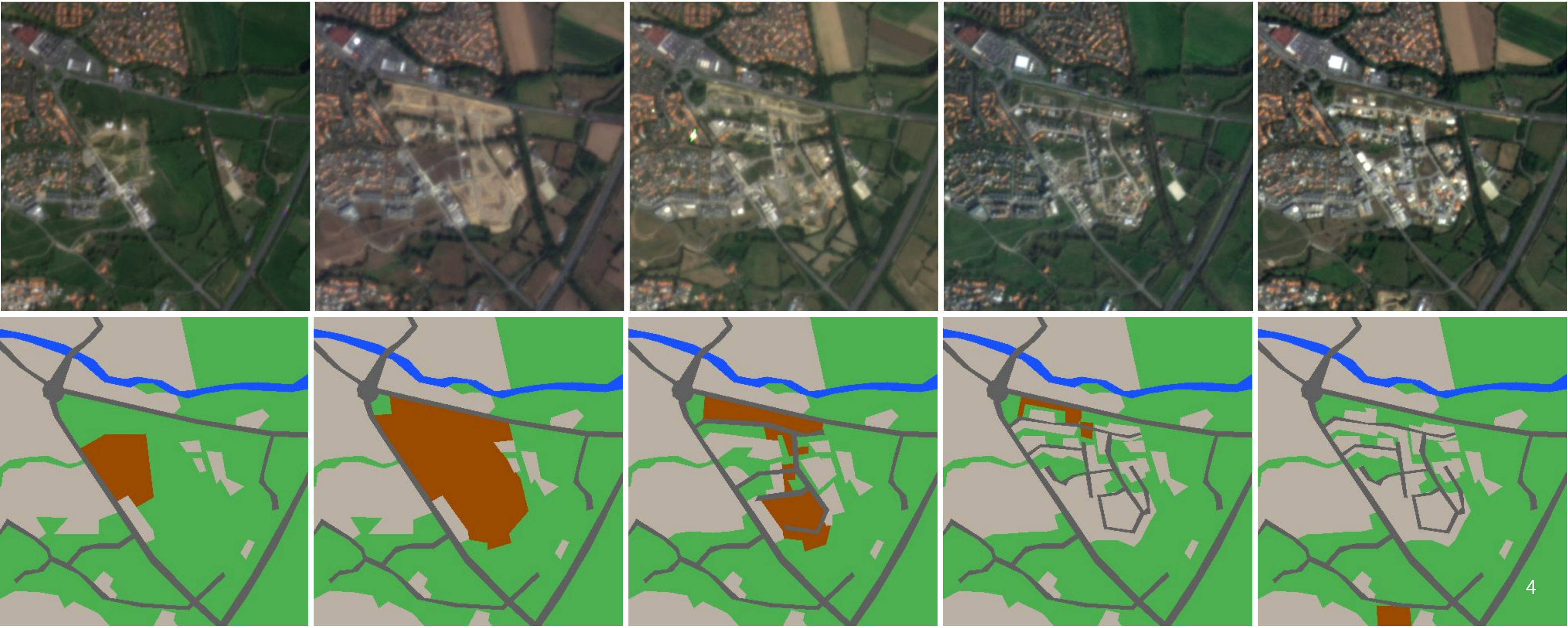
A toy example - Pixel-wise classification

# Satellite image time series

A toy example - Object-based classification

# Satellite image time series

A toy example - Semantic change detection



# Main challenge: scarcity of annotated data

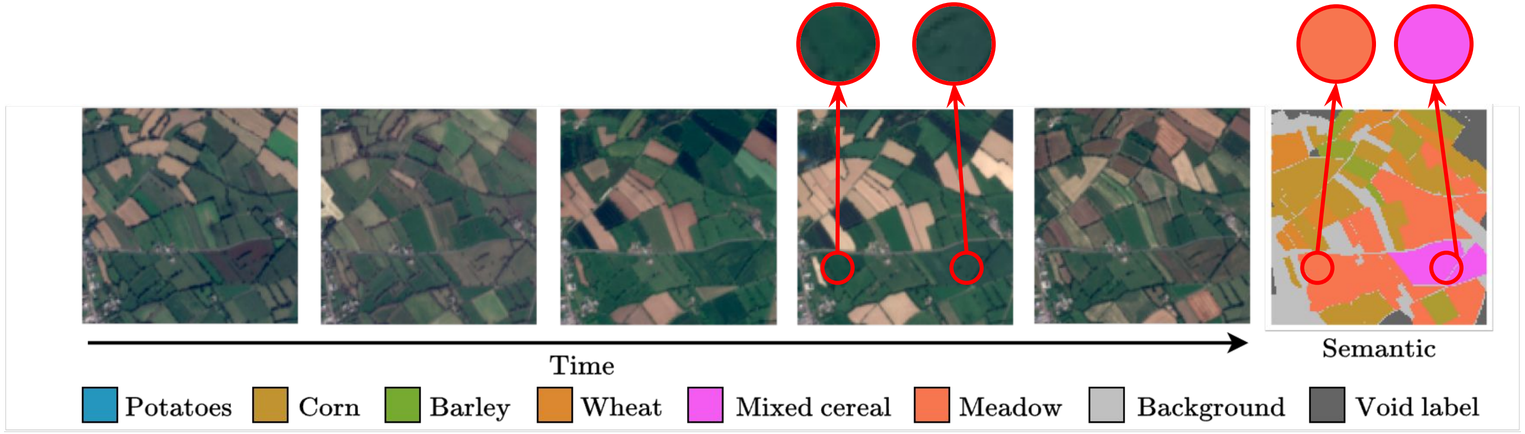
Datasets are "*small, sparse, spatio-temporally clustered, and specialized*"

E. Rolf et al. *Position: Mission Critical – Satellite Data is a Distinct Modality in Machine Learning*. ICML 2024.

# Scarcity of annotated data: why?

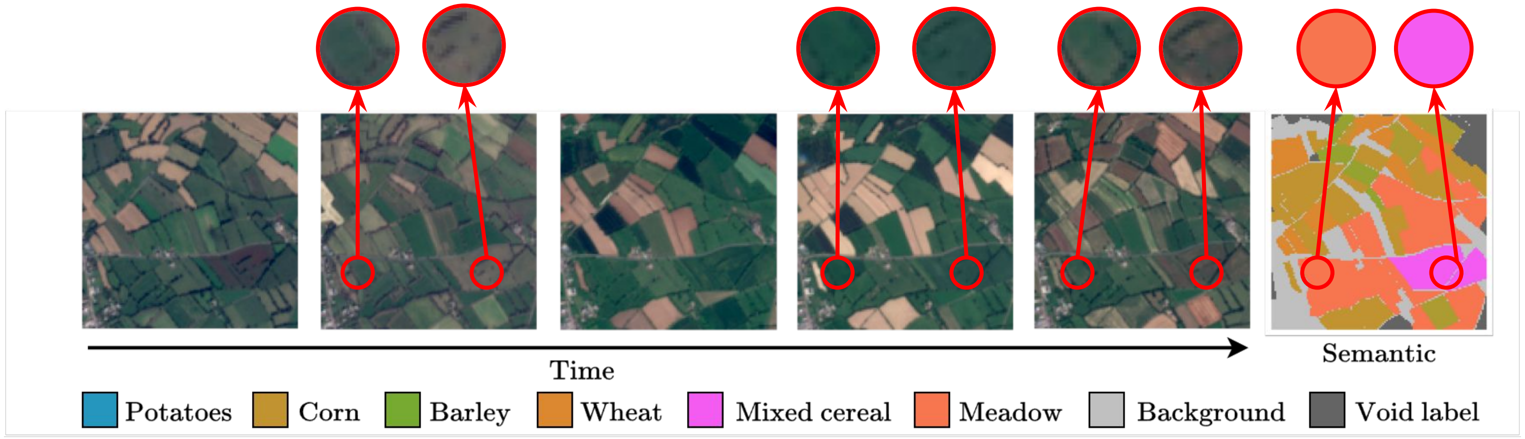
Need to annotate all pixels

# Scarcity of annotated data: why?



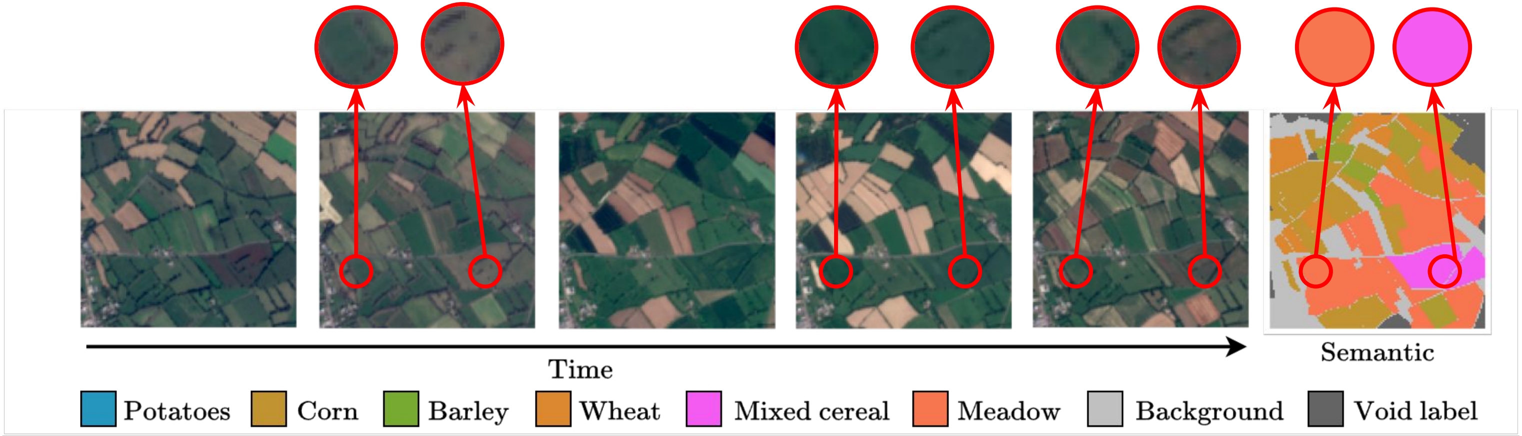

→ impossible to visually distinguish crops with a single image

# Scarcity of annotated data: why?



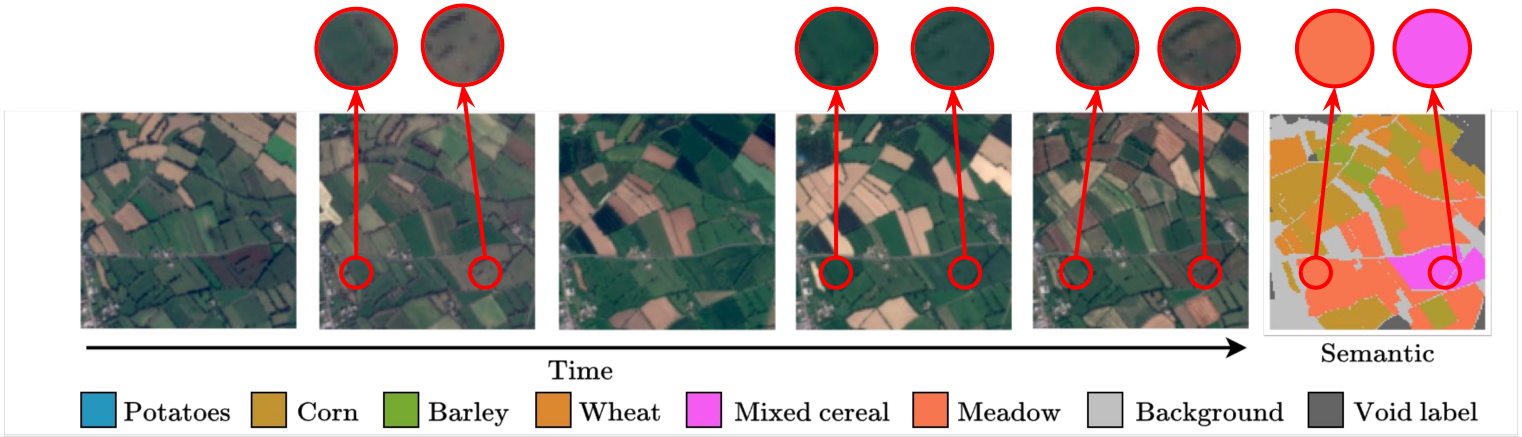

→ temporality allows to visually distinguish crop types

# Scarcity of annotated data: why?




→ **experts required to qualify crop types**
→ **external databases may be available**

# Scarcity of annotated data: why?




→ experts required to qualify crop types
→ external databases may be available
**→ costly, time-consuming**

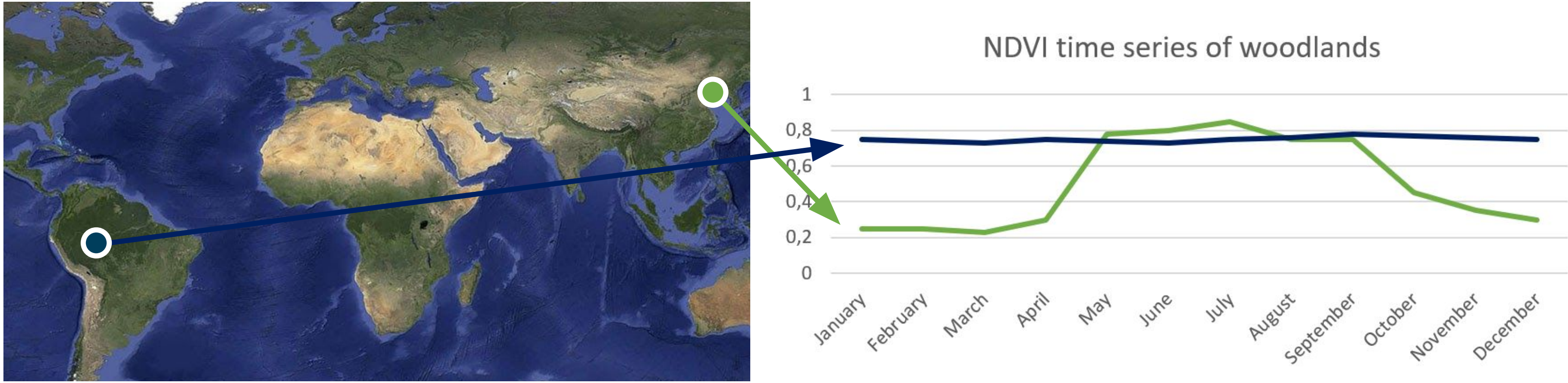
# Scarcity of annotated data: why is it an issue?

**Strong spatial and temporal domain shifts**

# Scarcity of annotated data: why is it an issue?

**Strong spatial and temporal domain shifts**

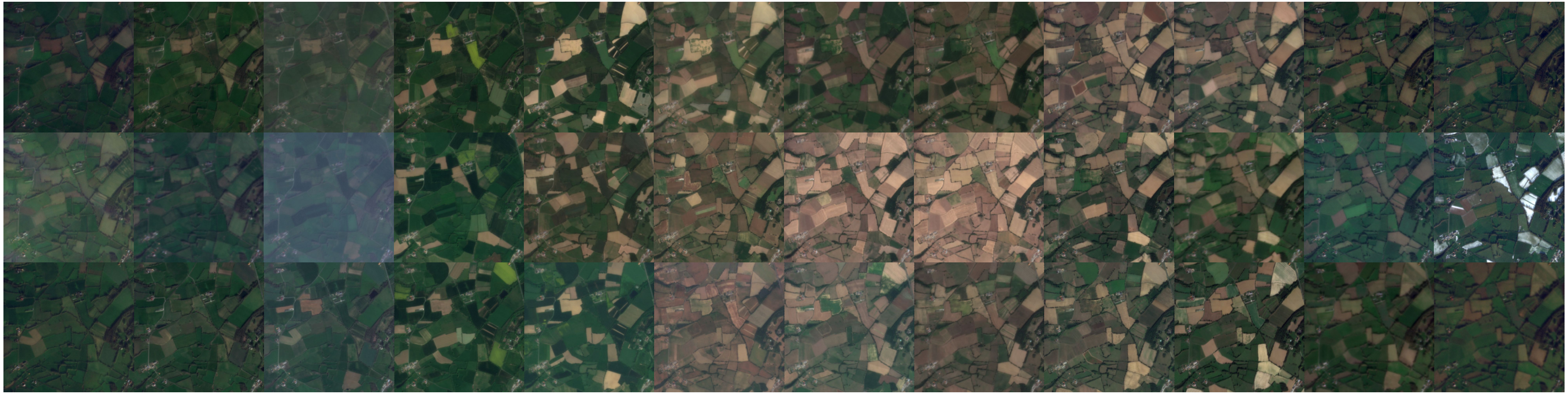
- **Spatial:** Variations due to geographical differences
  - Caused by atmospheric conditions, sensor characteristics
  - Example: Similar land covers (like different types of forests) appear differently



# Scarcity of annotated data: why is it an issue?

**Strong spatial and temporal domain shifts**

- **Spatial:** Variations due to geographical differences
- **Temporal:** Changes occurring over time
    - Caused by seasonal/weather variations, land-use changes, sensor degradation
    - Results in changing statistical properties of image data



2021
2022
2023

# Scarcity of annotated data: why is it an issue?

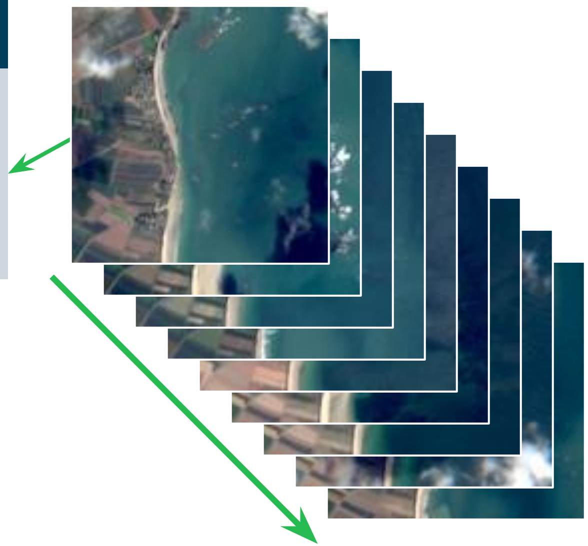
### A tridimensional data type

# Scarcity of annotated data: why is it an issue?

## A tridimensional data type

**Temporal**

- what temporal range?
- what temporal resolution?
- irregular sampling
- missing data

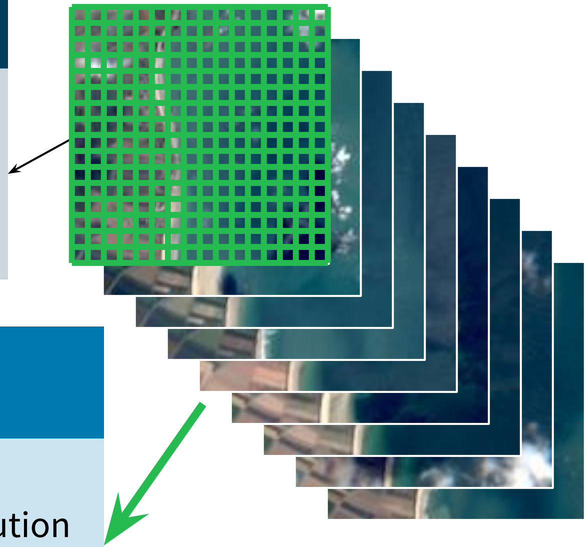


# Scarcity of annotated data: why is it an issue?

## A tridimensional data type

### Temporal

- what temporal range?
- what temporal resolution?
- irregular sampling
- missing data



### Spatial

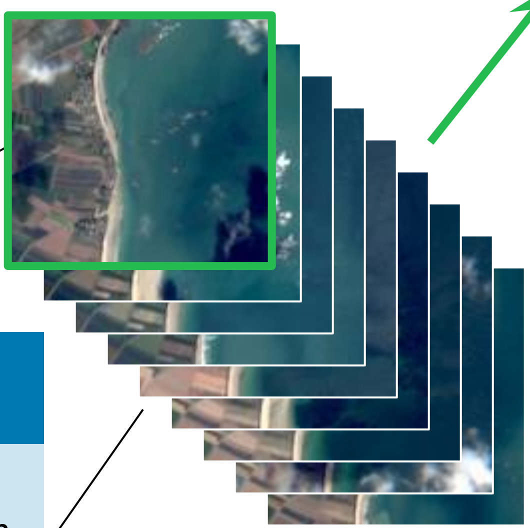
- what spatial resolution?
- trade-off with temporal resolution
- extent of a SITS often contain information irrelevant to the task

# Scarcity of annotated data: why is it an issue?

**A tridimensional data type**

### Temporal

- what temporal range?
- what temporal resolution?
- irregular sampling
- missing data

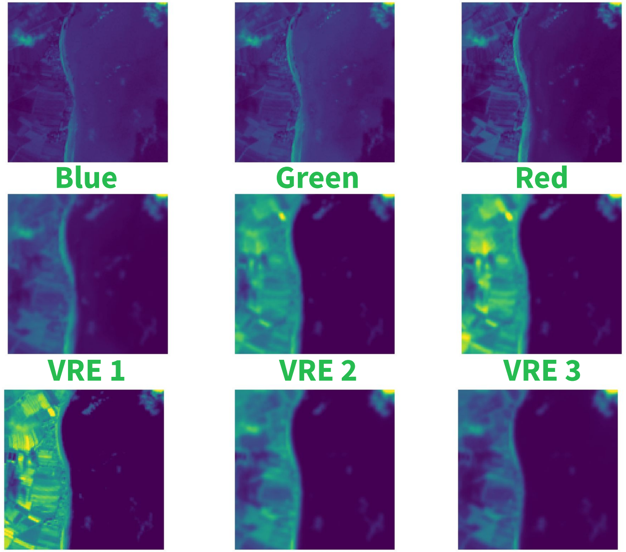


### Spatial

- what spatial resolution?
- trade-off with temporal resolution
- extent of a SITS often contain information irrelevant to the task

### Spectral

- from multi- to hyperspectral imagery
- heavy (storage, loading, ...)
- most pretrained vision model RGB only



# Contributions

**Scarcity of annotations**

- Designing unsupervised methods & methods efficient with few annotations
- Releasing annotated datasets for specific tasks and locations

Spatial/temporal shifts

- Evaluating the impact of domain shift

Tridimensional data

- Building architectures specific to SITS data

# Publications

*Pixel-wise Agricultural Image Time Series Classification: Comparison and a Deformable Prototype-based Approach*
**E. Vincent**, J. Ponce, M. Aubry – IGARSS 2025

*Satellite Image Time Series Semantic Change Detection: Novel Architecture and Analysis of Domain Shift*
**E. Vincent**, J. Ponce, M. Aubry – arXiv 2024

Best student paper award
*Detecting Looted Archaeological Sites from Satellite Image Time Series*
**E. Vincent**, M. Saroufim, J. Chemla, Y. Ubelmann, P. Marquis, J. Ponce, M. Aubry – EarthVision CVPR Workshop 2025

*CoDEx: Combining Domain Expertise for Spatial Generalization in Satellite Image Analysis*
A. Kuriyal, **E. Vincent**, M. Aubry, L. Landrieu – EarthVision CVPR Workshop 2025

*Unsupervised Layered Image Decomposition into Object Prototypes*
T. Monnier, **E. Vincent**, J. Ponce, M. Aubry – ICCV 2021

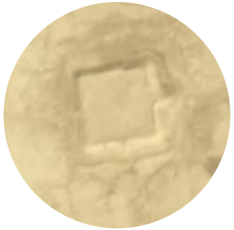
*A Model You Can Hear: Audio Identification with Playable Prototypes*
R. Loiseau, B. Bouvier, Y. Teytaut, **E. Vincent**, M. Aubry, L. Landrieu – ISMIR 2022

*Learnable Earth Parser: Discovering 3D Prototypes in Aerial Scans*
R. Loiseau, **E. Vincent**, M. Aubry, L. Landrieu – CVPR 2024

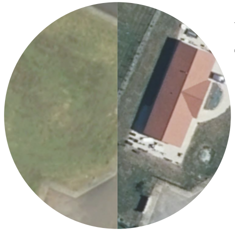
*OpenStreetView-5M: The Many Roads to Global Visual Geolocation*
G. Astruc, N. Dufour, I. Siglidis, C. Aronssohn, N. Bouia, S. Fu, R. Loiseau, V. Nguyen, C. Raude, **E. Vincent**, L. Xu, H. Zhou, L. Landrieu – CVPR 2024

*Historical Printed Ornaments: Dataset and Tasks*
S. Chaki, S. Baltaci, **E. Vincent**, R. Emonet, F. Vial-Bonacci, C. Bahier-Porte, M. Aubry, T. Fournel – ICDAR 2024

# Outline



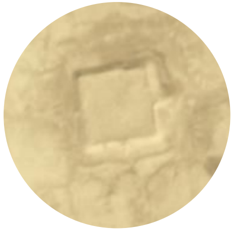
1 Afghan archaeological site
looting detection



2 Semantic change detection
and domain shift analysis

3 Crop-type classification
with few or no annotations

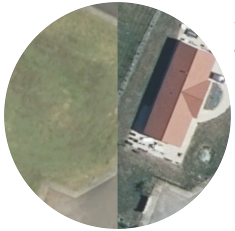
# Outline



1 Afghan archaeological site looting detection

*Detecting Looted Archaeological Sites from Satellite Image Time Series*
**E. Vincent**, M. Saroufim, J. Chemla, Y. Ubelmann, P. Marquis, J. Ponce, M. Aubry
EarthVision CVPR Workshop 2025

Best student paper award



2 Semantic change detection and domain shift analysis

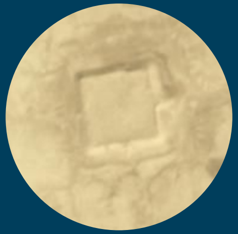
*Satellite Image Time Series Semantic Change Detection: Novel Architecture and Analysis of Domain Shift*
**E. Vincent**, J. Ponce, M. Aubry – arXiv 2024

*CoDEx: Combining Domain Expertise for Spatial Generalization in Satellite Image Analysis*
A. Kuriyal, **E. Vincent**, M. Aubry, L. Landrieu – EarthVision CVPR Workshop 2025

3 Crop-type classification with few or no annotations

*Pixel-wise Agricultural Image Time Series Classification: Comparison and a Deformable Prototype-based Approach*
**E. Vincent**, J. Ponce, M. Aubry – IGARSS 2025

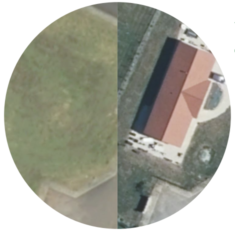
# Outline



1 Afghan archaeological site looting detection

*Detecting Looted Archaeological Sites from Satellite Image Time Series*
**E. Vincent**, M. Saroufim, J. Chemla, Y. Ubelmann, P. Marquis, J. Ponce, M. Aubry
EarthVision CVPR Workshop 2025

Best student paper award



2 Semantic change detection and domain shift analysis

*Satellite Image Time Series Semantic Change Detection: Novel Architecture and Analysis of Domain Shift*
**E. Vincent**, J. Ponce, M. Aubry – arXiv 2024

*CoDEx: Combining Domain Expertise for Spatial Generalization in Satellite Image Analysis*
A. Kuriyal, **E. Vincent**, M. Aubry, L. Landrieu – EarthVision CVPR Workshop 2025

3 Crop-type classification with few or no annotations

*Pixel-wise Agricultural Image Time Series Classification: Comparison and a Deformable Prototype-based Approach*
**E. Vincent**, J. Ponce, M. Aubry – IGARSS 2025

# Case study

- +5000 archaeological sites in Afghanistan
- First detected instance of looting at Dilberjin (DAFA, 2022)
- Ongoing looting activities
- Impossible ground surveys

→ **Need for automated monitoring processes**

AFGHANISTAN • INVESTIGATIONS

**Looting of Afghanistan archaeological site attributed to IS**

By Jacques Follorou

Published on April 8, 2023, at 7:00 am (Paris)

7 min read Lire en français

NEWS | **Dilberjin, the largest ancient urban center in the north of the country, suffered irreparable damage between 2019 and 2021. The destruction is attributable to criminal groups tied to the Islamic State organization, using methods previously seen in Syria and Iraq.**

*Le Monde* - April 8, 2023

# Monitoring archaeological sites from space

- A tool to assist archaeologist on the ground
- Several advantages (cost, rapidity, practicality)
- Rich literature on site monitoring with satellite/aerial images:
    - manually (comparison, counting)
    - automatically (change detection, detecting pits/holes)


*Casana et al., 2014 - Syria*

# Monitoring archaeological sites from space

- Deep learning methods have been evaluated, but:
    - no systematic comparison of baselines
    - often a single use case
    - very few datasets, few are released publicly

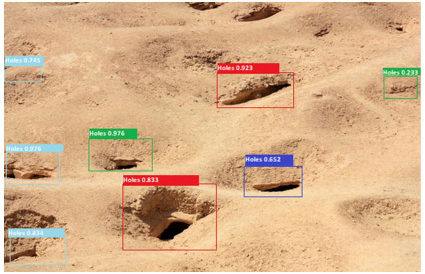
| | Open-access | Multi-temporal | Spatial resolution | Temporal resolution | Sensor | Location | Number of sites |
|---|---|---|---|---|---|---|---|
| Masini et al. (2020) | ✗ | ✓ | Varying | Yearly | Satellite | Syria | 2 |
| El Hajj (2021) | ✗ | ✗ | 15m/px | — | Satellite | Syria and Iraq | 9 |
| Payntar (2023) | ✗ | ✓ | 30m/px | Every 5 years | Satellite | Peru | 477 |
| Altaweel et al. (2024) | ✓ | ✗ | 3cm/px | — | UAV | Worldwide | 95 |

# Monitoring archaeological sites from space

- Deep learning methods have been evaluated, but:
  - no systematic comparison of baselines
  - often a single use case
  - very few datasets, few are released publicly

| | Open-access | Multi-temporal | Spatial resolution | Temporal resolution | Sensor | Location | Number of sites |
|---|---|---|---|---|---|---|---|
| Masini et al. (2020) | ✗ | ✓ | Varying | Yearly | Satellite | Syria | 2 |
| El Hajj (2021) | ✗ | ✗ | 15m/px | — | Satellite | Syria and Iraq | 9 |
| Payntar (2023) | ✗ | ✓ | 30m/px | Every 5 years | Satellite | Peru | 477 |
| Altaweel et al. (2024) | ✓ | ✗ | 3cm/px | — | UAV | Worldwide | 95 |

YOLO to detect and count looting pits on UAV (drones) images



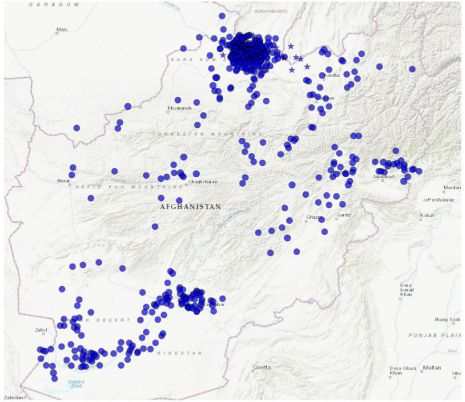
# Monitoring archaeological sites from space

- Deep learning methods have been evaluated, but:
  - no systematic comparison of baselines
  - often a single use case
  - very few datasets, few are released publicly

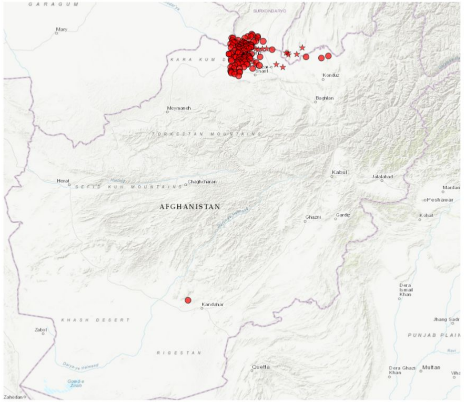
| | Open-access | Multi-temporal | Spatial resolution | Temporal resolution | Sensor | Location | Number of sites |
|---|---|---|---|---|---|---|---|
| Masini et al. (2020) | ✗ | ✓ | Varying | Yearly | Satellite | Syria | 2 |
| El Hajj (2021) | ✗ | ✗ | 15m/px | — | Satellite | Syria and Iraq | 9 |
| Payntar (2023) | ✗ | ✓ | 30m/px | Every 5 years | Satellite | Peru | 477 |
| Altaweel et al. (2024) | ✓ | ✗ | 3cm/px | — | UAV | Worldwide | 95 |
| **DAFA-LS (ours)** | ✓ | ✓ | 3.8m/px | Monthly | Satellite | Afghanistan | 675 |

# Introducing DAFA-LS

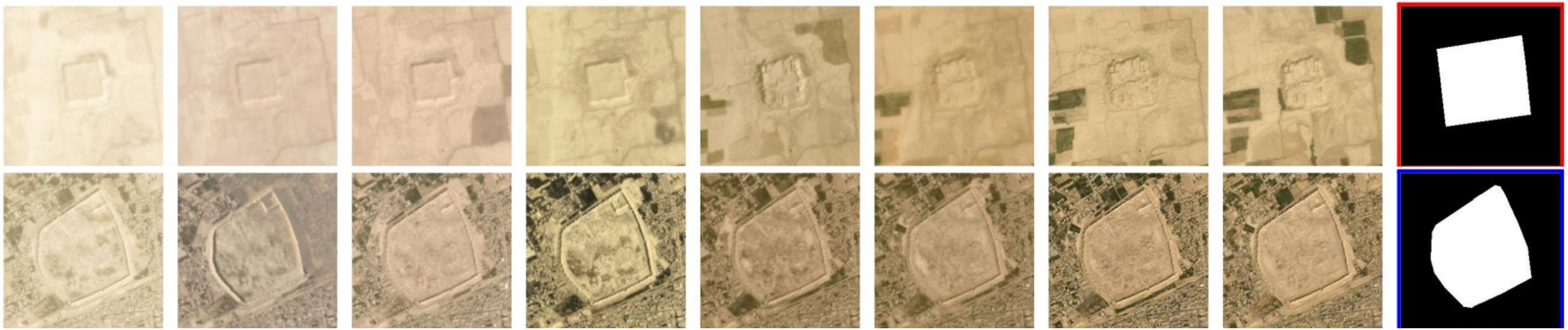
- 55,480 satellite images over 8 years (2016-2023)
- 675 archaeological sites
- 135 were looted during the period
- Monthly Planet satellite image time series (SITS)
- Preservation status + coarse location mask



(a) Map of preserved sites



(b) Map of looted sites

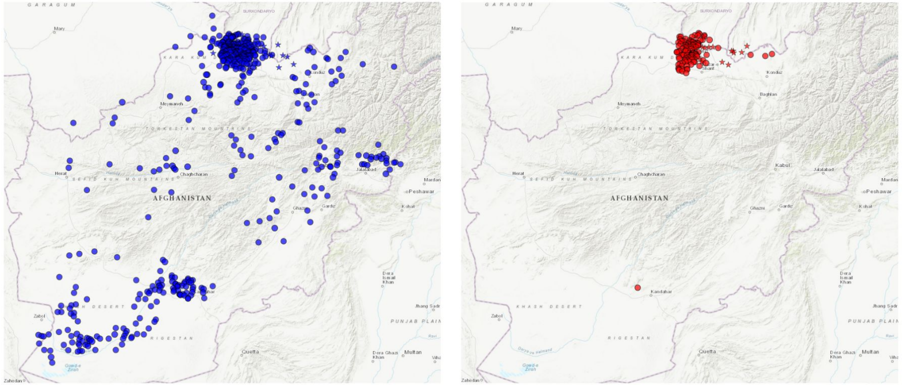


2016 2017 2018 2019 2020 2021 2022 2023 Mask

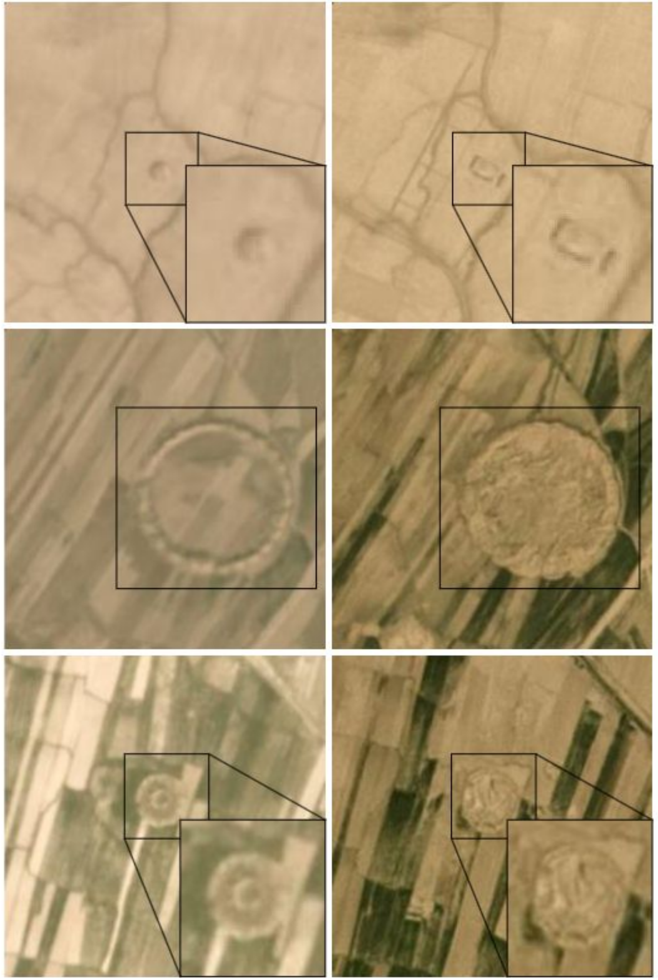
# Introducing DAFA-LS

Special care in our data splits formulation, limiting

- **surface bias:** larger sites more likely to be looted
- **geographical bias:** northern sites more likely to be looted
- + 5 spatially separated train/val folds




(a) Map of preserved sites (b) Map of looted sites



*Examples of before/after looting marks*

# Benchmarking deep learning methods

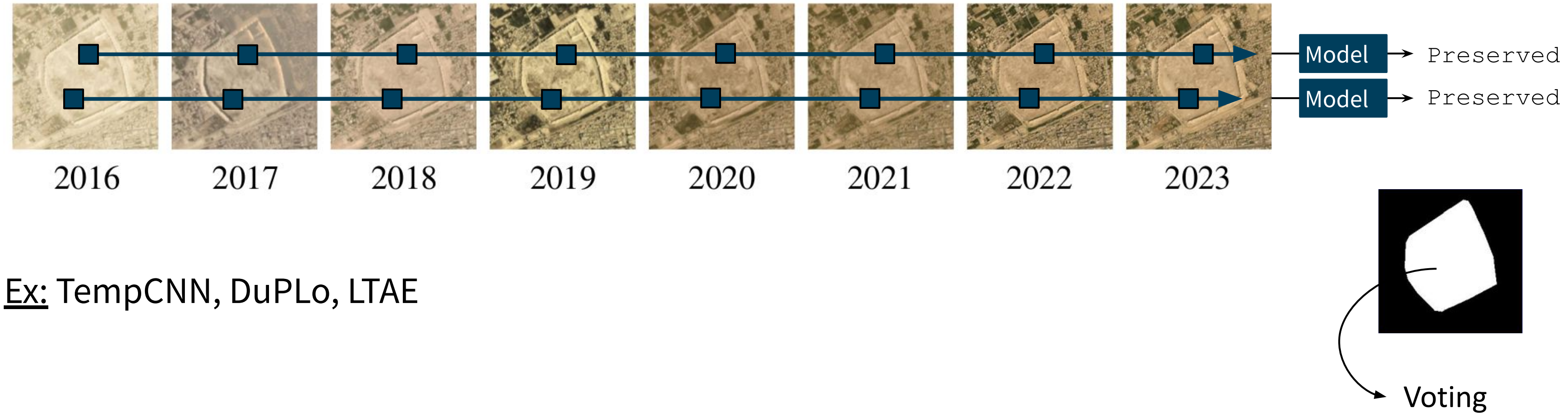
1. Multi-frame pixel-wise methods, *1 prediction for each pixel sequence,*
   → *aggregating predictions <u>spatially</u>*

2. Single-frame whole-image methods, *1 prediction for each time step,*
   → *aggregating predictions <u>temporally</u>*

3. Multi-frame whole-image methods, *direct prediction*

# Benchmarking deep learning methods

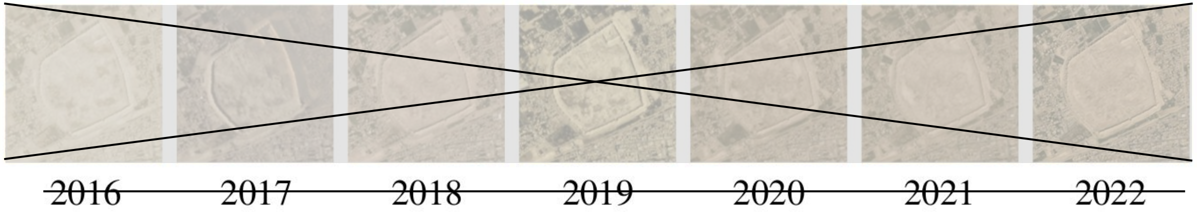
1. Multi-frame pixel-wise methods, *1 prediction for each pixel sequence*

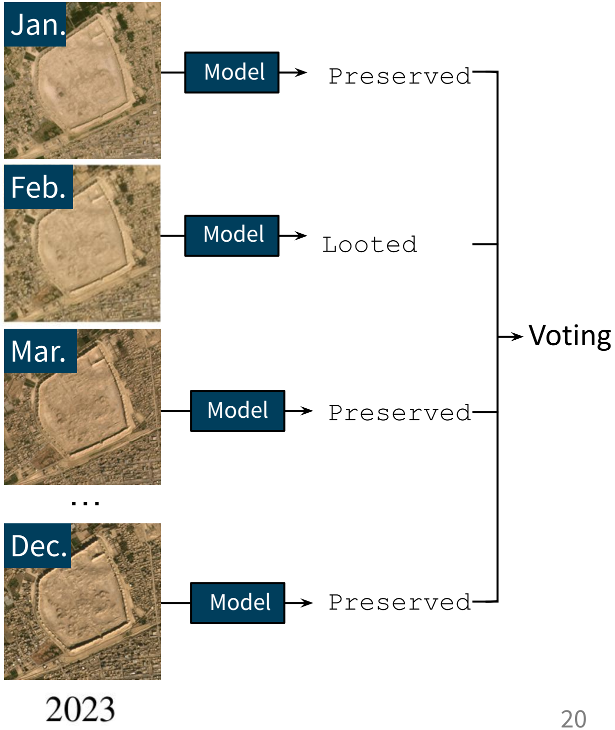



Ex: TempCNN, DuPLo, LTAE

# Benchmarking deep learning methods

2. Single-frame whole-image methods, *1 prediction for each time step*

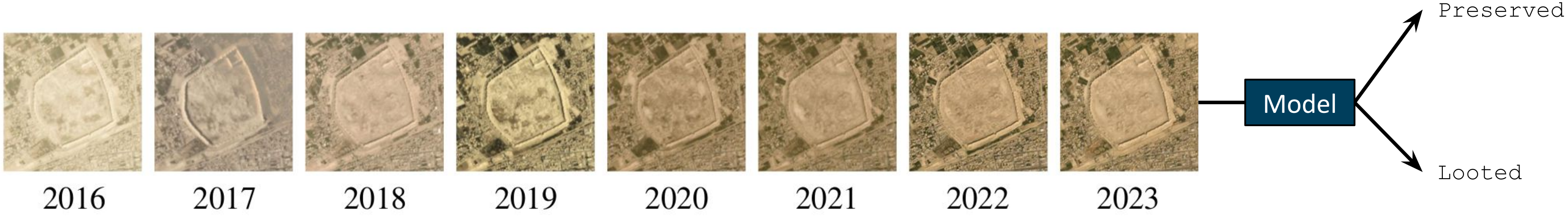



<u>Ex:</u> ResNets, ViTs

# Benchmarking deep learning methods

3. Multi-frame whole-image methods, *direct prediction*




<u>Ex:</u> ViTs + LTAE

→ Temporal attention

→ Good image representation, pre-trained (SatMAE, Scale-MAE, DOFA)

# Results

| Method | #param (x1000) | OA↑ | F1↑ | AUROC↑ |
|---|---|---|---|---|
| *Single-frame methods* | | | | |
| ResNet20 | 269.2 | 54.7 (8.9) | 54.5 (17.1) | 75.3 (3.1) |
| ResNet18 | 11,177.5 | 71.8 (2.6) | 64.1 (5.4) | 84.5 (1.5) |
| ResNet34 | 21,285.7 | 74.1 (3.2) | <u>68.9</u> (6.3) | <u>85.2</u> (1.7) |
| SatMAE | 2.1 | 63.6 (0.7) | 41.9 (0.4) | 75.3 (0.2) |
| Scale-MAE | 2.1 | 62.6 (0.7) | 39.3 (1.9) | 76.0 (0.3) |
| DOFA | 1.5 | <u>76.7</u> (2.8) | 67.0 (4.2) | 84.0 (1.4) |
| *Multi-frame methods* | | | | |
| *Pixel-wise methods* | | | | |
| DuPLo | 86.8 | 52.1 (2.8) | 50.4 (4.9) | 50.9 (3.7) |
| TempCNN | 28.5 | 55.7 (3.4) | 44.2 (9.7) | 58.8 (1.8) |
| Transformer | 38.5 | 56.4 (3.7) | 63.5 (3.2) | 62.7 (4.1) |
| LTAE | 32.2 | 52.5 (7.8) | 58.0 (4.6) | 62.0 (8.5) |
| *Whole-image methods* | | | | |
| PSE+LTAE | 34.0 | 55.1 (9.8) | 47.7 (6.2) | 59.5 (6.3) |
| UTAE | 68.9 | 62.0 (3.5) | 58.9 (2.3) | 64.5 (4.5) |
| TSViT (cls. head) | 236.9 | 64.3 (1.2) | 53.0 (3.7) | 70.8 (2.3) |
| TSViT (seg. head) | 237.4 | 64.6 (3.5) | 60.2 (7.1) | 69.6 (4.2) |
| SatMAE+LTAE | 1,627.9 | 67.9 (4.7) | 64.7 (4.0) | 75.2 (3.7) |
| Scale-MAE+LTAE | 1,627.9 | 68.5 (2.4) | 56.4 (7.7) | 77.6 (0.8) |
| DOFA+LTAE | 926.1 | **78.7** (2.3) | **74.9** (3.5) | **87.1** (3.0) |

Big table,
Wow, lots of numbers!

# Results

| Method | #param (x1000) | OA↑ | F1↑ | AUROC↑ |
|---|---|---|---|---|
| *Single-frame methods* | | | | |
| ResNet20 | 269.2 | 54.7 (8.9) | 54.5 (17.1) | 75.3 (3.1) |
| ResNet18 | 11,177.5 | 71.8 (2.6) | 64.1 (5.4) | 84.5 (1.5) |
| ResNet34 | 21,285.7 | 74.1 (3.2) | <u>68.9</u> (6.3) | <u>85.2</u> (1.7) |
| SatMAE | 2.1 | 63.6 (0.7) | 41.9 (0.4) | 75.3 (0.2) |
| Scale-MAE | 2.1 | 62.6 (0.7) | 39.3 (1.9) | 76.0 (0.3) |
| DOFA | 1.5 | <u>76.7</u> (2.8) | 67.0 (4.2) | 84.0 (1.4) |
| *Multi-frame methods* | | | | |
| *Pixel-wise methods* | | | | |
| DuPLo | 86.8 | 52.1 (2.8) | 50.4 (4.9) | 50.9 (3.7) |
| TempCNN | 28.5 | 55.7 (3.4) | 44.2 (9.7) | 58.8 (1.8) |
| Transformer | 38.5 | 56.4 (3.7) | 63.5 (3.2) | 62.7 (4.1) |
| LTAE | 32.2 | 52.5 (7.8) | 58.0 (4.6) | 62.0 (8.5) |
| *Whole-image methods* | | | | |
| PSE+LTAE | 34.0 | 55.1 (9.8) | 47.7 (6.2) | 59.5 (6.3) |
| UTAE | 68.9 | 62.0 (3.5) | 58.9 (2.3) | 64.5 (4.5) |
| TSViT (cls. head) | 236.9 | 64.3 (1.2) | 53.0 (3.7) | 70.8 (2.3) |
| TSViT (seg. head) | 237.4 | 64.6 (3.5) | 60.2 (7.1) | 69.6 (4.2) |
| SatMAE+LTAE | 1,627.9 | 67.9 (4.7) | 64.7 (4.0) | 75.2 (3.7) |
| Scale-MAE+LTAE | 1,627.9 | 68.5 (2.4) | 56.4 (7.7) | 77.6 (0.8) |
| DOFA+LTAE | 926.1 | **78.7** (2.3) | **74.9** (3.5) | **87.1** (3.0) |

Pixel-wise methods are clearly outperformed by others → importance of spatial context

# Results

| Method | #param (x1000) | OA↑ | F1↑ | AUROC↑ |
|---|---|---|---|---|
| *Single-frame methods* | | | | |
| ResNet20 | 269.2 | 54.7 (8.9) | 54.5 (17.1) | 75.3 (3.1) |
| ResNet18 | 11,177.5 | 71.8 (2.6) | 64.1 (5.4) | 84.5 (1.5) |
| ResNet34 | 21,285.7 | 74.1 (3.2) | <u>68.9</u> (6.3) | <u>85.2</u> (1.7) |
| SatMAE | 2.1 | 63.6 (0.7) | 41.9 (0.4) | 75.3 (0.2) |
| Scale-MAE | 2.1 | 62.6 (0.7) | 39.3 (1.9) | 76.0 (0.3) |
| DOFA | 1.5 | <u>76.7</u> (2.8) | 67.0 (4.2) | 84.0 (1.4) |
| *Multi-frame methods* | | | | |
| *Pixel-wise methods* | | | | |
| DuPLo | 86.8 | 52.1 (2.8) | 50.4 (4.9) | 50.9 (3.7) |
| TempCNN | 28.5 | 55.7 (3.4) | 44.2 (9.7) | 58.8 (1.8) |
| Transformer | 38.5 | 56.4 (3.7) | 63.5 (3.2) | 62.7 (4.1) |
| LTAE | 32.2 | 52.5 (7.8) | 58.0 (4.6) | 62.0 (8.5) |
| *Whole-image methods* | | | | |
| PSE+LTAE | 34.0 | 55.1 (9.8) | 47.7 (6.2) | 59.5 (6.3) |
| UTAE | 68.9 | 62.0 (3.5) | 58.9 (2.3) | 64.5 (4.5) |
| TSViT (cls. head) | 236.9 | 64.3 (1.2) | 53.0 (3.7) | 70.8 (2.3) |
| TSViT (seg. head) | 237.4 | 64.6 (3.5) | 60.2 (7.1) | 69.6 (4.2) |
| SatMAE+LTAE | 1,627.9 | 67.9 (4.7) | 64.7 (4.0) | 75.2 (3.7) |
| Scale-MAE+LTAE | 1,627.9 | 68.5 (2.4) | 56.4 (7.7) | 77.6 (0.8) |
| DOFA+LTAE | 926.1 | **78.7** (2.3) | **74.9** (3.5) | **87.1** (3.0) |

Temporal methods improve over single-frame methods

Image representation

<u>On average:</u>

+6% Accuracy

+37% F1 score

Image representation + temporal attention

# Results

| Method | #param (x1000) | OA↑ | F1↑ | AUROC↑ |
|---|---|---|---|---|
| *Single-frame methods* | | | | |
| ResNet20 | 269.2 | 54.7 (8.9) | 54.5 (17.1) | 75.3 (3.1) |
| ResNet18 | 11,177.5 | 71.8 (2.6) | 64.1 (5.4) | 84.5 (1.5) |
| ResNet34 | 21,285.7 | 74.1 (3.2) | <u>68.9</u> (6.3) | <u>85.2</u> (1.7) |
| SatMAE | 2.1 | 63.6 (0.7) | 41.9 (0.4) | 75.3 (0.2) |
| Scale-MAE | 2.1 | 62.6 (0.7) | 39.3 (1.9) | 76.0 (0.3) |
| DOFA | 1.5 | <u>76.7</u> (2.8) | 67.0 (4.2) | 84.0 (1.4) |
| *Multi-frame methods* | | | | |
| *Pixel-wise methods* | | | | |
| DuPLo | 86.8 | 52.1 (2.8) | 50.4 (4.9) | 50.9 (3.7) |
| TempCNN | 28.5 | 55.7 (3.4) | 44.2 (9.7) | 58.8 (1.8) |
| Transformer | 38.5 | 56.4 (3.7) | 63.5 (3.2) | 62.7 (4.1) |
| LTAE | 32.2 | 52.5 (7.8) | 58.0 (4.6) | 62.0 (8.5) |
| *Whole-image methods* | | | | |
| PSE+LTAE | 34.0 | 55.1 (9.8) | 47.7 (6.2) | 59.5 (6.3) |
| UTAE | 68.9 | 62.0 (3.5) | 58.9 (2.3) | 64.5 (4.5) |
| TSViT (cls. head) | 236.9 | 64.3 (1.2) | 53.0 (3.7) | 70.8 (2.3) |
| TSViT (seg. head) | 237.4 | 64.6 (3.5) | 60.2 (7.1) | 69.6 (4.2) |
| SatMAE+LTAE | 1,627.9 | 67.9 (4.7) | 64.7 (4.0) | 75.2 (3.7) |
| Scale-MAE+LTAE | 1,627.9 | 68.5 (2.4) | 56.4 (7.7) | 77.6 (0.8) |
| DOFA+LTAE | 926.1 | **78.7** (2.3) | **74.9** (3.5) | **87.1** (3.0) |

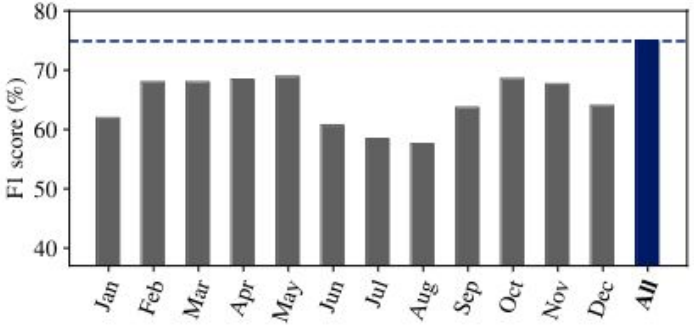
“Foundation models” provide strong representations for this downstream task

**Best performing model:** DOFA (pretrained, frozen) + LTAE

# Temporal analysis

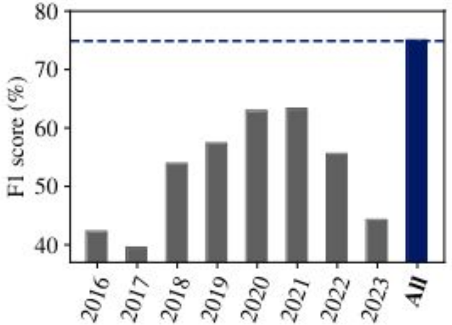
- Inference experiments with DOFA+LTAE

→ **The more time steps, the better**



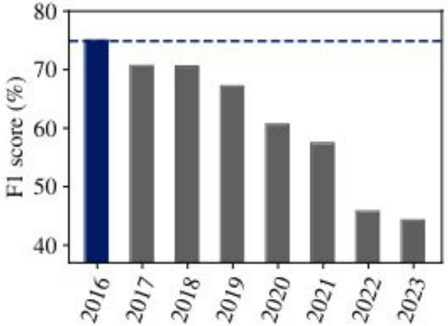

Month-specific time series

*(i) Seasonal behaviour*




Year-specific time series

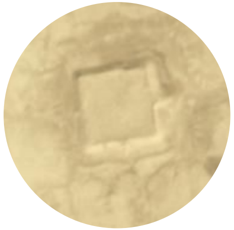
*(ii) Indicating looting activities?*



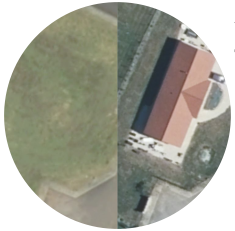

Sequence of decreasing size (starting year in *x*)

*(iii) Importance of temporal range*

# Outline



1 Afghan archaeological site looting detection

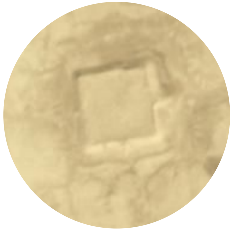


2 Semantic change detection and domain shift analysis

3 Crop-type classification with few or no annotations

✔ Providing labeled data for a specific task/location
✔ Making use of pre-trained off-the-shelf models

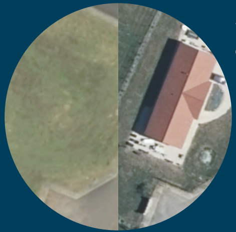
# Outline



1 Afghan archaeological site looting detection

*Detecting Looted Archaeological Sites from Satellite Image Time Series*
**E. Vincent**, M. Saroufim, J. Chemla, Y. Ubelmann, P. Marquis, J. Ponce, M. Aubry
EarthVision CVPR Workshop 2025

Best student paper award



2 Semantic change detection and domain shift analysis

*Satellite Image Time Series Semantic Change Detection: Novel Architecture and Analysis of Domain Shift*
**E. Vincent**, J. Ponce, M. Aubry – arXiv 2024

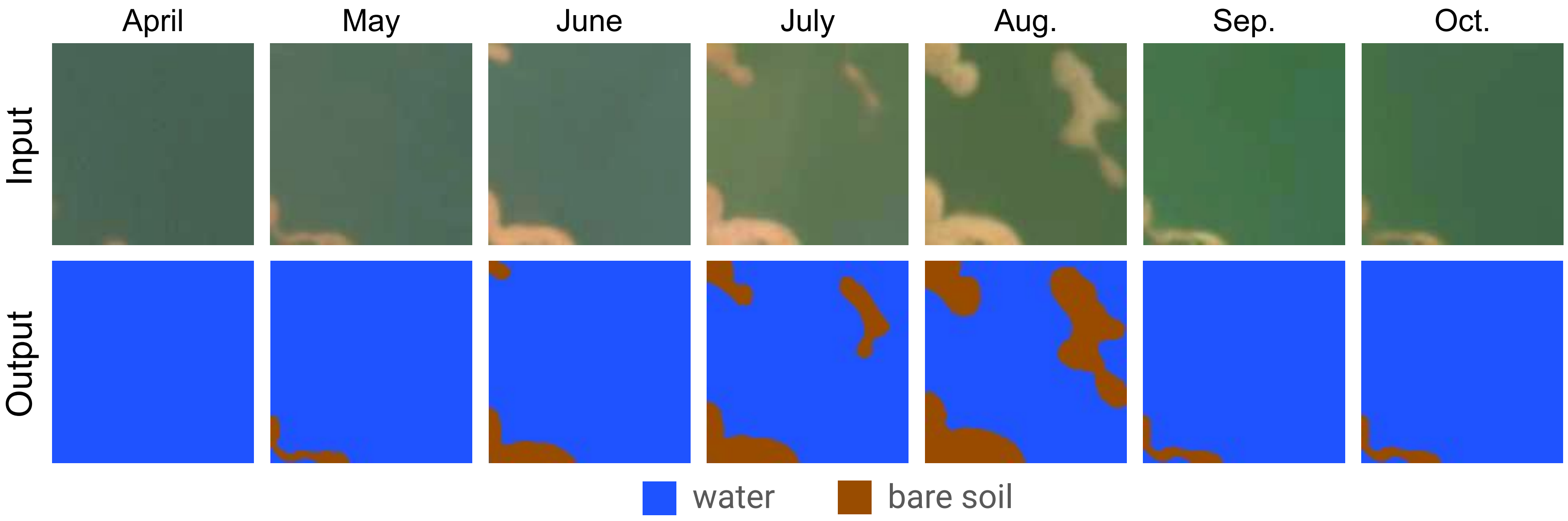
*CoDEx: Combining Domain Expertise for Spatial Generalization in Satellite Image Analysis*
A. Kuriyal, **E. Vincent**, M. Aubry, L. Landrieu – EarthVision CVPR Workshop 2025

3 Crop-type classification with few or no annotations

*Pixel-wise Agricultural Image Time Series Classification: Comparison and a Deformable Prototype-based Approach*
**E. Vincent**, J. Ponce, M. Aubry – IGARSS 2025

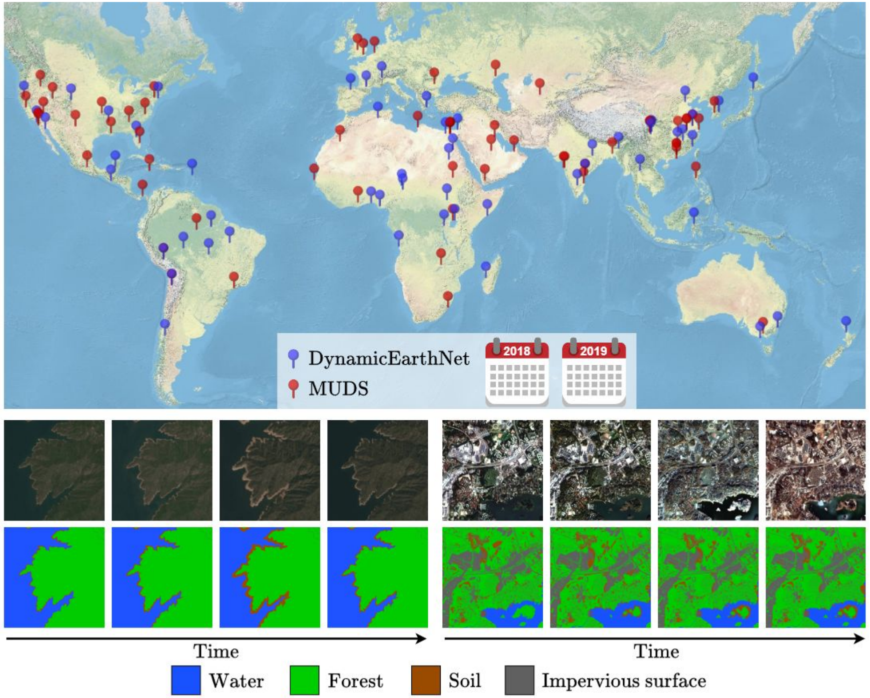
# Semantic Change Detection

**Time series** allow to spot land cover change at high frequency (e.g. monthly)



# Spatial and temporal domain shift

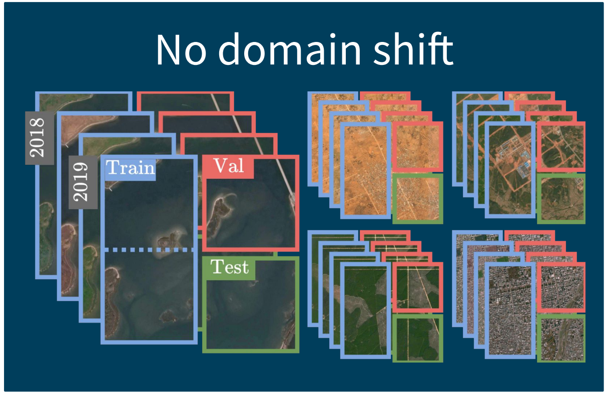
Methods evaluated in settings with **domain shift**




→ **global** and **multi-year** datasets
→ allow us to define challenging dataset splits exhibiting either spatial or temporal domain shift

# Domain shift settings

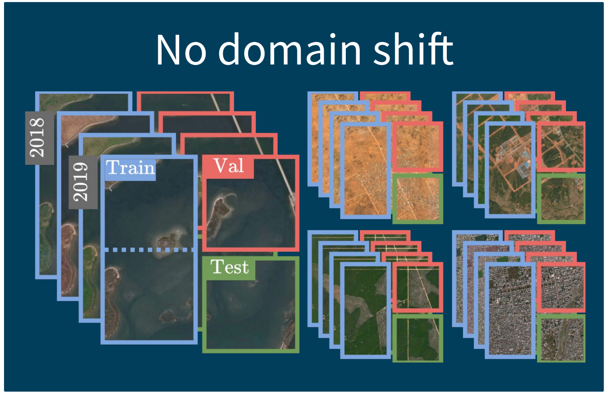
- 2 land-cover SITS datasets: DynamicEarthNet and MUDS
  - global spatial coverage
  - multi-year temporal coverage
- 3 domain shift settings

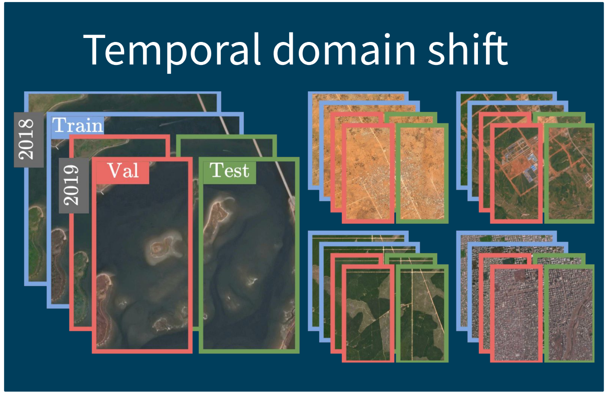



Toker, A., et al. "Dynamicearthnet: Daily multi-spectral satellite dataset for semantic change segmentation" *CVPR* 2022.
Van Etten, A., et al. "The multi-temporal urban development spacenet dataset" *CVPR* 2021.

# Domain shift settings

- 2 land-cover SITS datasets: DynamicEarthNet and MUDS
  - global spatial coverage
  - multi-year temporal coverage
- 3 domain shift settings

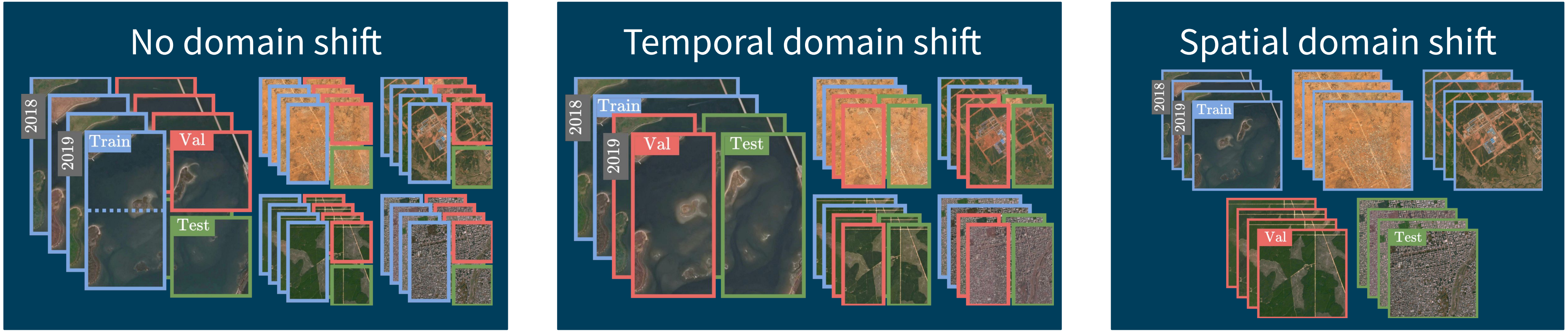



Toker, A., et al. "Dynamicearthnet: Daily multi-spectral satellite dataset for semantic change segmentation" *CVPR* 2022.
Van Etten, A., et al. "The multi-temporal urban development spacenet dataset" *CVPR* 2021.

# Domain shift settings

- 2 land-cover SITS datasets: DynamicEarthNet and MUDS
  - global spatial coverage
  - multi-year temporal coverage
- 3 domain shift settings





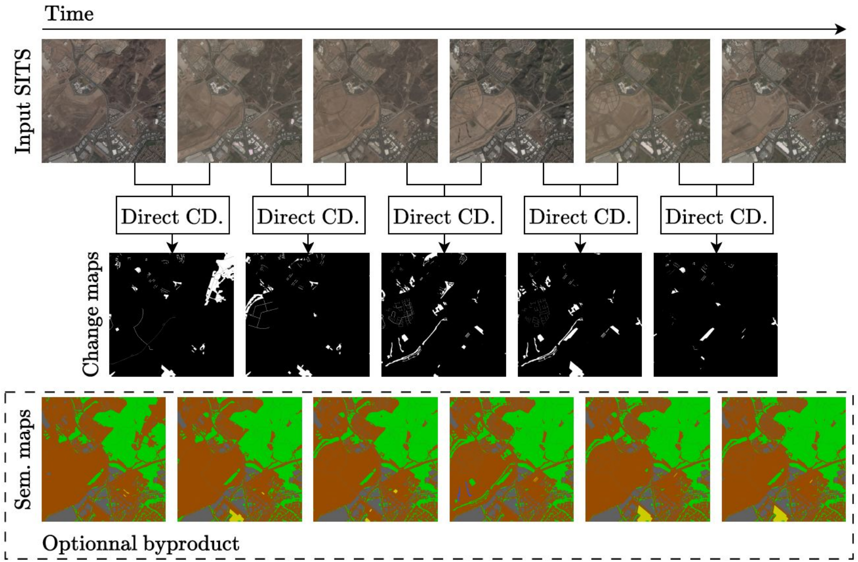
Toker, A., et al. "Dynamicearthnet: Daily multi-spectral satellite dataset for semantic change segmentation" *CVPR* 2022.
Van Etten, A., et al. "The multi-temporal urban development spacenet dataset" *CVPR* 2021.

# Semantic Change Detection

**Time series** allow to spot land cover change at high frequency (e.g. monthly)

But most change detection approaches are either:

- **bi-temporal** (process image pairs)



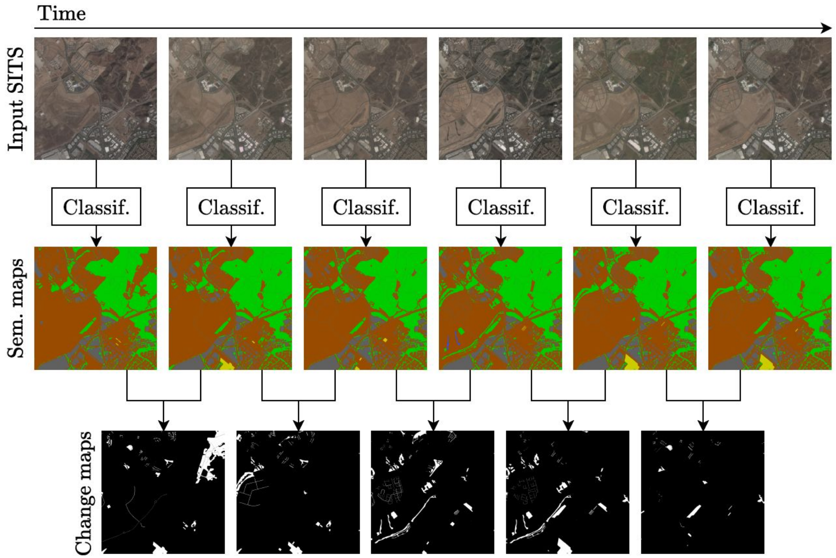

→ siamese architectures, 3-branch networks, multi-task learning, etc.

# Semantic Change Detection

**Time series** allow to spot land cover change at high frequency (e.g. monthly)

But most change detection approaches are either:

- **bi-temporal** (process image pairs)
- **mono-frame** (post-classification methods)



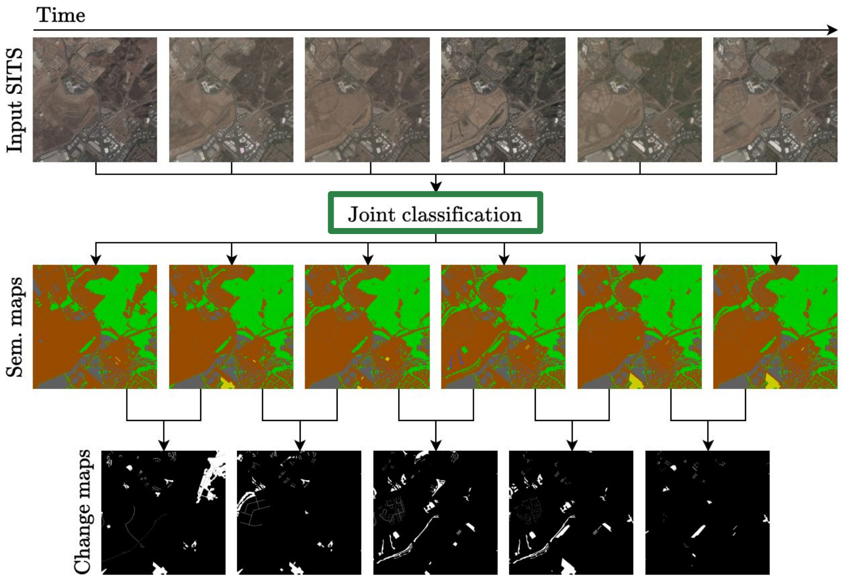

→ build on SOTA segmentation methods

# SITS-Semantic Change Detection

**Time series** allow to spot land cover change at high frequency (e.g. monthly)

But most change detection approaches are either:

- ~~**bi-temporal** (process image pairs)~~
- ~~**mono-frame** (post-classification methods)~~




**Instead** → perform the classification **jointly**

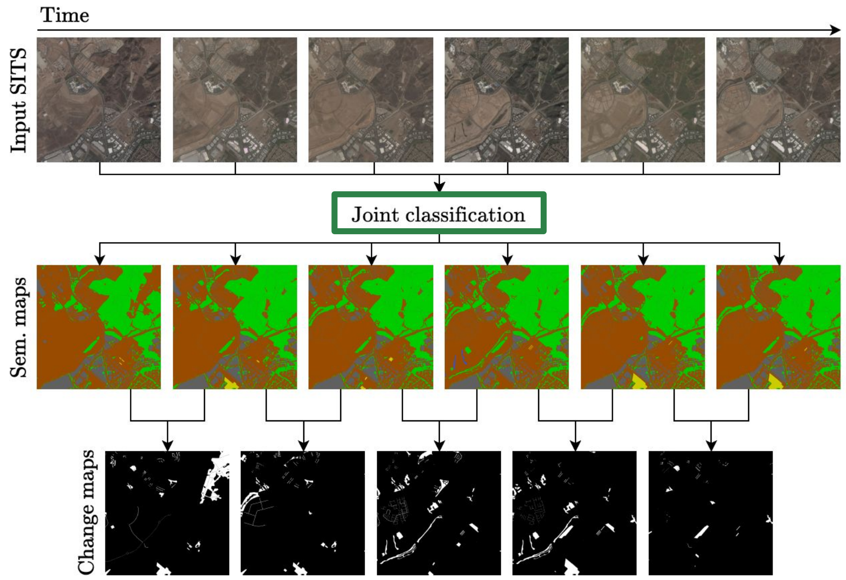
- Sequence-to-sequence model
- Leverage long-range temporal information
- Less prone to false positives

# SITS-Semantic Change Detection

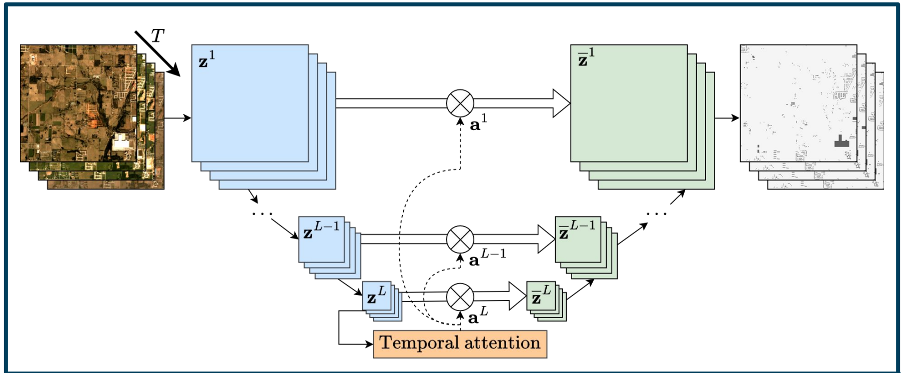
**Time series** allow to spot land cover change at high frequency (e.g. monthly)

But most change detection approaches are either:

- ~~**bi-temporal** (process image pairs)~~
- ~~**mono-frame** (post-classification methods)~~




**Instead** → perform the classification **jointly**




**Editing the temporal attention mechanism of UTAE**

V. Sainte Fare Garnot *et al*. *Panoptic segmentation of satellite image time series with convolutional temporal attention networks*. CVPR 2021.

# Results

| Dataset | Method | Input type | Strategy |
|---|---|---|---|
| DynamicEarthNet | Random | — | — |
| | TSViT monthly | Single image | Mono |
| | UTAE monthly | Single image | Mono |
| | TSViT weekly | SITS | Mono |
| | UTAE weekly | SITS | Mono |
| | A2Net | Image pair | Bi |
| | SCanNet | Image pair | Bi |
| | TSSCD | Pixel-wise SITS | Multi |
| | Ours | SITS | Multi |
| MUDS | Random | — | — |
| | TSViT monthly | Single image | Mono |
| | UTAE monthly | Single image | Mono |
| | A2Net | Image pair | Bi |
| | SCanNet | Image pair | Bi |
| | TSSCD | Pixel-wise SITS | Multi |
| | Ours | SITS | Multi |

Mono-frame methods

Bi-temporal methods

Sequence-to-sequence (pixel-wise)

Sequence-to-sequence

# Results

| Dataset | Method | Input type | Strategy | No domain shift BC↑ | No domain shift mIoU↑ | Temporal domain shift BC↑ | Temporal domain shift mIoU↑ | Spatial domain shift BC↑ | Spatial domain shift mIoU↑ |
|---|---|---|---|---|---|---|---|---|---|
| DynamicEarthNet | Random | — | — | | | | | | |
| | TSViT monthly | Single image | Mono | | | | | | |
| | UTAE monthly | Single image | Mono | | | | | | |
| | TSViT weekly | SITS | Mono | | | | | | |
| | UTAE weekly | SITS | Mono | | | | | | |
| | A2Net | Image pair | Bi | | | | | | |
| | SCanNet | Image pair | Bi | | | | | | |
| | TSSCD | Pixel-wise SITS | Multi | | | | | | |
| | Ours | SITS | Multi | | | | | | |
| MUDS | Random | — | — | | | | | | |
| | TSViT monthly | Single image | Mono | | | | | | |
| | UTAE monthly | Single image | Mono | | | | | | |
| | A2Net | Image pair | Bi | | | | | | |
| | SCanNet | Image pair | Bi | | | | | | |
| | TSSCD | Pixel-wise SITS | Multi | | | | | | |
| | Ours | SITS | Multi | | | | | | |

Binary change

Semantic segmentation

# Results

Binary change

Semantic segmentation

| | Method | Input type | Strategy | No domain shift BC↑ | No domain shift mIoU↑ | Temporal domain shift BC↑ | Temporal domain shift mIoU↑ | Spatial domain shift BC↑ | Spatial domain shift mIoU↑ |
|---|---|---|---|---|---|---|---|---|---|
| DynamicEarthNet | Random | — | — | 4.9 | 7.3 | 5.0 | 7.3 | 4.9 | 7.1 |
| | TSViT monthly | Single image | Mono | 11.8 | 50.5 | 9.9 | 47.3 | 7.9 | 31.2 |
| | UTAE monthly | Single image | Mono | 13.8 | 53.7 | 10.9 | 53.7 | 9.0 | 36.9 |
| | TSViT weekly | SITS | Mono | 12.5 | 50.9 | 10.9 | 51.4 | 7.4 | 32.2 |
| | UTAE weekly | SITS | Mono | 14.3 | 54.4 | 11.3 | 54.7 | 8.7 | 37.8 |
| | A2Net | Image pair | Bi | 11.5 | 47.2 | 11.0 | 46.7 | 8.2 | 37.9 |
| | SCanNet | Image pair | Bi | 13.9 | 53.0 | 13.1 | 55.6 | 9.3 | 37.3 |
| | TSSCD | Pixel-wise SITS | Multi | 4.7 | 33.9 | 5.2 | 29.4 | 5.2 | 22.9 |
| | Ours | SITS | Multi | **22.4** | **60.5** | **15.3** | **61.7** | **10.1** | **38.5** |
| MUDS | Random | — | — | 0.1 | 28.1 | 0.1 | 28.1 | 0.1 | 28.1 |
| | TSViT monthly | Single image | Mono | 0.5 | 60.2 | 0.5 | 56.8 | 0.4 | 49.8 |
| | UTAE monthly | Single image | Mono | 0.6 | 67.1 | 0.6 | 66.0 | 0.6 | 63.0 |
| | A2Net | Image pair | Bi | 0.5 | 61.5 | 0.6 | 56.1 | 0.5 | 53.0 |
| | SCanNet | Image pair | Bi | 0.7 | 64.9 | 0.8 | 62.8 | 0.4 | 58.8 |
| | TSSCD | Pixel-wise SITS | Multi | 0.1 | 47.7 | 0.2 | 49.6 | 0.1 | 43.6 |
| | Ours | SITS | Multi | **1.7** | **72.0** | **1.9** | **71.1** | **0.7** | **66.2** |

Semantic segmentation

# Results

| | Method | Input type | Strategy | No domain shift BC↑ | No domain shift mIoU↑ | Temporal domain shift BC↑ | Temporal domain shift mIoU↑ | Spatial domain shift BC↑ | Spatial domain shift mIoU↑ |
|---|---|---|---|---|---|---|---|---|---|
| DynamicEarthNet | Random | — | — | 4.9 | 7.3 | 5.0 | 7.3 | 4.9 | 7.1 |
| | TSViT monthly | Single image | Mono | 11.8 | 50.5 | 9.9 | 47.3 | 7.9 | 31.2 |
| | UTAE monthly | Single image | Mono | 13.8 | 53.7 | 10.9 | 53.7 | 9.0 | 36.9 |
| | TSViT weekly | SITS | Mono | 12.5 | 50.9 | 10.9 | 51.4 | 7.4 | 32.2 |
| | UTAE weekly | SITS | Mono | 14.3 | 54.4 | 11.3 | 54.7 | 8.7 | 37.8 |
| | A2Net | Image pair | Bi | 11.5 | 47.2 | 11.0 | 46.7 | 8.2 | 37.9 |
| | SCanNet | Image pair | Bi | 13.9 | 53.0 | 13.1 | 55.6 | 9.3 | 37.3 |
| | TSSCD | Pixel-wise SITS | Multi | 4.7 | 33.9 | 5.2 | 29.4 | 5.2 | 22.9 |
| | Ours | SITS | Multi | **22.4** | **60.5** | **15.3** | **61.7** | **10.1** | **38.5** |
| MUDS | Random | — | — | 0.1 | 28.1 | 0.1 | 28.1 | 0.1 | 28.1 |
| | TSViT monthly | Single image | Mono | 0.5 | 60.2 | 0.5 | 56.8 | 0.4 | 49.8 |
| | UTAE monthly | Single image | Mono | 0.6 | 67.1 | 0.6 | 66.0 | 0.6 | 63.0 |
| | A2Net | Image pair | Bi | 0.5 | 61.5 | 0.6 | 56.1 | 0.5 | 53.0 |
| | SCanNet | Image pair | Bi | 0.7 | 64.9 | 0.8 | 62.8 | 0.4 | 58.8 |
| | TSSCD | Pixel-wise SITS | Multi | 0.1 | 47.7 | 0.2 | 49.6 | 0.1 | 43.6 |
| | Ours | SITS | Multi | **1.7** | **72.0** | **1.9** | **71.1** | **0.7** | **66.2** |

# Results

Binary change

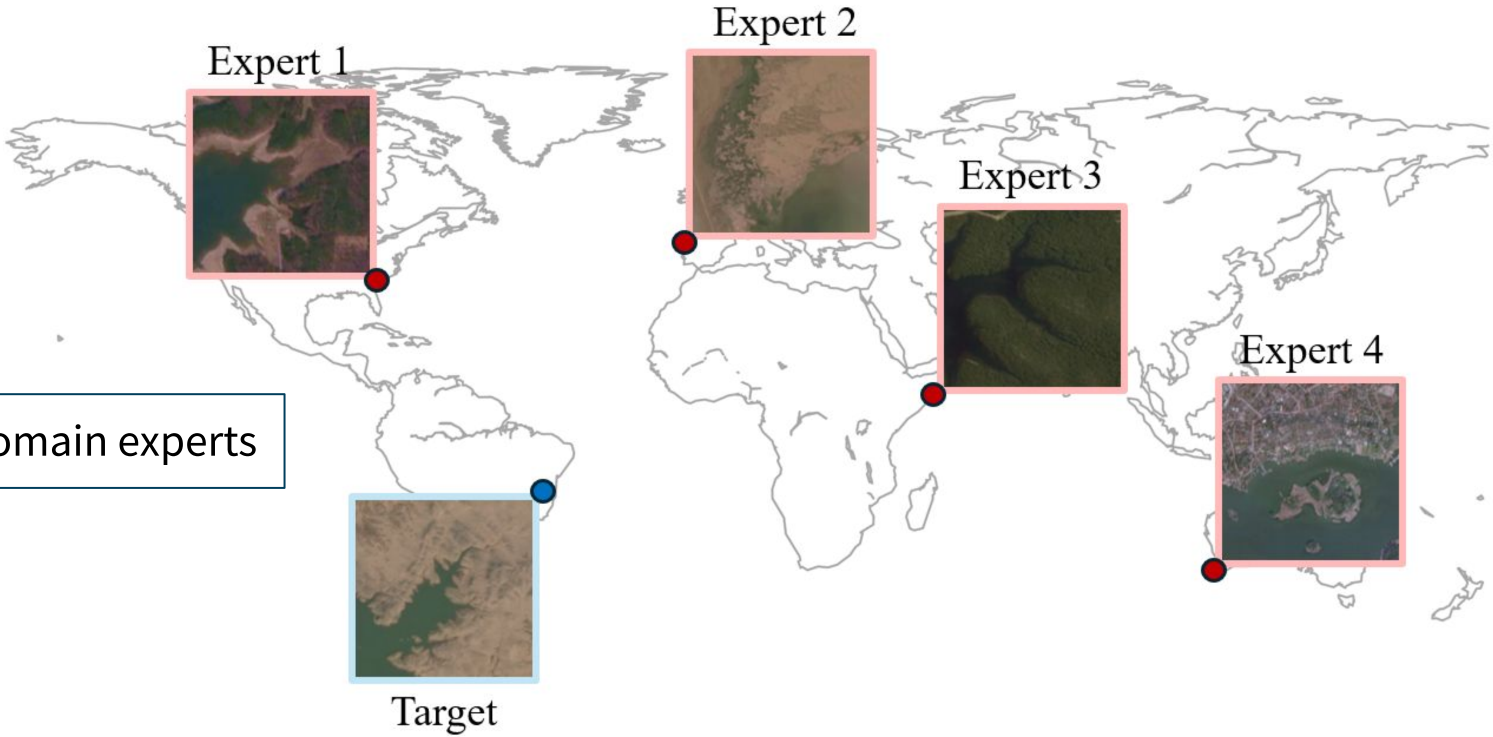
| | Method | Input type | Strategy | No domain shift | | Temporal domain shift | | Spatial domain shift | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | BC↑ | mIoU↑ | BC↑ | mIoU↑ | BC↑ | mIoU↑ |
| DynamicEarthNet | Random | — | — | 4.9 | 7.3 | 5.0 | 7.3 | 4.9 | 7.1 |
| | TSViT monthly | Single image | Mono | 11.8 | 50.5 | 9.9 | 47.3 | 7.9 | 31.2 |
| | UTAE monthly | Single image | Mono | 13.8 | 53.7 | 10.9 | 53.7 | 9.0 | 36.9 |
| | TSViT weekly | SITS | Mono | 12.5 | 50.9 | 10.9 | 51.4 | 7.4 | 32.2 |
| | UTAE weekly | SITS | Mono | 14.3 | 54.4 | 11.3 | 54.7 | 8.7 | 37.8 |
| | A2Net | Image pair | Bi | 11.5 | 47.2 | 11.0 | 46.7 | 8.2 | 37.9 |
| | SCanNet | Image pair | Bi | 13.9 | 53.0 | 13.1 | 55.6 | 9.3 | 37.3 |
| | TSSCD | Pixel-wise SITS | Multi | 4.7 | 33.9 | 5.2 | 29.4 | 5.2 | 22.9 |
| | Ours | SITS | Multi | **22.4** | **60.5** | **15.3** | **61.7** | **10.1** | **38.5** |
| MUDS | Random | — | — | 0.1 | 28.1 | 0.1 | 28.1 | 0.1 | 28.1 |
| | TSViT monthly | Single image | Mono | 0.5 | 60.2 | 0.5 | 56.8 | 0.4 | 49.8 |
| | UTAE monthly | Single image | Mono | 0.6 | 67.1 | 0.6 | 66.0 | 0.6 | 63.0 |
| | A2Net | Image pair | Bi | 0.5 | 61.5 | 0.6 | 56.1 | 0.5 | 53.0 |
| | SCanNet | Image pair | Bi | 0.7 | 64.9 | 0.8 | 62.8 | 0.4 | 58.8 |
| | TSSCD | Pixel-wise SITS | Multi | 0.1 | 47.7 | 0.2 | 49.6 | 0.1 | 43.6 |
| | Ours | SITS | Multi | **1.7** | **72.0** | **1.9** | **71.1** | **0.7** | **66.2** |

# Results

Dramatic impact of spatial domain shift overall!

| Dataset | Method | Input type | Strategy | No domain shift BC↑ | No domain shift mIoU↑ | Temporal domain shift BC↑ | Temporal domain shift mIoU↑ | Spatial domain shift BC↑ | Spatial domain shift mIoU↑ |
|---|---|---|---|---|---|---|---|---|---|
| DynamicEarthNet | Random | — | — | 4.9 | 7.3 | 5.0 | 7.3 | 4.9 | 7.1 |
| | TSViT monthly | Single image | Mono | 11.8 | 50.5 | 9.9 | 47.3 | 7.9 | 31.2 |
| | UTAE monthly | Single image | Mono | 13.8 | 53.7 | 10.9 | 53.7 | 9.0 | 36.9 |
| | TSViT weekly | SITS | Mono | 12.5 | 50.9 | 10.9 | 51.4 | 7.4 | 32.2 |
| | UTAE weekly | SITS | Mono | 14.3 | 54.4 | 11.3 | 54.7 | 8.7 | 37.8 |
| | A2Net | Image pair | Bi | 11.5 | 47.2 | 11.0 | 46.7 | 8.2 | 37.9 |
| | SCanNet | Image pair | Bi | 13.9 | 53.0 | 13.1 | 55.6 | 9.3 | 37.3 |
| | TSSCD | Pixel-wise SITS | Multi | 4.7 | 33.9 | 5.2 | 29.4 | 5.2 | 22.9 |
| | Ours | SITS | Multi | **22.4** | **60.5** | **15.3** | **61.7** | **10.1** | **38.5** |
| MUDS | Random | — | — | 0.1 | 28.1 | 0.1 | 28.1 | 0.1 | 28.1 |
| | TSViT monthly | Single image | Mono | 0.5 | 60.2 | 0.5 | 56.8 | 0.4 | 49.8 |
| | UTAE monthly | Single image | Mono | 0.6 | 67.1 | 0.6 | 66.0 | 0.6 | 63.0 |
| | A2Net | Image pair | Bi | 0.5 | 61.5 | 0.6 | 56.1 | 0.5 | 53.0 |
| | SCanNet | Image pair | Bi | 0.7 | 64.9 | 0.8 | 62.8 | 0.4 | 58.8 |
| | TSSCD | Pixel-wise SITS | Multi | 0.1 | 47.7 | 0.2 | 49.6 | 0.1 | 43.6 |
| | Ours | SITS | Multi | **1.7** | **72.0** | **1.9** | **71.1** | **0.7** | **66.2** |

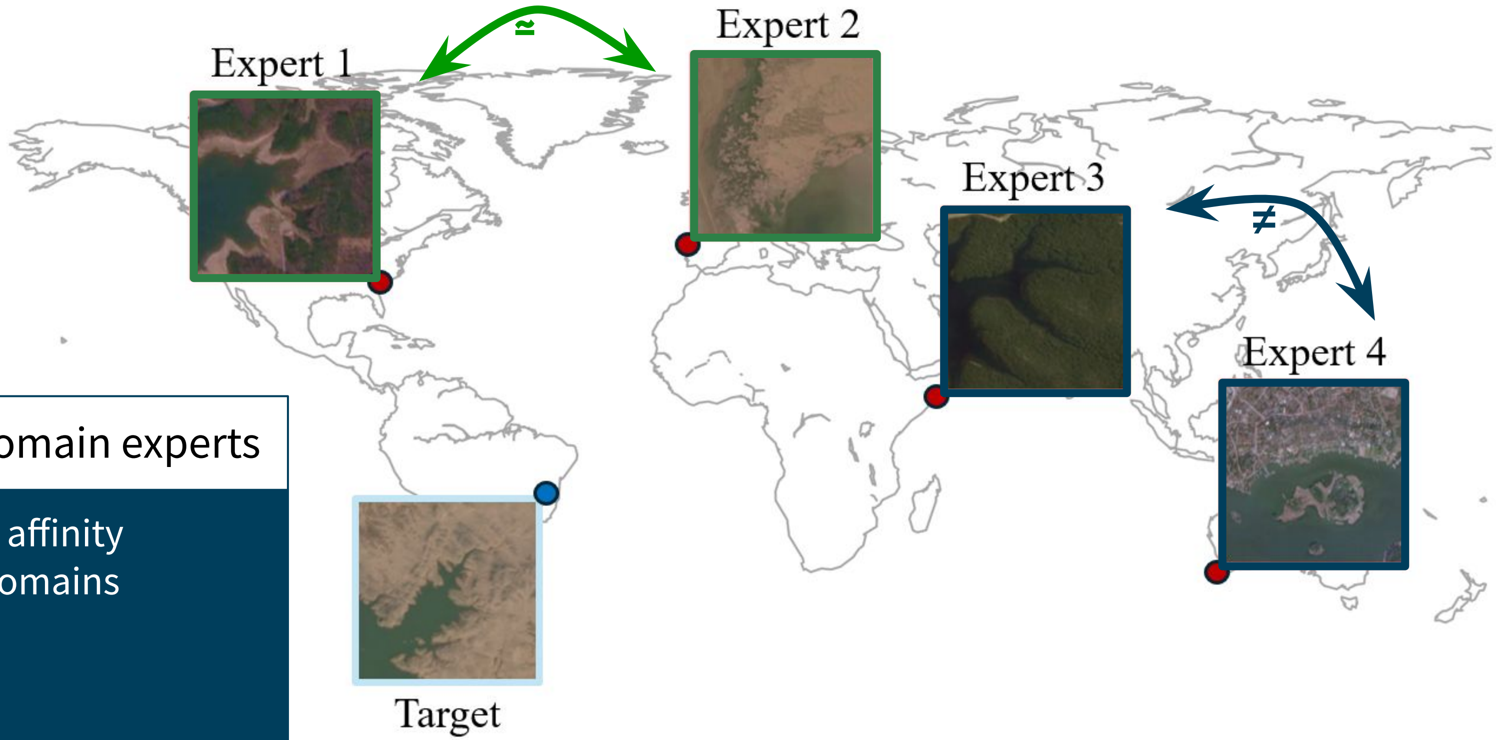
# Addressing spatial domain shift




… with domain experts

A. Kuriyal, **E. Vincent,** M. Aubry, L. Landrieu. *CoDEx: Combining Domain Expertise for Spatial Generalization in Satellite Image Analysis*. CVPRW 2025

# Addressing spatial domain shift

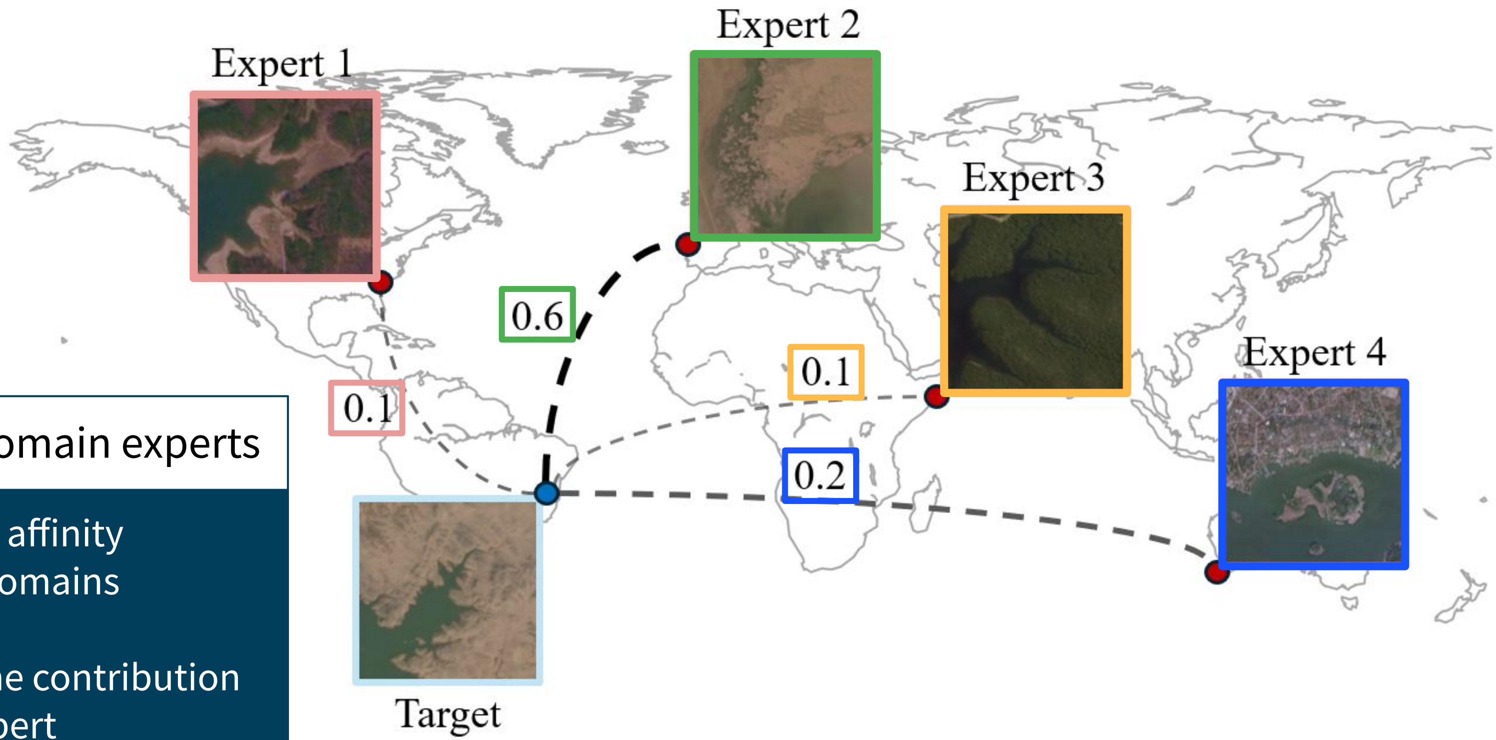



… with domain experts

- Learn the affinity between domains

A. Kuriyal, **E. Vincent,** M. Aubry, L. Landrieu. *CoDEx: Combining Domain Expertise for Spatial Generalization in Satellite Image Analysis*. CVPRW 2025

# Addressing spatial domain shift




… with domain experts

- Learn the affinity between domains

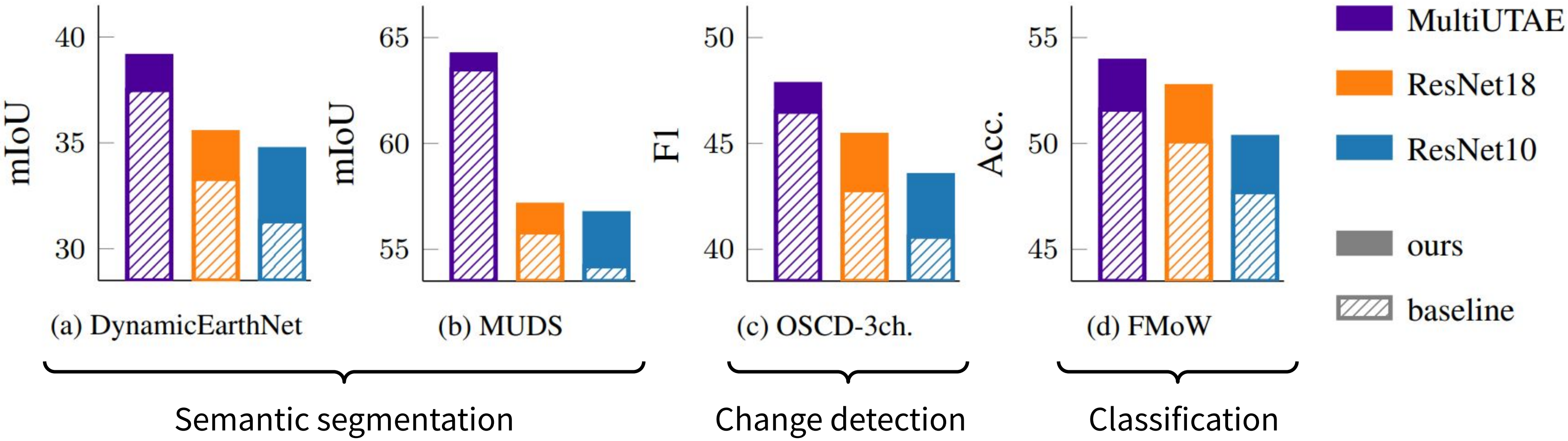
- Predict the contribution of each expert

A. Kuriyal, **E. Vincent,** M. Aubry, L. Landrieu. *CoDEx: Combining Domain Expertise for Spatial Generalization in Satellite Image Analysis*. CVPRW 2025

# Addressing spatial domain shift

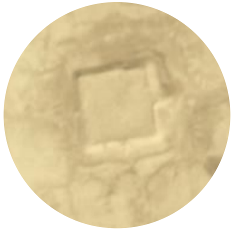
Improving performance:

- across 3 tasks and 3 baselines
- on 4 satellite image datasets

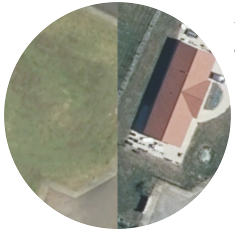



A. Kuriyal, **E. Vincent,** M. Aubry, L. Landrieu. *CoDEx: Combining Domain Expertise for Spatial Generalization in Satellite Image Analysis*. CVPRW 2025

# Outline



1 Afghan archaeological site looting detection

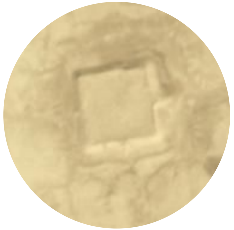


2 Semantic change detection and domain shift analysis

✔ Evaluating the impact of temporal/spatial shift
✔ Addressing spatial shift with domain experts

3 Crop-type classification with few or no annotations

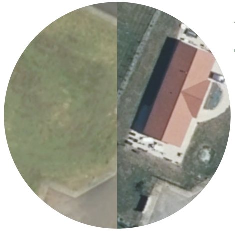
# Outline



1 Afghan archaeological site looting detection

*Detecting Looted Archaeological Sites from Satellite Image Time Series*
**E. Vincent**, M. Saroufim, J. Chemla, Y. Ubelmann, P. Marquis, J. Ponce, M. Aubry
EarthVision CVPR Workshop 2025

Best student paper award



2 Semantic change detection and domain shift analysis

*Satellite Image Time Series Semantic Change Detection: Novel Architecture and Analysis of Domain Shift*
**E. Vincent**, J. Ponce, M. Aubry – arXiv 2024

*CoDEx: Combining Domain Expertise for Spatial Generalization in Satellite Image Analysis*
A. Kuriyal, **E. Vincent**, M. Aubry, L. Landrieu – EarthVision CVPR Workshop 2025

3 Crop-type classification with few or no annotations

*Pixel-wise Agricultural Image Time Series Classification: Comparison and a Deformable Prototype-based Approach*
**E. Vincent**, J. Ponce, M. Aubry – IGARSS 2025

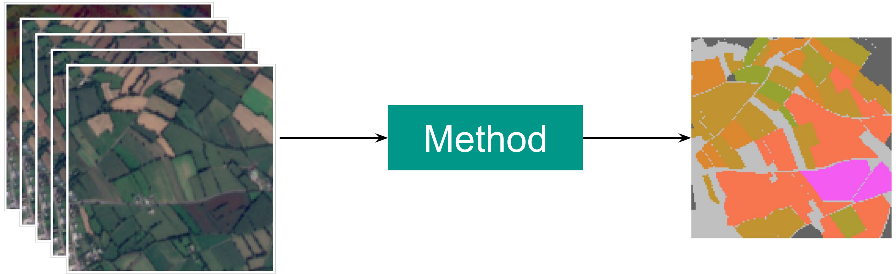
# Task

Agricultural satellite image time series (SITS) classification


Crop-type pixel-wise classification (wheat, oat, potatoes, …)

# Related Work

Agricultural satellite image time series (SITS) classification


**Whole-image methods**

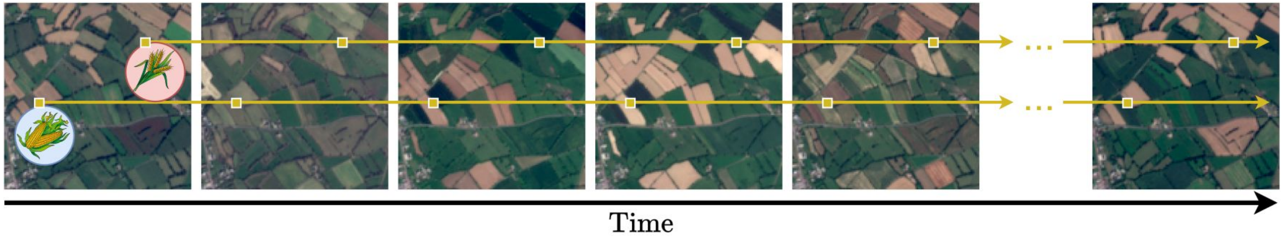
- Explicitly leverage the image structure
- U-Net + temporal aggregation (3D-Unet)
- U-Net + temporal attention encoder  (UTAE)

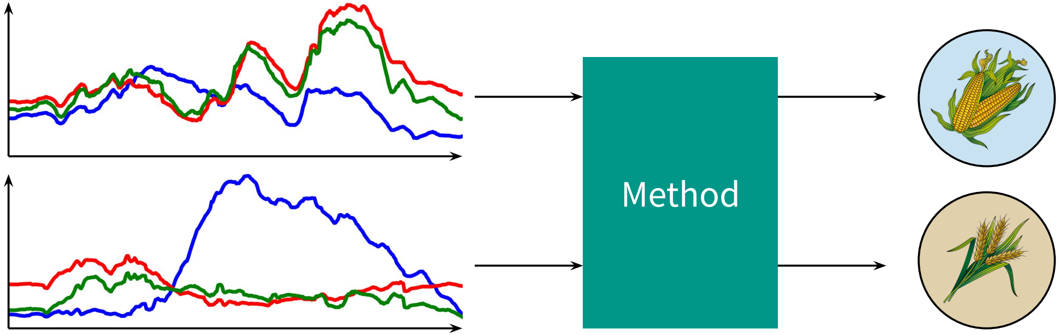
Designed for SITS

R. Rustowicz *et al. Semantic segmentation of crop type in Africa: A novel dataset and analysis of deep learning methods*. CVPR workshops 2019.
V. Sainte Fare Garnot *et al. Panoptic segmentation of satellite image time series with convolutional temporal attention networks*. CVPR 2021.

# Related Work

**Agricultural satellite image time series (SITS) classification**




**Time series-based methods**

- Whole-series based (1NN, NCC)
- Feature based (BoP, shapelet based, deep encoders)

Not necessarily designed for SITS specifically

→ generic methods for multivariate time series classification (MTSC)

V. Sainte Fare Garnot *et al*. *Lightweight temporal self-attention for classifying satellite images time series*. AALTD 2020.
W. Tang *et al. Omni-scale CNNs: a simple and effective kernel size configuration for time series classification.* ICLR 2022.

# Related Work

| Methods introduced so far → Supervised |
| --- |
| - require vast amount of labeled data<br>- low interpretability |

# Related Work

Methods introduced so far → Supervised

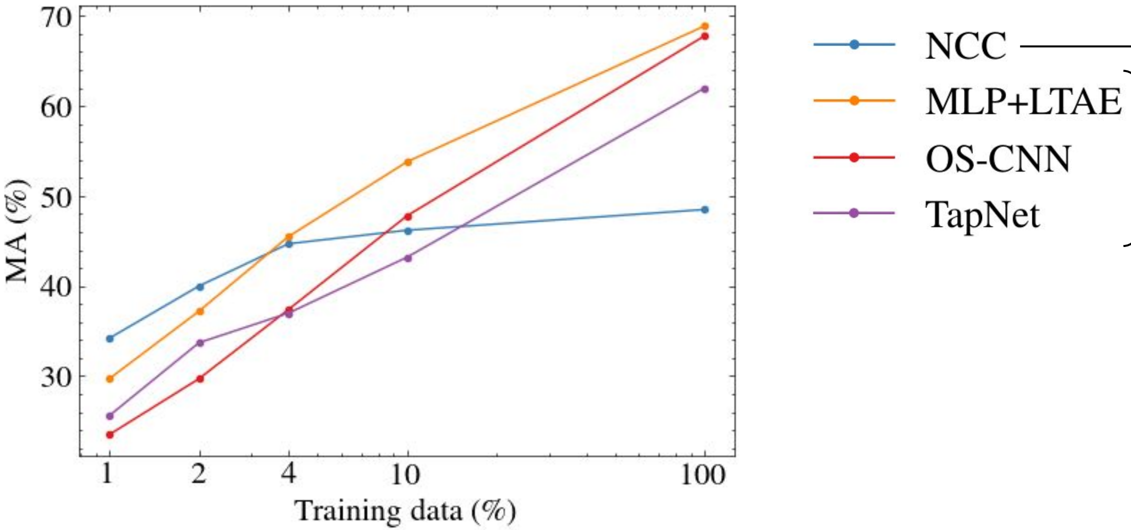
- require vast amount of labeled data
- low interpretability

How do they perform in low-data regime on the crop-type classification task?

# Related Work

**Methods introduced so far → Supervised**

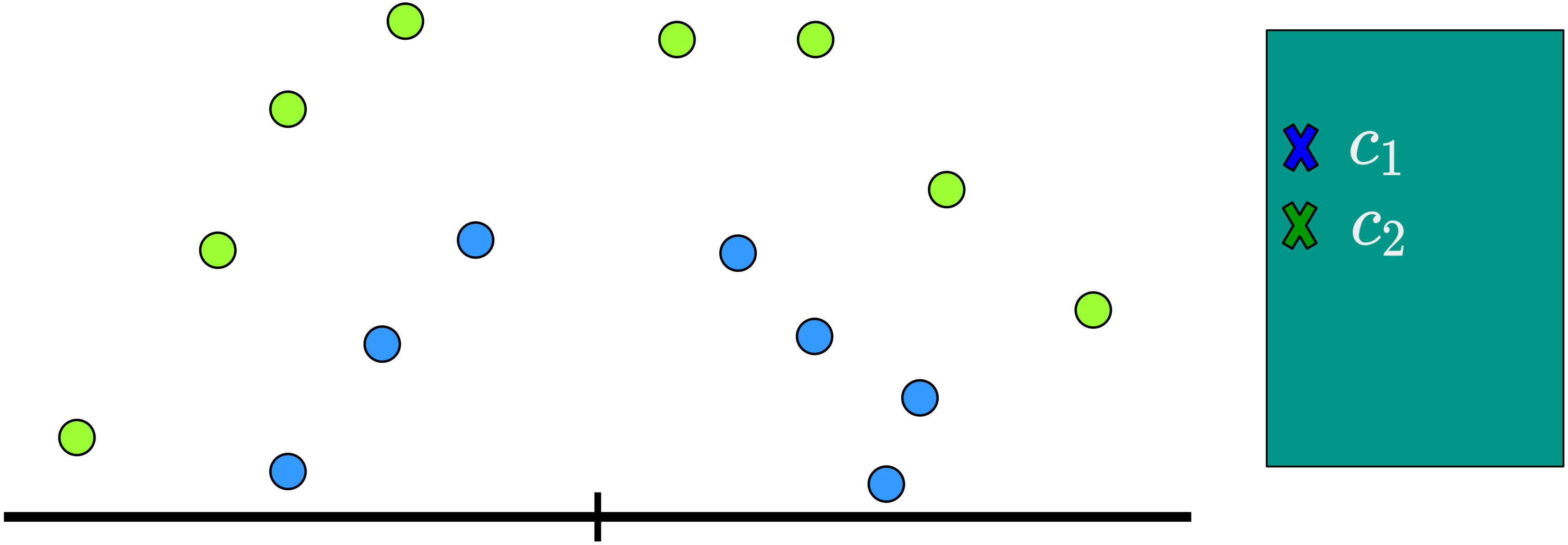
- require vast amount of labeled data
- low interpretability



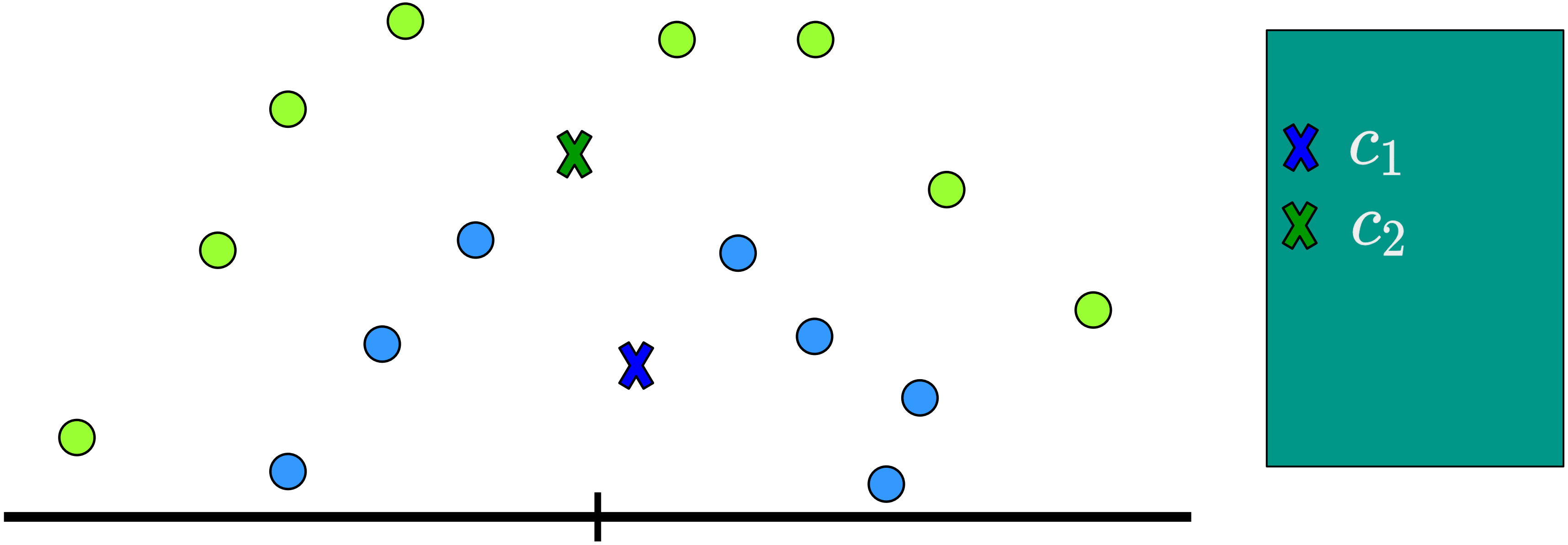

**Nearest centroid classifier**
Efficient in low-data regime

Best-performing MTSC methods on PASTIS dataset

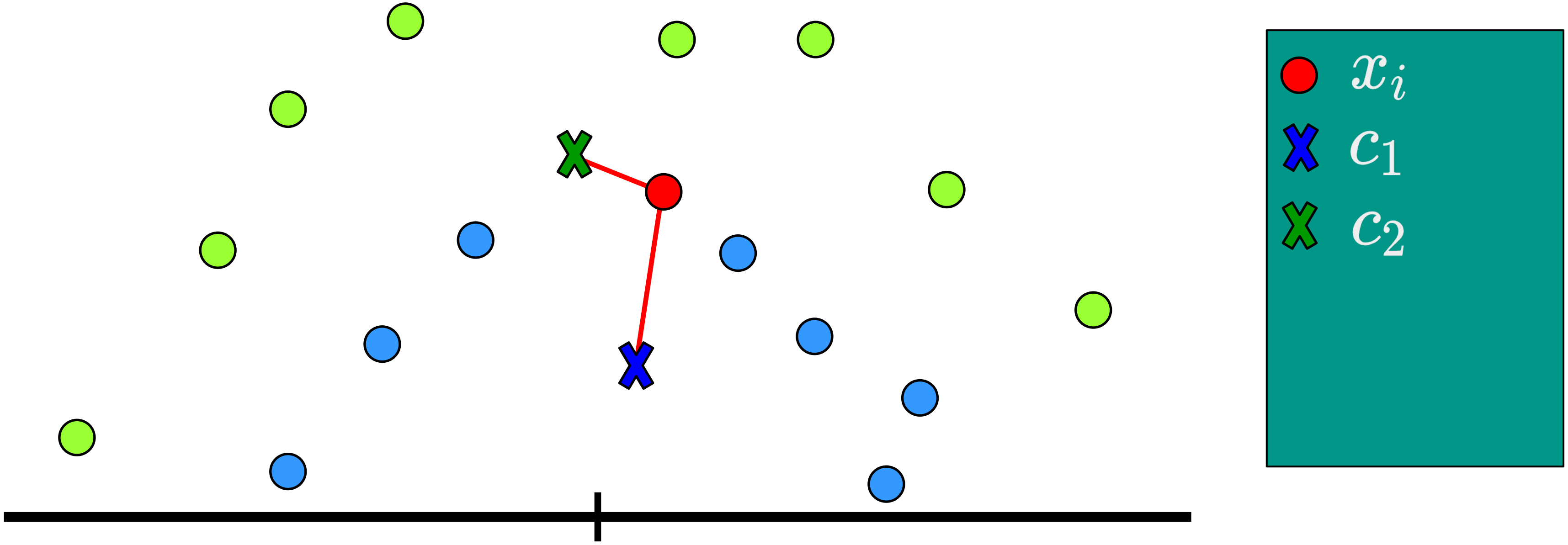
# Nearest centroid classifier



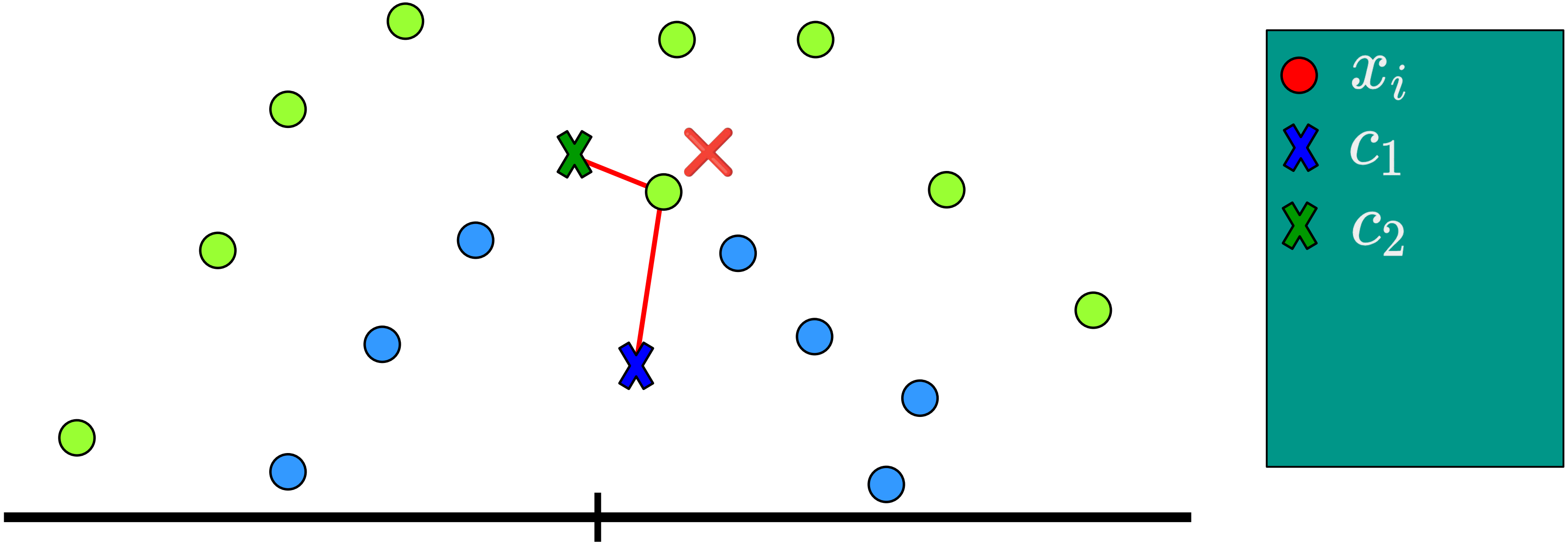
# Nearest centroid classifier



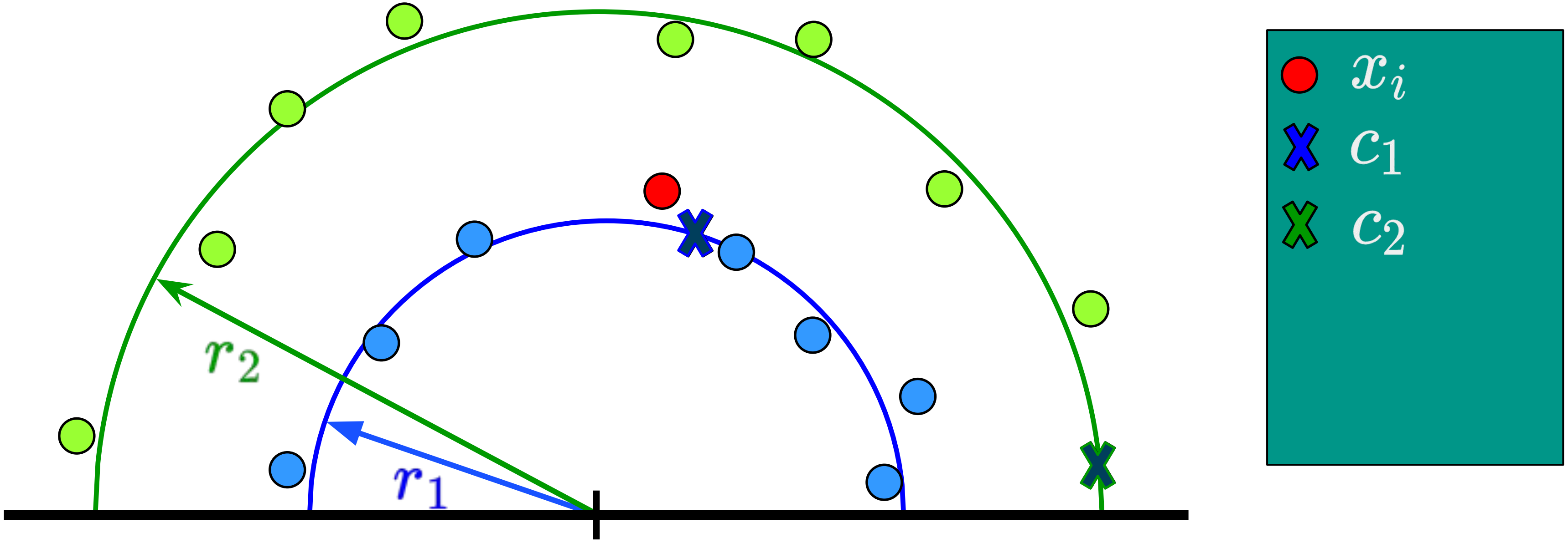
# Nearest centroid classifier



# Nearest centroid classifier

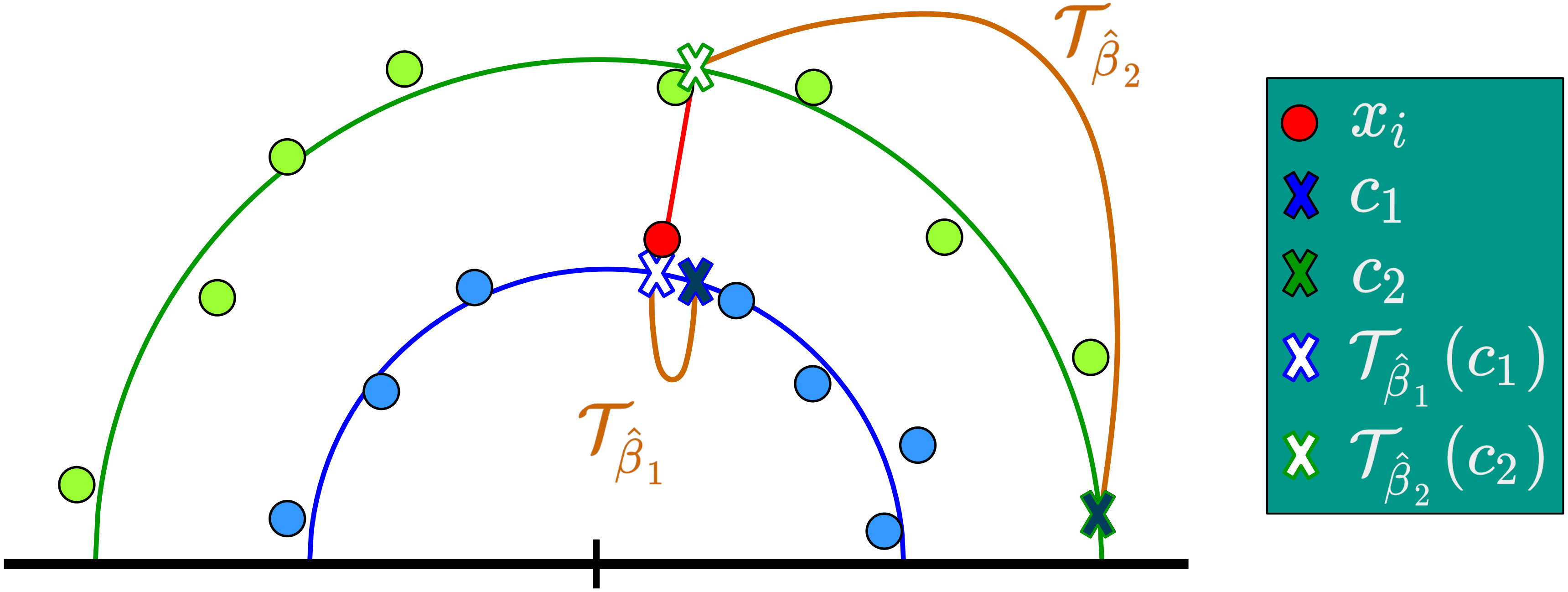


# Adding invariance to transformations



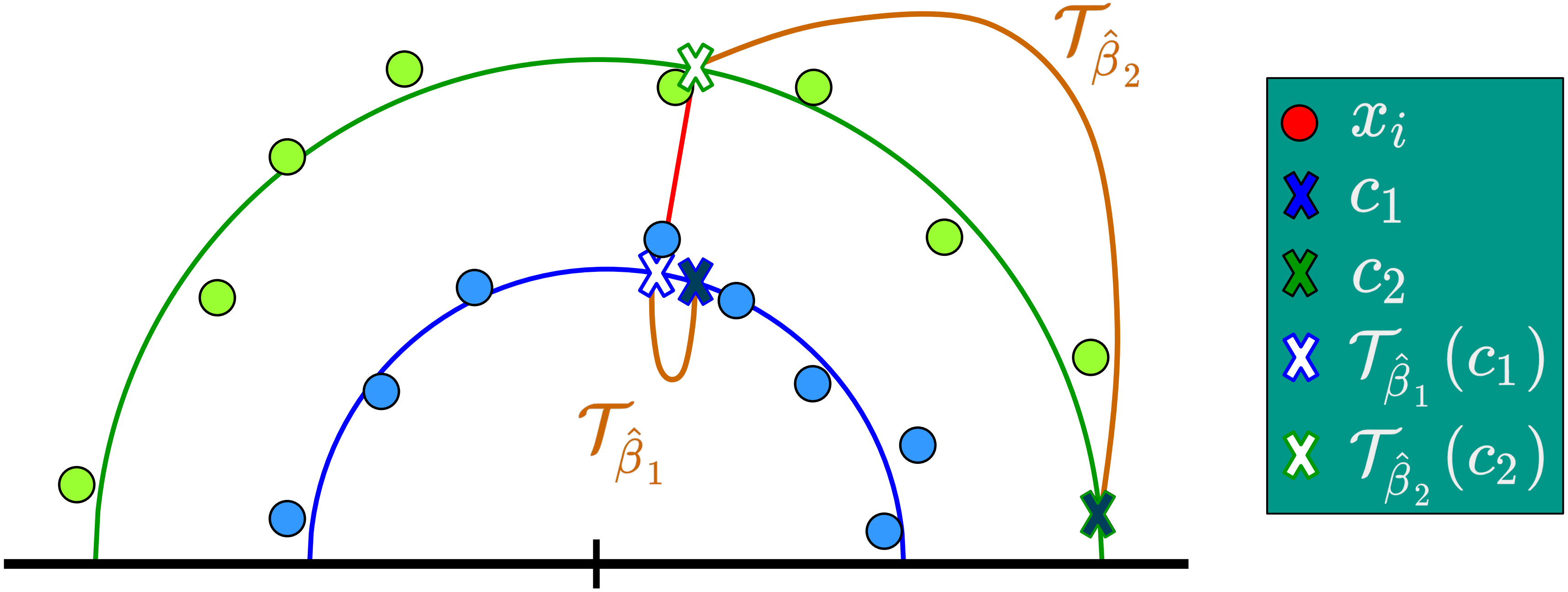

T. Monnier *et al*. *Deep Transformation-Invariant Clustering*. NeurIPS 2020.

# Adding invariance to transformations



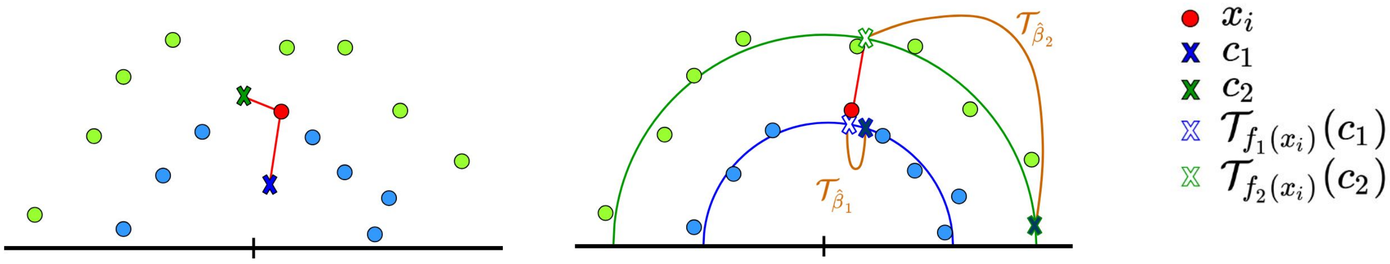

T. Monnier *et al*. *Deep Transformation-Invariant Clustering*. NeurIPS 2020.

# Adding invariance to transformations




T. Monnier *et al*. *Deep Transformation-Invariant Clustering*. NeurIPS 2020.

# Adding invariance to transformations



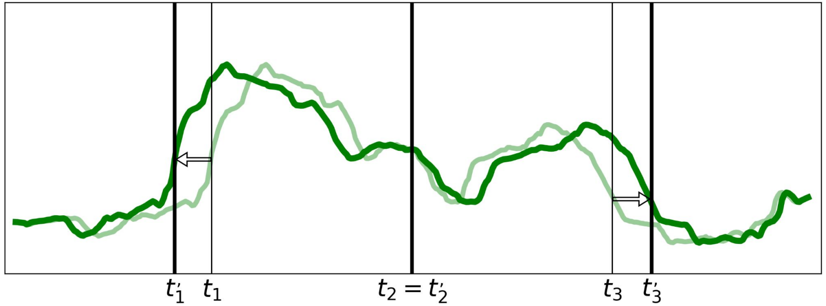

Classic NCC

Transf. invariant NCC

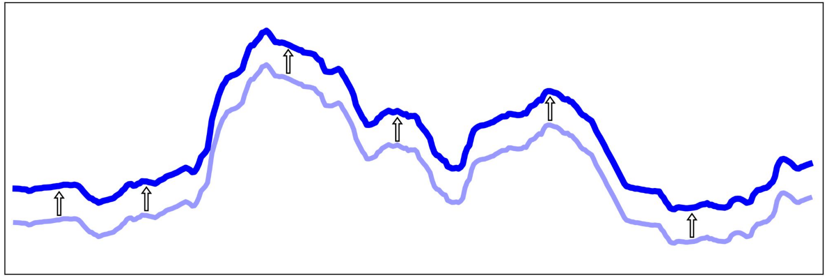
| Transformations for SITS |
| --- |
| - Time warping<br>- Offset |

# DTI-TS: Transformations

Time Warping




Offset



Time warping

1D TPS with control points

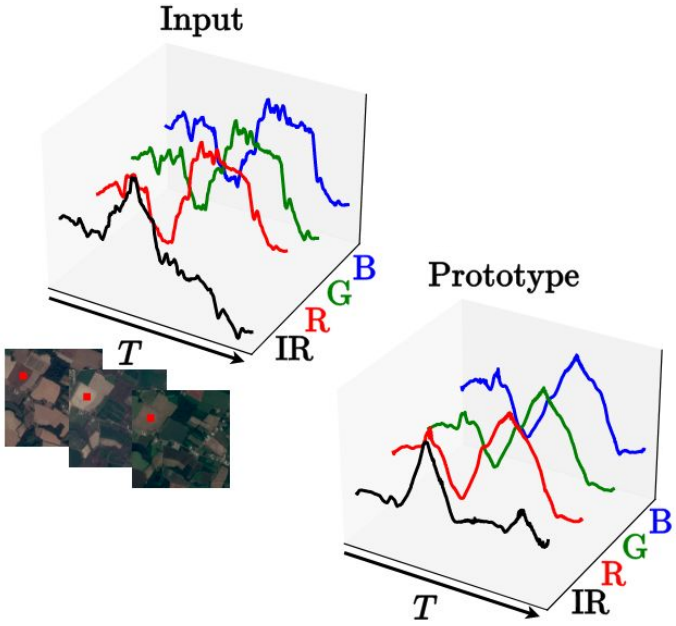
$$\mathcal{T}_k^{\mathrm{tw}}(x, \theta)$$

Offset

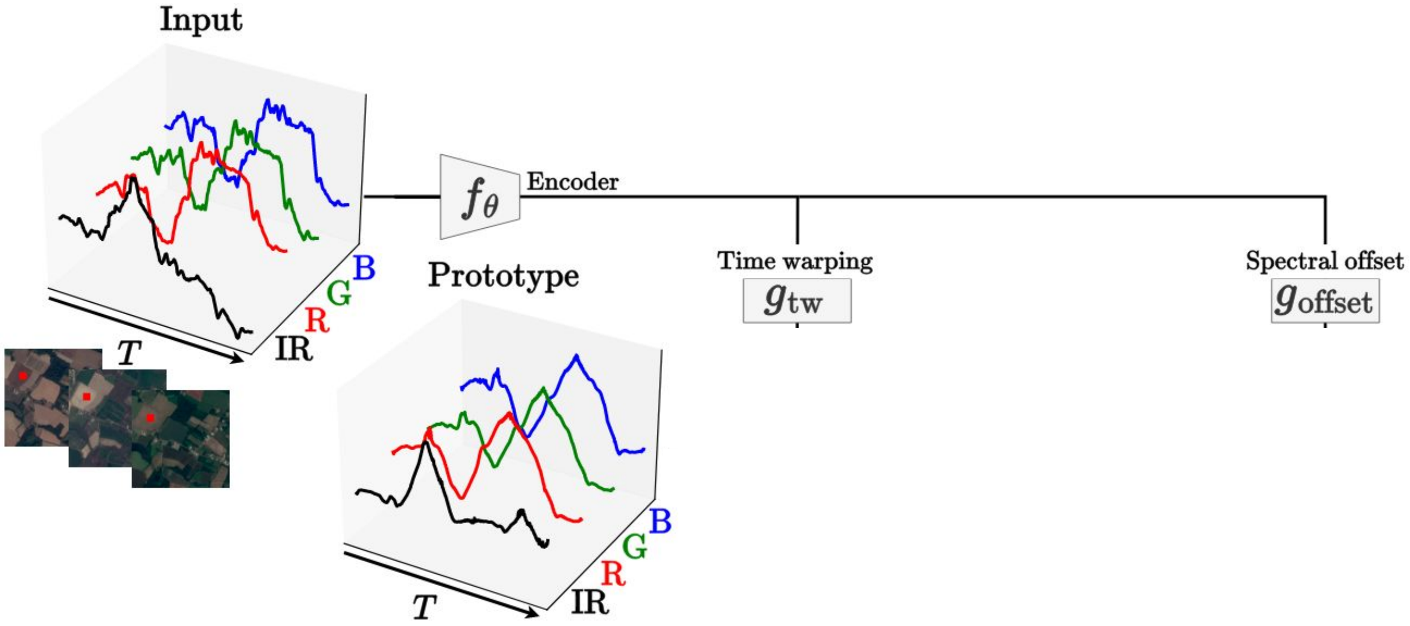
Channel wise

$$\mathcal{T}_k^{\mathrm{offset}}(x, \theta)$$

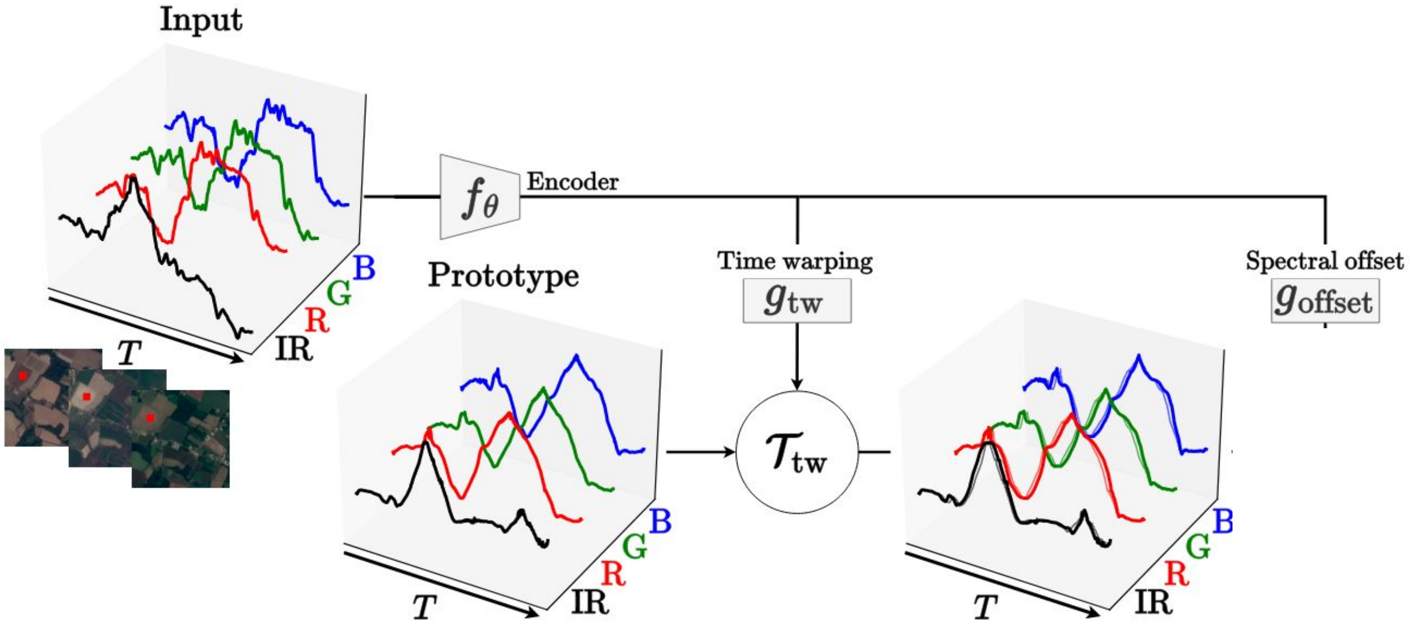
# DTI-TS: Overview



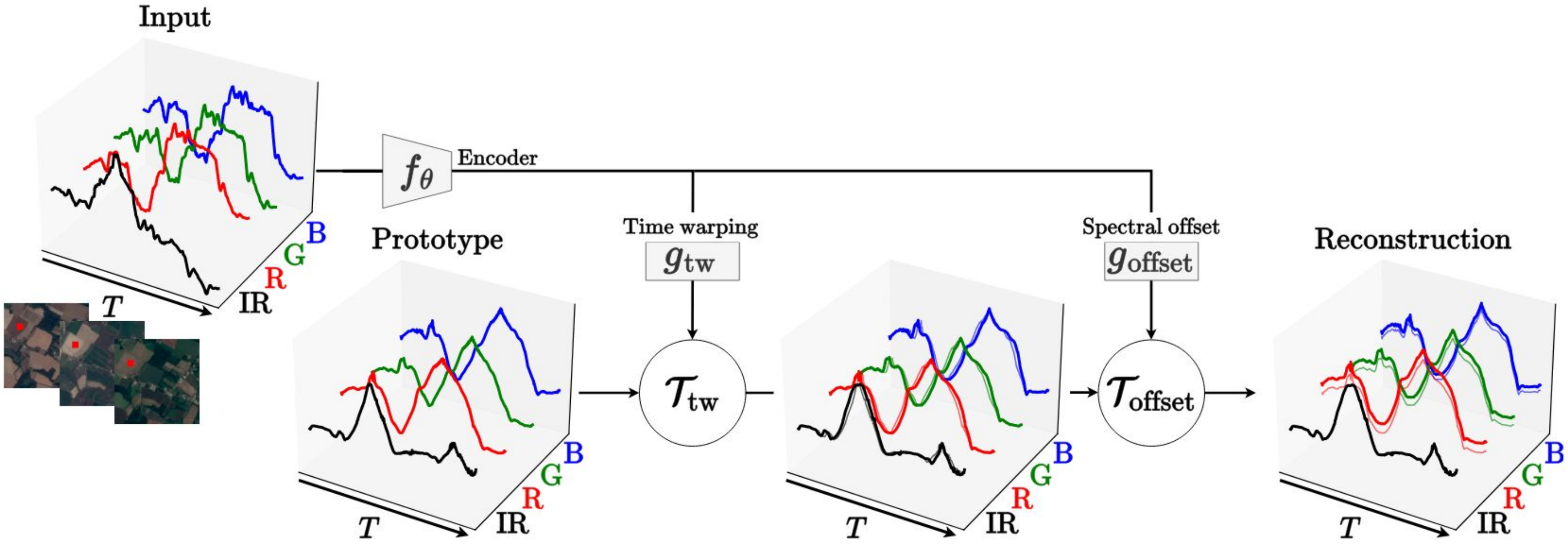
# DTI-TS: Overview



# DTI-TS: Overview



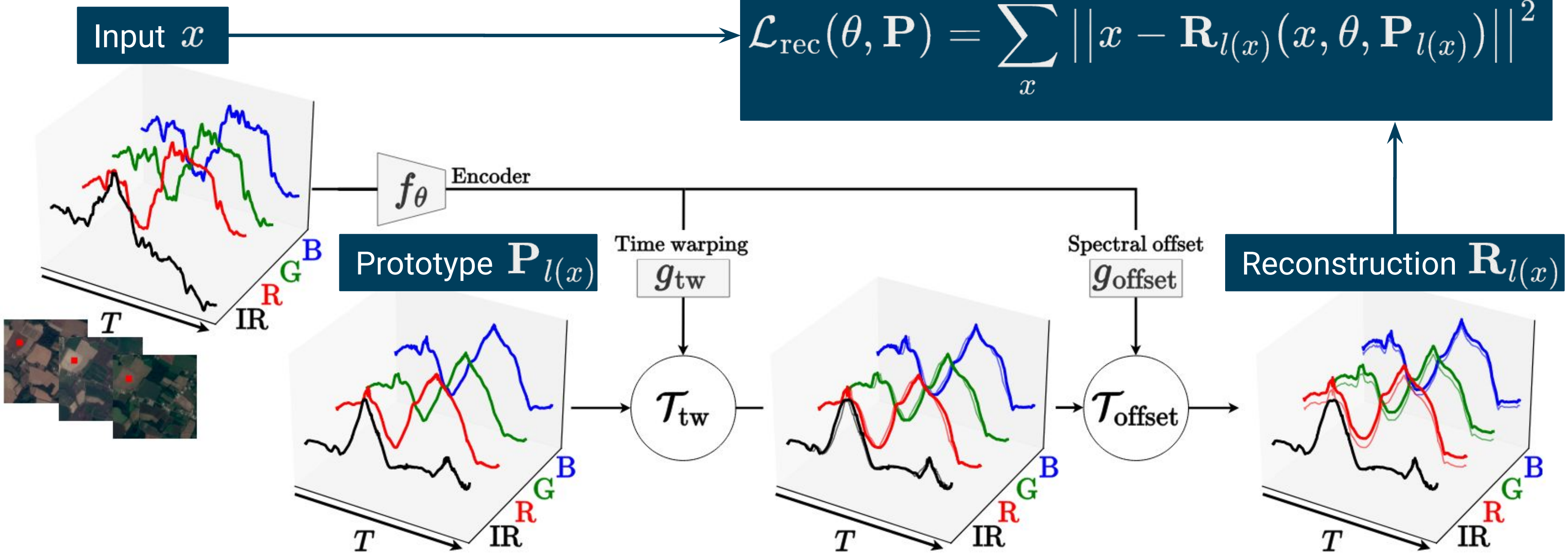
# DTI-TS: Overview



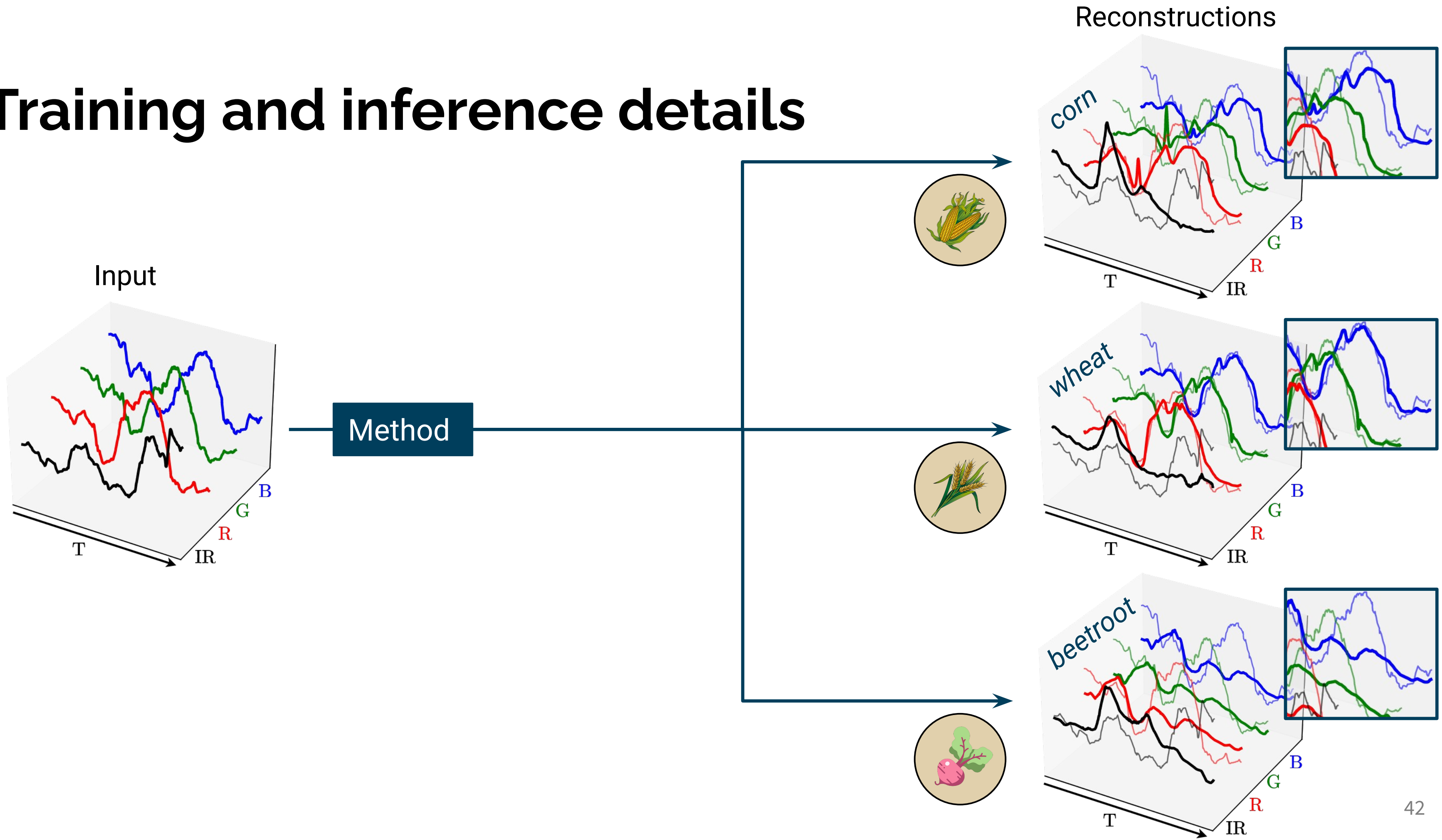
# DTI-TS: Overview

**Reconstruction loss**

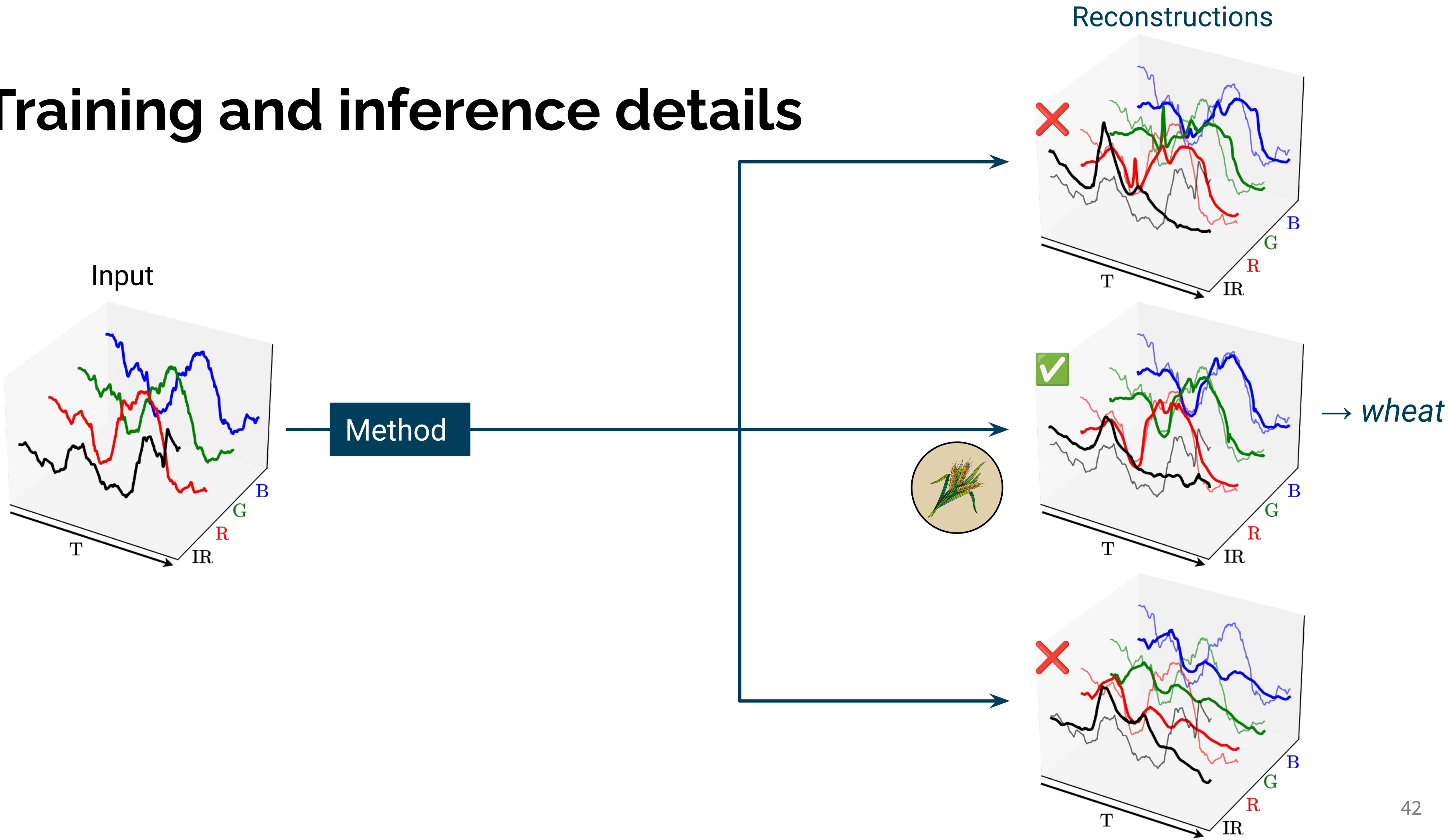
$$\mathcal{L}_{\text{rec}}(\theta, \mathbf{P}) = \sum_{x} \left|\left| x - \mathbf{R}_{l(x)}(x, \theta, \mathbf{P}_{l(x)}) \right|\right|^2$$



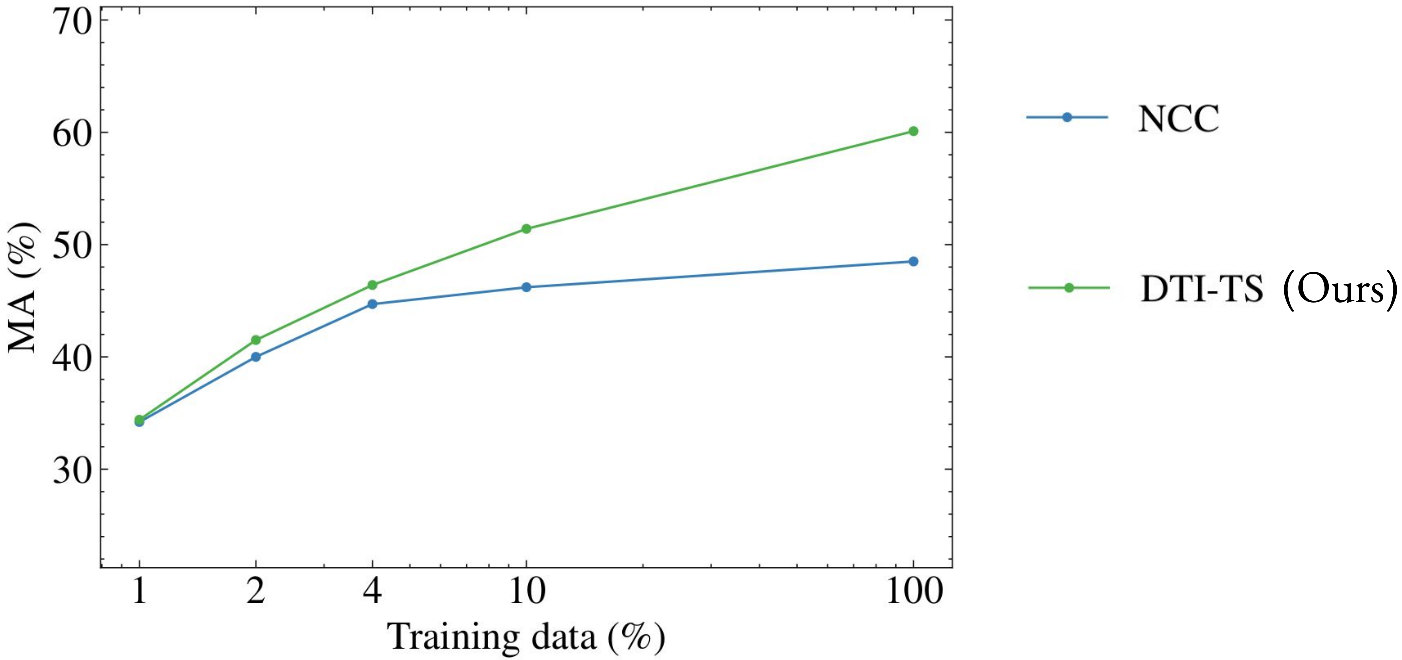
# Training and inference details



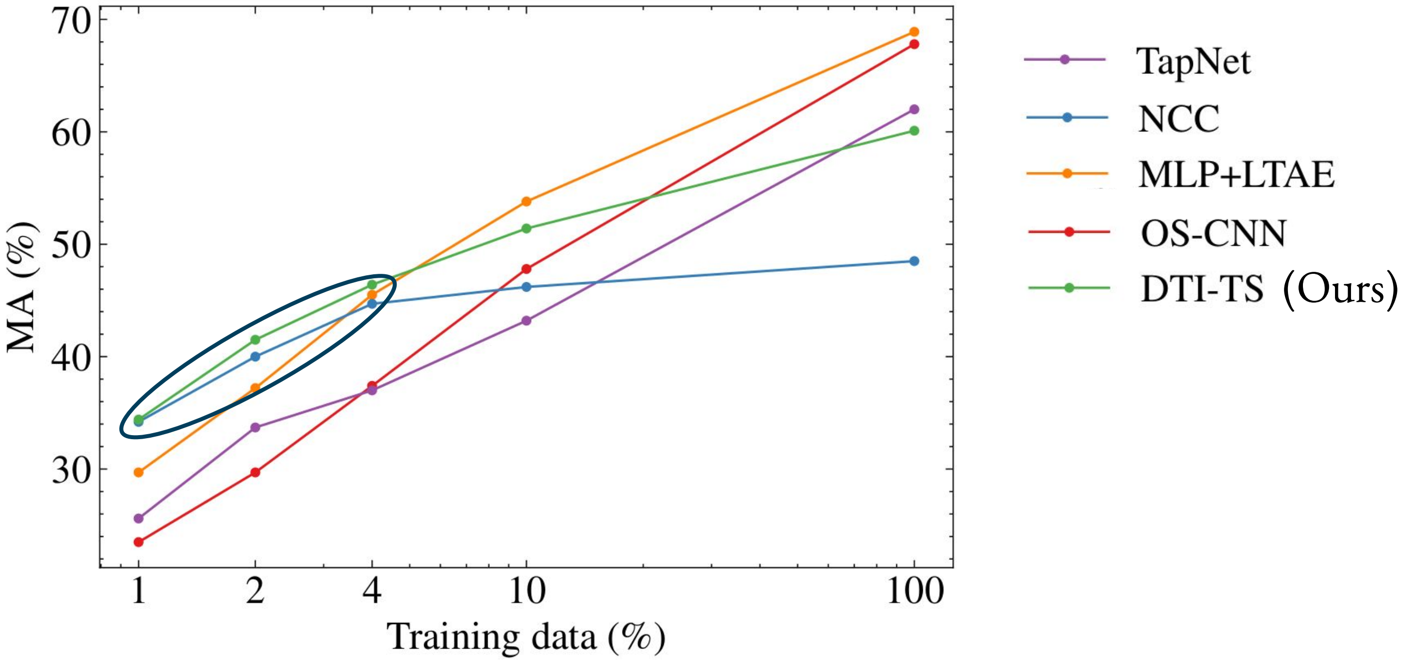
# Training and inference details



# Efficient in low-data regime



# Efficient in low-data regime



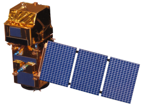
# Efficient in temporal shift settings

| Method | PASTIS MA↑ | TS2C MA↑ | SA MA↑ | DENET. MA↑ |
|---|---|---|---|---|
| MLP + LTAE |  |  |  |  |
| OS-CNN | | | | |
| TapNet | | | | |
| MLSTM-FCN | | | | |
| SVM | | | | |
| Random Forest | | | | |
| 1NN-DTW | | | | |
| 1NN | | | | |
| NCC | | | | |
| DTI-TS: NCC + Time warping + Offset | | | | |

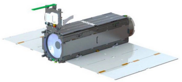
V. Sainte Fare Garnot et al. Panoptic segmentation of satellite image time series with convolutional temporal attention networks. ICCV, 2021.
G. Weikmann et al. Timesen2crop: A million labeled samples dataset of sentinel 2 image time series for crop-type classification. JSTARS, 2021.
L. Kondmann et al. Denethor: The dynamicearthnet dataset for harmonized, inter-operable, analysis-ready, daily crop monitoring from space. NeurIPS, 2021.
L. Kondmann et al. Early crop type classification with satellite imagery: an empirical analysis. ICLR Workshop, 2022.

# Efficient in temporal shift settings

| Method | PASTIS MA↑ | TS2C MA↑ | SA MA↑ | DENET. MA↑ |
|---|---|---|---|---|
| MLP + LTAE | | | | |
| OS-CNN | | | | |
| TapNet | | | | |
| MLSTM-FCN | | | | |
| SVM | | | | |
| Random Forest | | | | |
| 1NN-DTW | | | | |
| 1NN | | | | |
| NCC | | | | |
| DTI-TS: NCC + Time warping + Offset | | | | |



Sentinel 2



PlanetScope

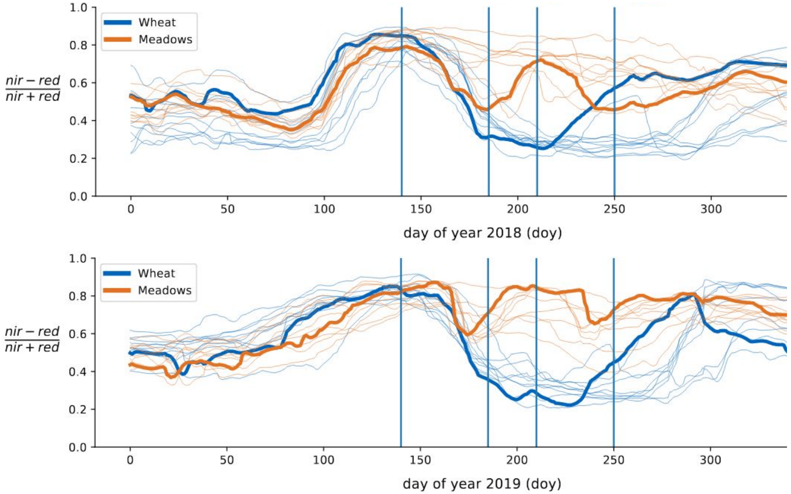

V. Sainte Fare Garnot et al. Panoptic segmentation of satellite image time series with convolutional temporal attention networks. ICCV, 2021.
G. Weikmann et al. Timesen2crop: A million labeled samples dataset of sentinel 2 image time series for crop-type classification. JSTARS, 2021.
L. Kondmann et al. Denethor: The dynamicearthnet dataset for harmonized, inter-operable, analysis-ready, daily crop monitoring from space. NeurIPS, 2021.
L. Kondmann et al. Early crop type classification with satellite imagery: an empirical analysis. ICLR Workshop, 2022.

# Efficient in temporal shift settings

| Method | PASTIS MA↑ | TS2C MA↑ | SA MA↑ | DENET. MA↑ |
|---|---|---|---|---|
| MLP + LTAE | | | | |
| OS-CNN | | | | |
| TapNet | | | | |
| MLSTM-FCN | | | | |
| SVM | | | | |
| Random Forest | | | | |
| 1NN-DTW | | | | |
| 1NN | | | | |
| NCC | | | | |
| DTI-TS: NCC + Time warping + Offset | | | | |

No domain shift

Temporal domain shift




L. Kondmann et al. Denethor: The dynamicearthnet dataset for harmonized, inter-operable, analysis-ready, daily crop monitoring from space. NeurIPS, 2021.

# Efficient in temporal shift settings

| Method | PASTIS MA↑ | TS2C MA↑ | SA MA↑ | DENET. MA↑ |
|---|---|---|---|---|
| MLP + LTAE | 65.9 | 80.9 | **63.7** | 43.6 |
| OS-CNN | **68.1** | **81.2** | 60.3 | 39.2 |
| TapNet | 60.3 | 77.3 | 56.7 | 43.7 |
| MLSTM-FCN | 10.9 | 44.0 | 47.9 | 48.3 |
| SVM | 48.7 | 56.1 | 52.8 | 28.6 |
| Random Forest | 46.6 | 50.2 | 61.3 | 51.6 |
| 1NN-DTW | — | 23.0 | — | — |
| 1NN | 40.1 | 35.0 | 54.9 | 48.2 |
| NCC | 48.4 | 49.9 | 46.4 | 55.5 |
| DTI-TS: NCC + Time warping | 51.4 | 52.3 | 49.7 | 56.4 |
| + Offset | 53.8 | 55.0 | 50.0 | **62.9** |

No domain shift

Temporal domain shift

# Can also be trained without supervision

**Supervised**

$$\mathcal{L}_{\text{rec}}(\theta, \mathbf{P}) = \sum_x \left|\left| x - \mathbf{R}_{l(x)}(x, \theta, \mathbf{P}_{l(x)}) \right|\right|^2$$

**Unsupervised**

$$\mathcal{L}_{\text{rec}}(\theta, \mathbf{P}) = \sum_x \min_k \left|\left| x - \mathbf{R}_k(x, \theta, \mathbf{P}_k) \right|\right|^2$$

# Can also be trained without supervision

**Supervised**

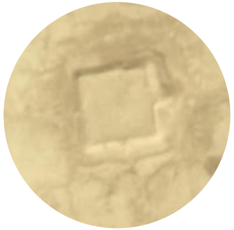
$$\mathcal{L}_{\text{rec}}(\theta, \mathbf{P}) = \sum_x \left|\left| x - \mathbf{R}_{l(x)}(x, \theta, \mathbf{P}_{l(x)}) \right|\right|^2$$

**Unsupervised**

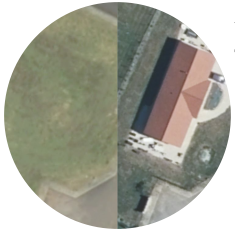
$$\mathcal{L}_{\text{rec}}(\theta, \mathbf{P}) = \sum_x \min_k \left|\left| x - \mathbf{R}_{k}(x, \theta, \mathbf{P}_{k}) \right|\right|^2$$

| Method | PASTIS MA↑ | TS2C MA↑ | SA MA↑ | DENET. MA↑ |
|---|---|---|---|---|
| K-means-DTW | — | 26.8 | — | — |
| USRL+K-means | 20.4 | 23.6 | 48.6 | 46.4 |
| DTAN+K-means | 21.4 | 29.3 | 48.6 | 36.9 |
| K-means | 29.8 | 32.5 | 47.8 | 48.5 |
| DTI-TS: K-means + Time warping | **30.4** | **36.0** | **51.7** | 51.1 |
| + Offset | 28.6 | 35.5 | 50.4 | **52.6** |

# Progress Recap



1 Afghan archaeological site looting detection



2 Semantic change detection and domain shift analysis

3 Crop-type classification with few or no annotations

✔ Learning with low to no data
✔ Efficiency in temporal shift settings

# Conclusion


✔ Providing labeled data for a specific task/location

✔ Making use of pre-trained off-the-shelf models

✔ Evaluating the impact of temporal/spatial shift

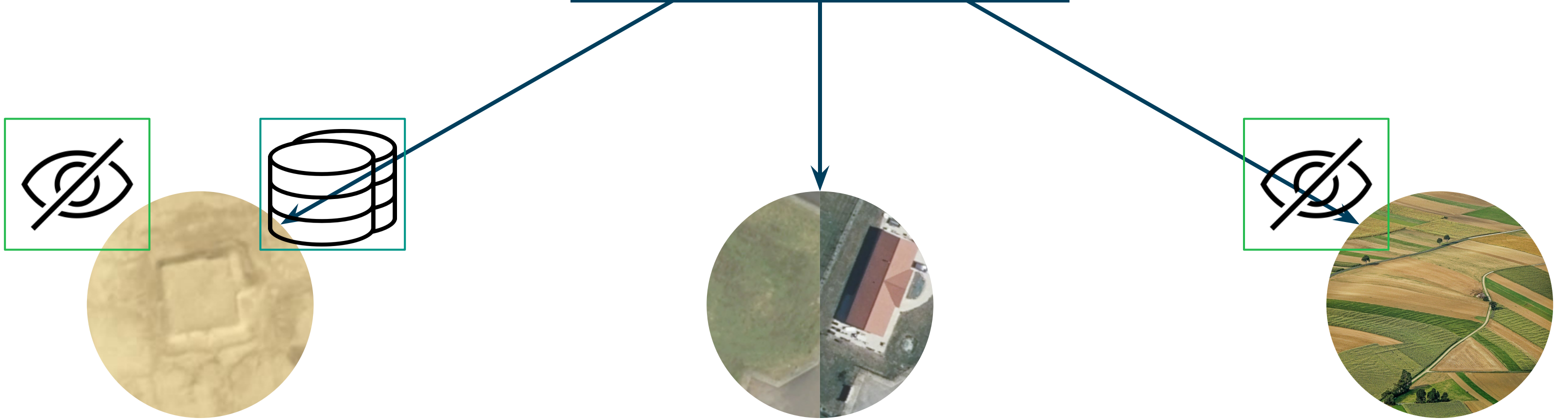
✔ Addressing spatial shift with domain experts

✔ Learning with low to no data

✔ Efficiency in temporal shift settings

# Conclusion

Scarcity of annotated data

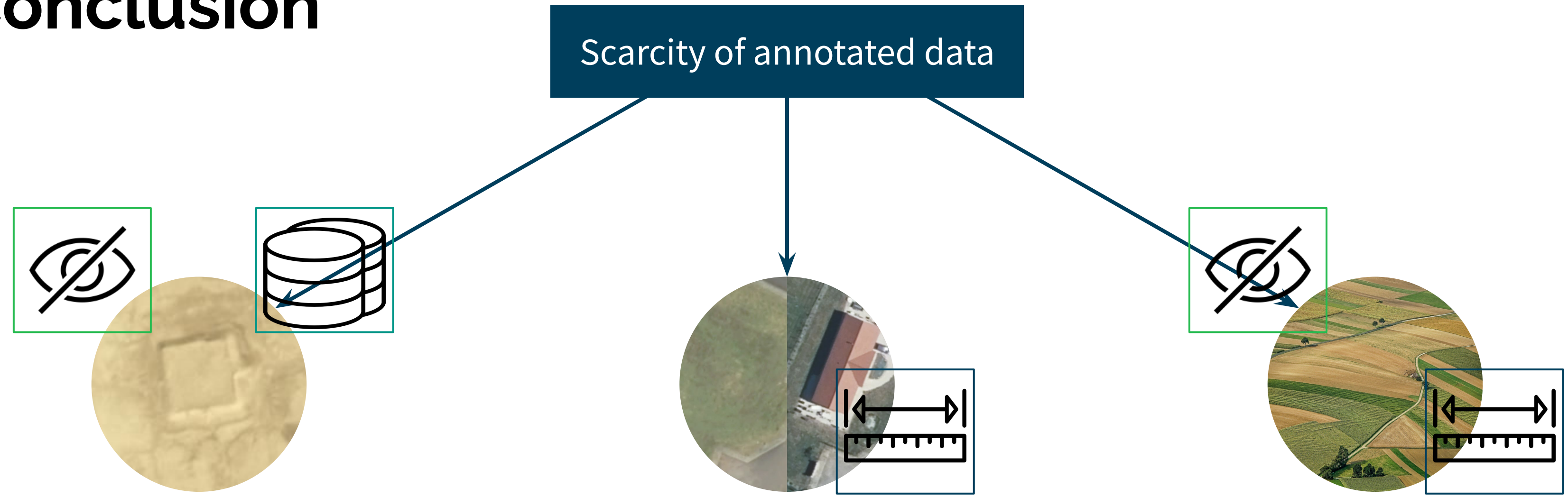


✔ Providing labeled data for a specific task/location
✔ Making use of pre-trained off-the-shelf models

✔ Evaluating the impact of temporal/spatial shift
✔ Addressing spatial shift with domain experts

✔ Learning with low to no data
✔ Efficiency in temporal shift settings

# Conclusion



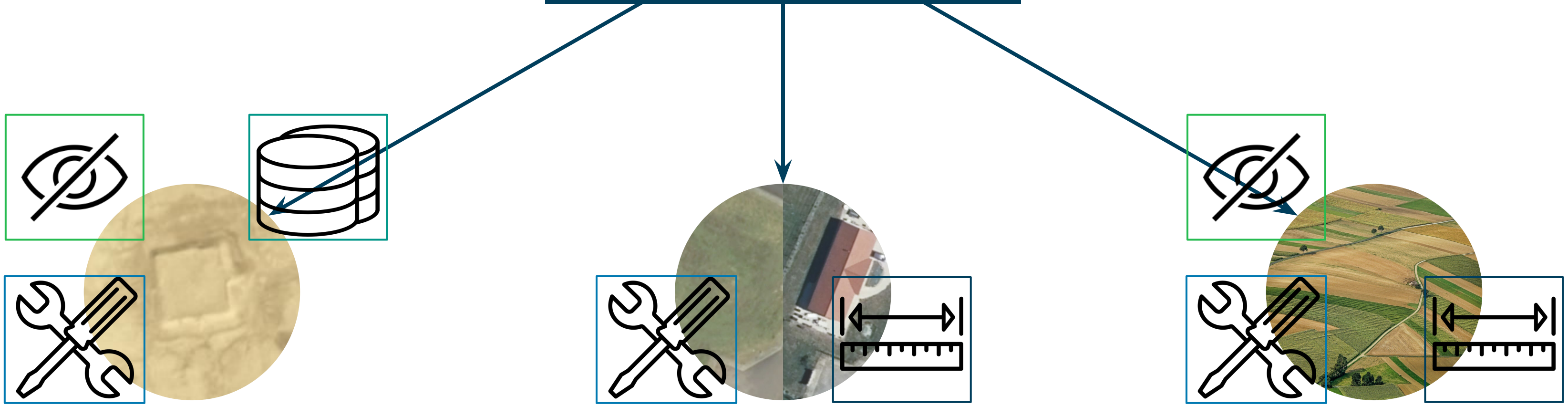

✔ Providing labeled data for a specific task/location
✔ Making use of pre-trained off-the-shelf models

✔ Evaluating the impact of temporal/spatial shift
✔ Addressing spatial shift with domain experts

✔ Learning with low to no data
✔ Efficiency in temporal shift settings

# Conclusion

Scarcity of annotated data



✔ Providing labeled data for a specific task/location
✔ Making use of pre-trained off-the-shelf models

✔ Evaluating the impact of temporal/spatial shift
✔ Addressing spatial shift with domain experts

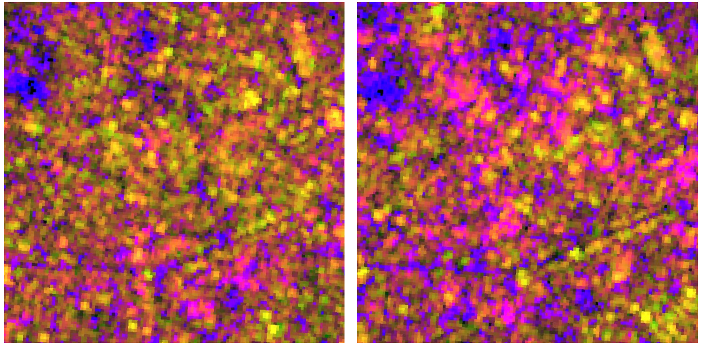
✔ Learning with low to no data
✔ Efficiency in temporal shift settings

# Future work

Towards increased **temporal** multimodality:

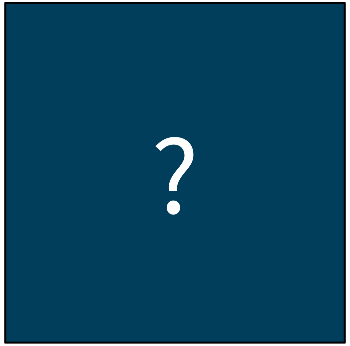
- sensors, resolutions

- radar data



- **3D data**

- ???

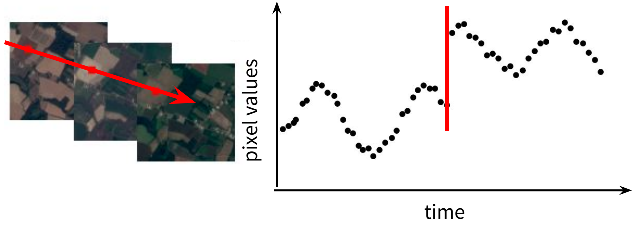


# Future work

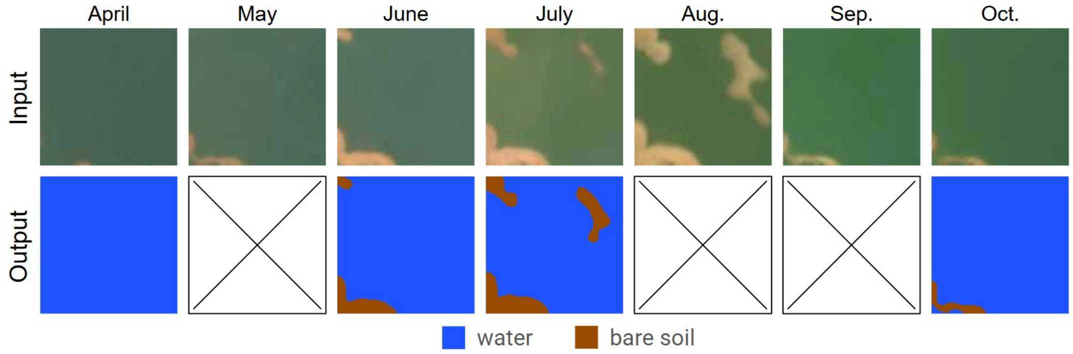
Towards increased **temporal** multimodality:

Improving **weakly-supervised** and **unsupervised** methods

Leveraging change point detection techniques




Learning from incomplete annotations



# Analysis of satellite image time series for classification and change detection

Elliot Vincent - May 27th, 2025

Committee:
Sébastien LEFEVRE
Jan Dirk WEGNER
Pauline LUC
Charlotte PELLETIER
Gabriele FACCIOLO
Mathieu AUBRY (advisor)
Jean PONCE (co-advisor)

Thanks to all my co-authors!