Pixel-wise Agricultural Image Time Series Classification: Comparisons and a Deformable Prototype-based Approach

Elliot Vincent - 31/03/2023
– IGN Reading group –

About Me

2nd year PhD student

Research interests: Unsupervised learning, Remote sensing, Prototype-based methods

Elliot Vincent [1,2]

Advisors

Mathieu Aubry [1]

Jean Ponce [3,4]

1 École des Ponts ParisTech

2 Inria

3 ENS | PSL

4 NYU

About Me

Previous projects → mainly, unsupervised prototype learning

- *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

- *Pixel-wise Agricultural Image Time Series Classification: Comparison and a Deformable Prototype-based Approach*
  E. Vincent, J. Ponce, M. Aubry – Preprint 2023







About Me

Previous projects → mainly, unsupervised prototype learning

- *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

- *Pixel-wise Agricultural Image Time Series Classification: Comparison and a Deformable Prototype-based Approach*
  E. Vincent, J. Ponce, M. Aubry – Preprint 2023

Today's program!

Task

Agricultural satellite image time series (SITS) classification

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

Context

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)

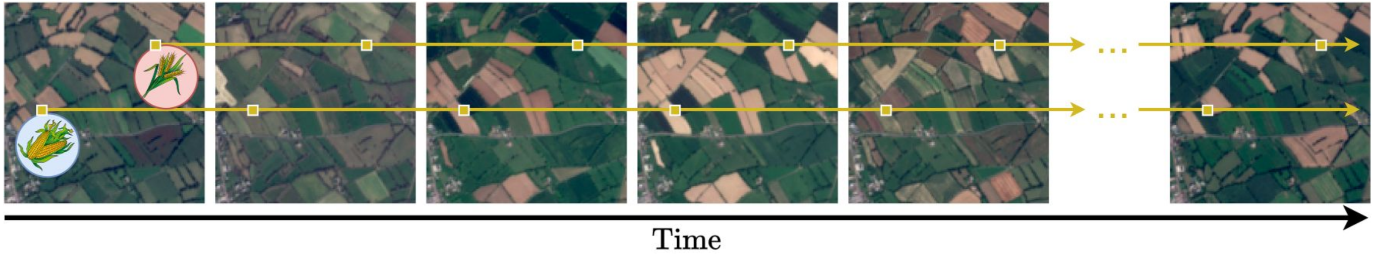
Designed for SITS

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.

Context

Agricultural satellite image time series (SITS) classification



Time series-based methods

- Whole-series based (1NN, prototype-based)
- 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.

Context

Methods introduced so far → Supervised

- require vast amount of labeled data
- low interpretability

What about unsupervised methods for MTSC?

- K-means (Euclidean distance, DTW)
- K-means on learned features
  (= Unsupervised representation learning + K-means)

F. Petitjean *et al*. *Clustering of satellite image time series under Time Warping*. MultiTemp 2011.

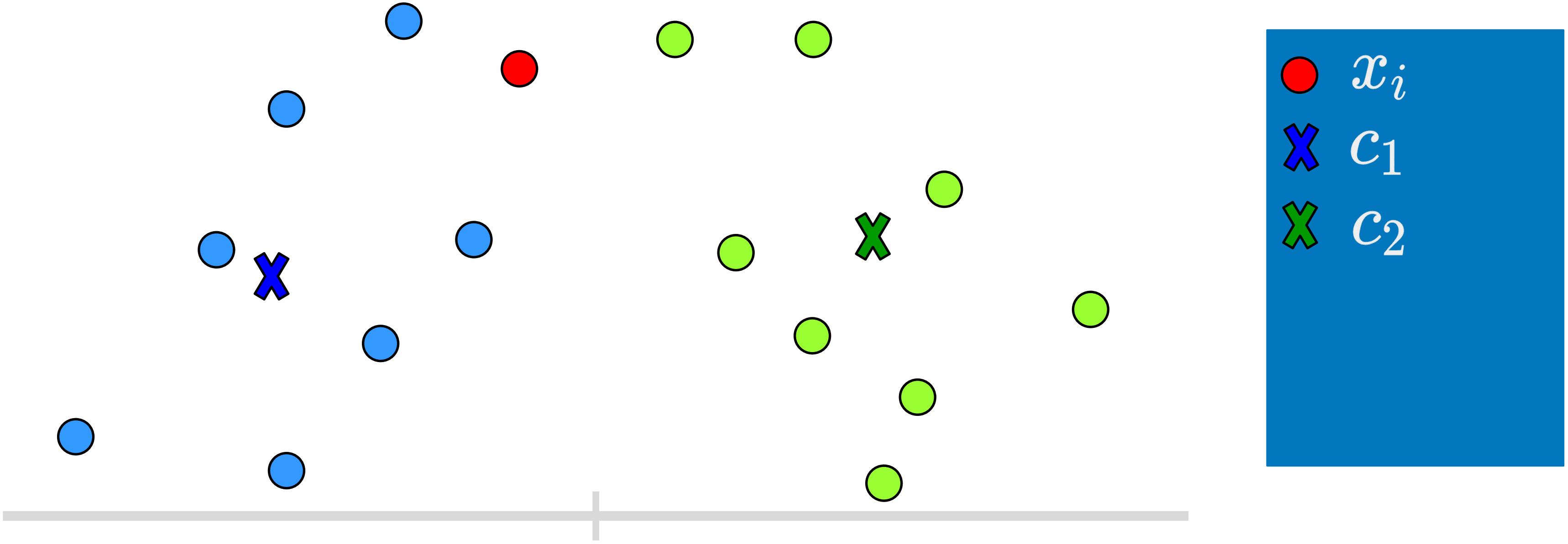
J.-Y. Franceschi *et al*. *Unsupervised scalable representation learning for multivariate time series*. NeurIPS 2019

Contributions

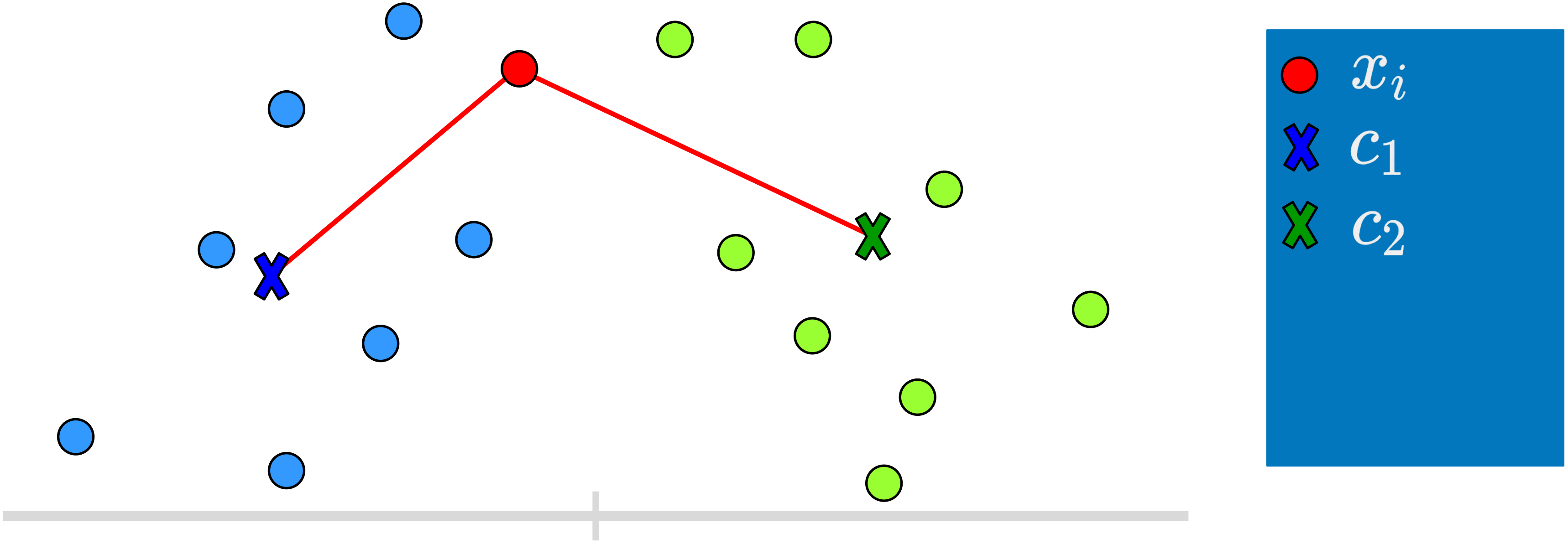
- Introduce a prototype-based method
  - Insisting on **unsupervised learning**
  - And **interpretability**

- **Benchmark** multivariate time series classification
  - **supervised** and **unsupervised**
  - **4 recent SITS datasets**

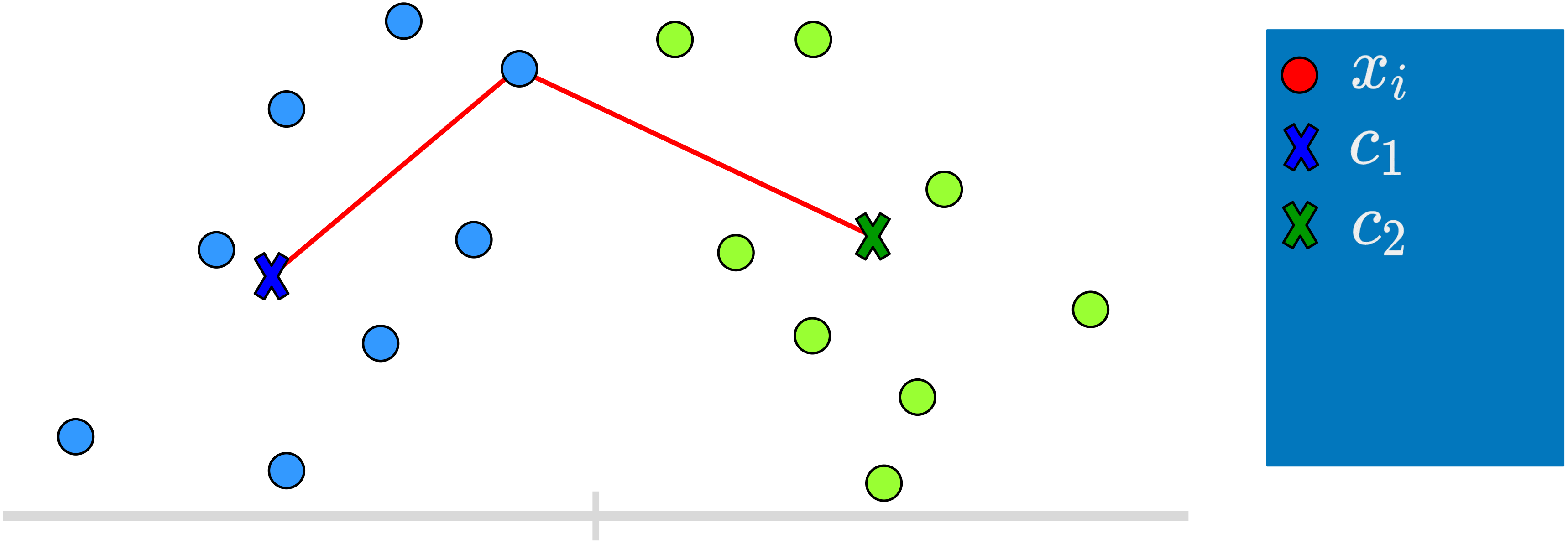
Classic clustering



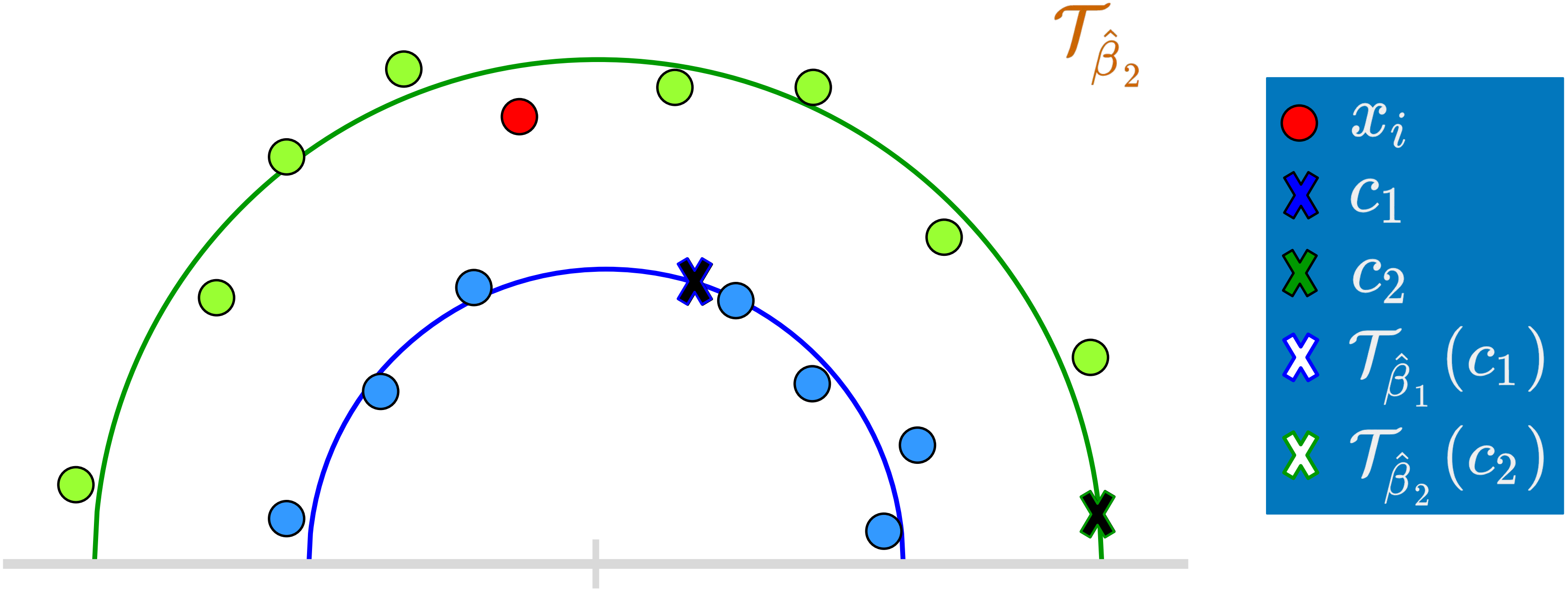
Classic clustering



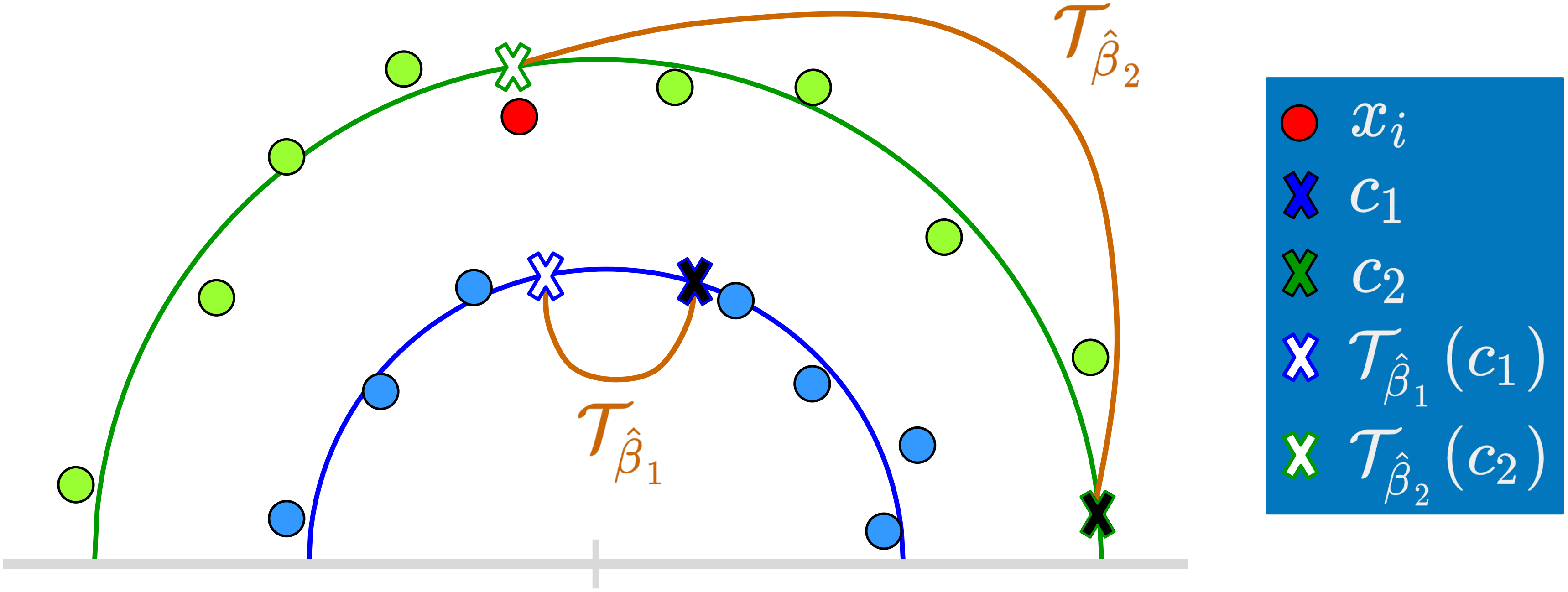
Classic clustering



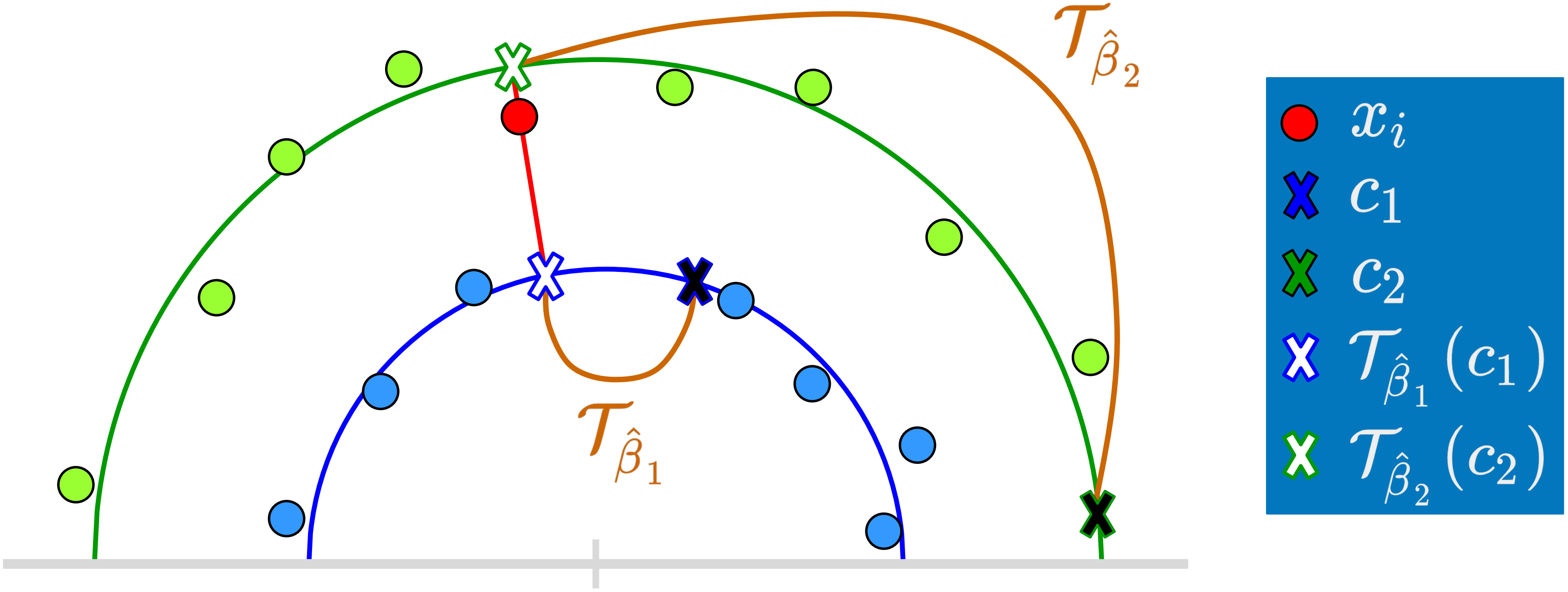
DTI clustering



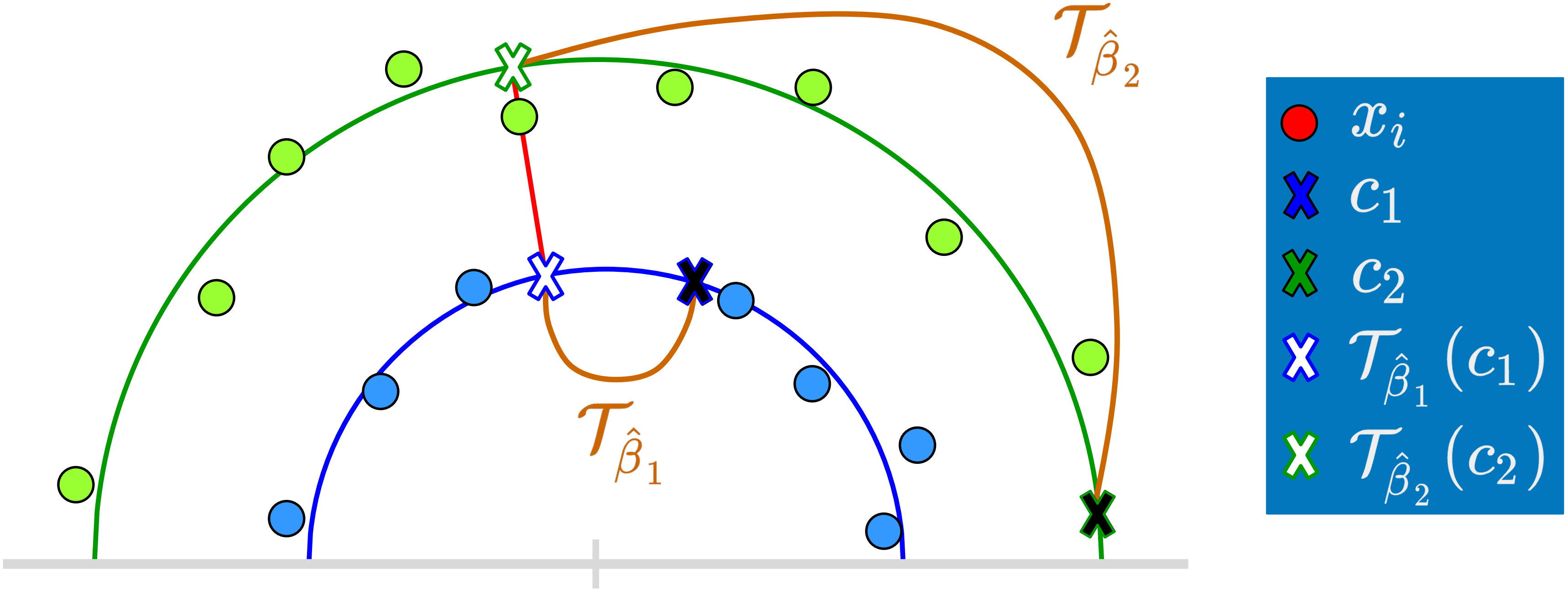
DTI clustering



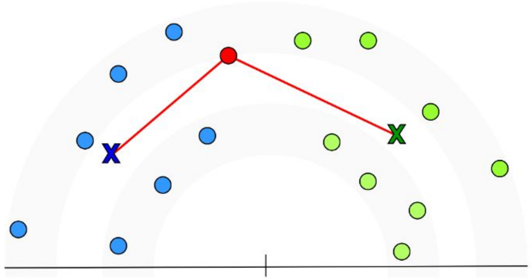
DTI clustering



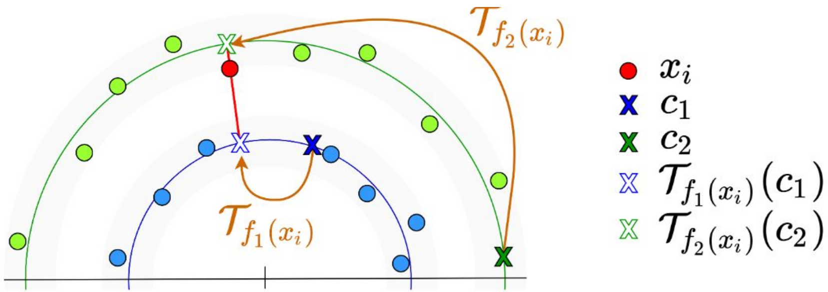
DTI clustering



DTI clustering



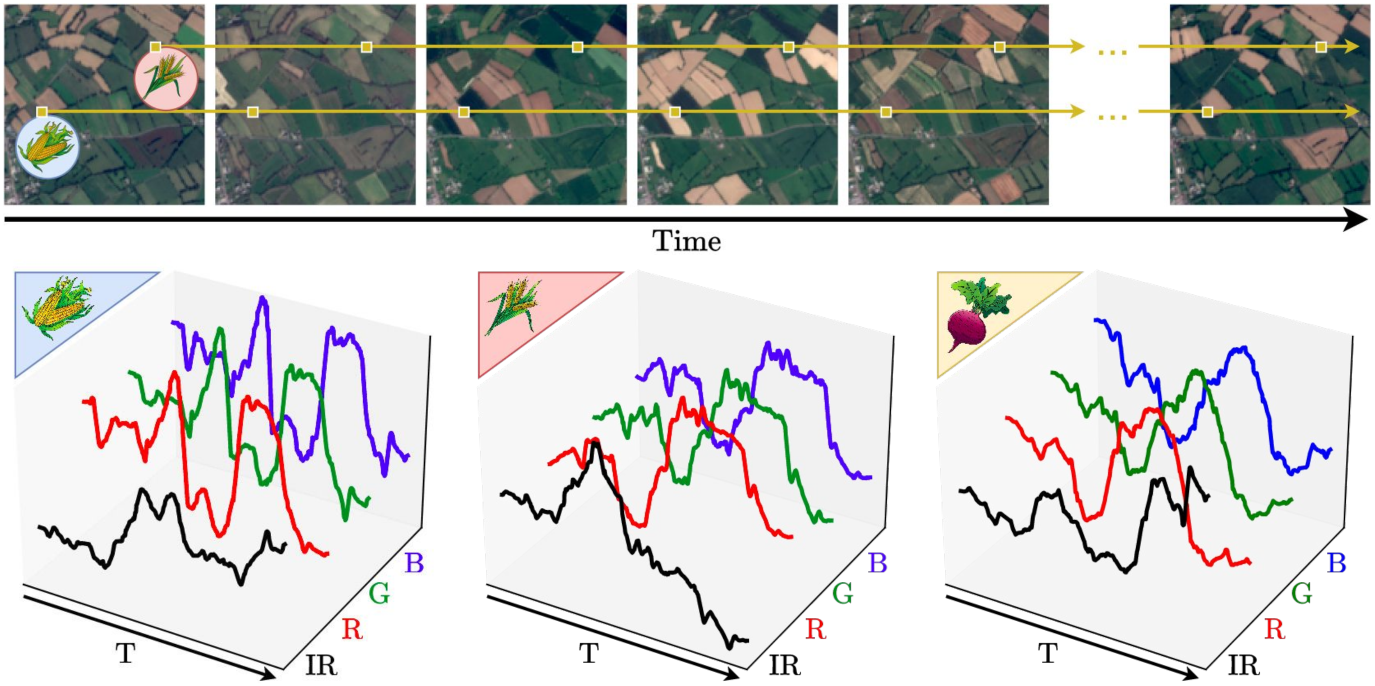
Classic clustering



DTI clustering

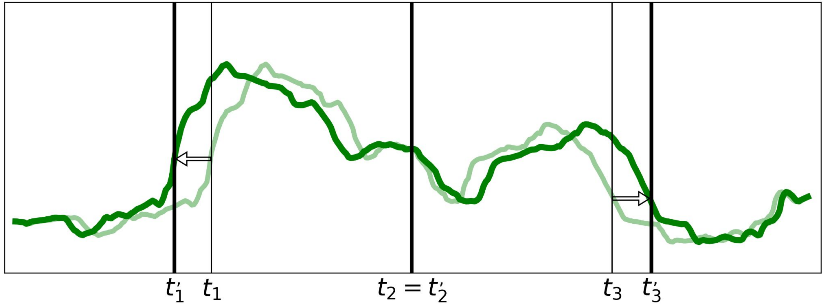
T. Monnier *et al*. *Deep Transformation-Invariant Clustering*. NeurIPS 2020.

Classic clustering

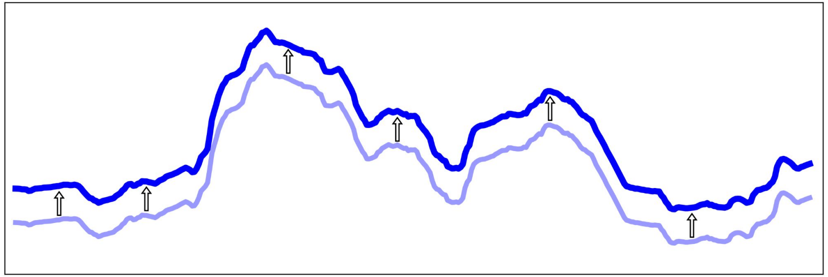


Our Method: Transformations

Time Warping



Offset



Time warping

1D TPS with control points

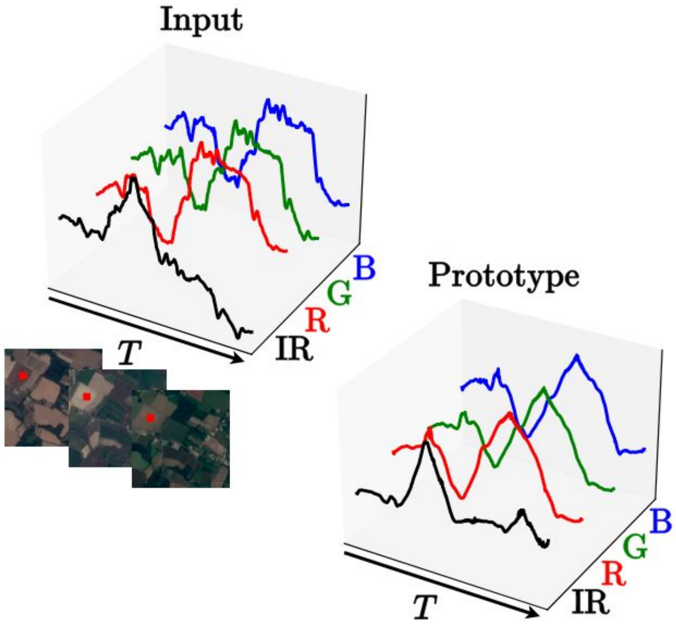
$$\mathcal{T}_k^{\mathrm{tw}}(x, \theta)$$

Offset

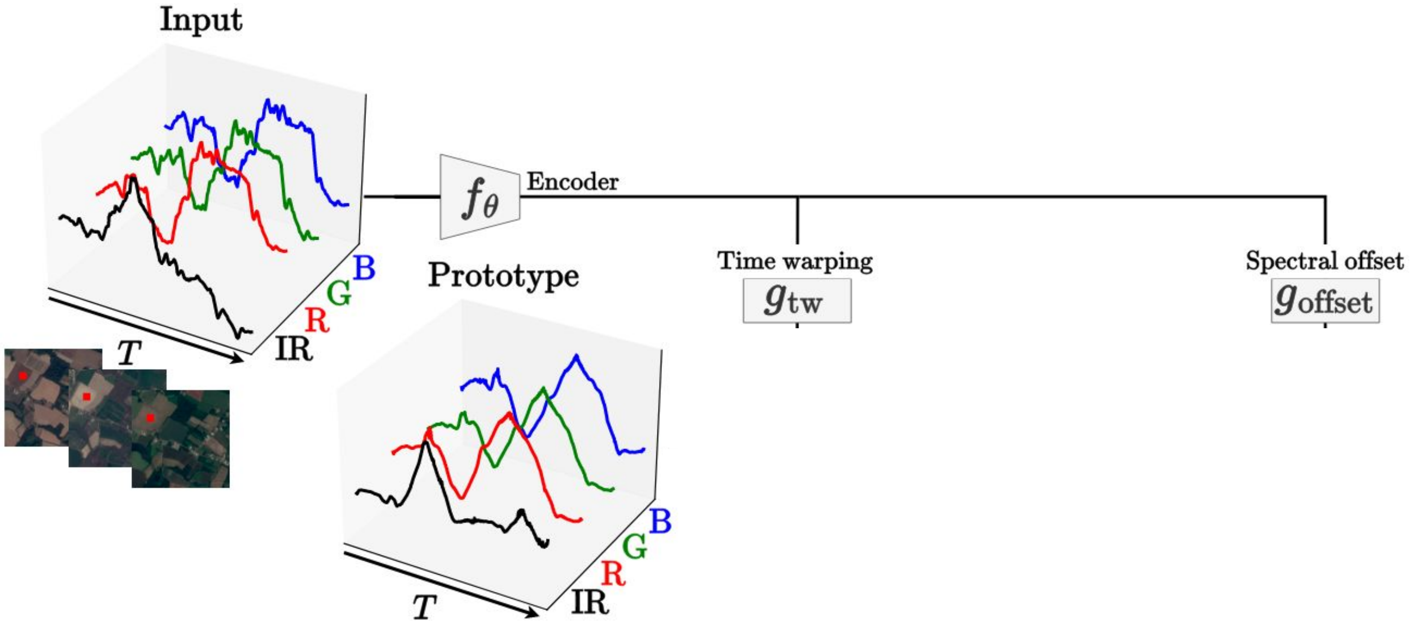
Channel wise

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

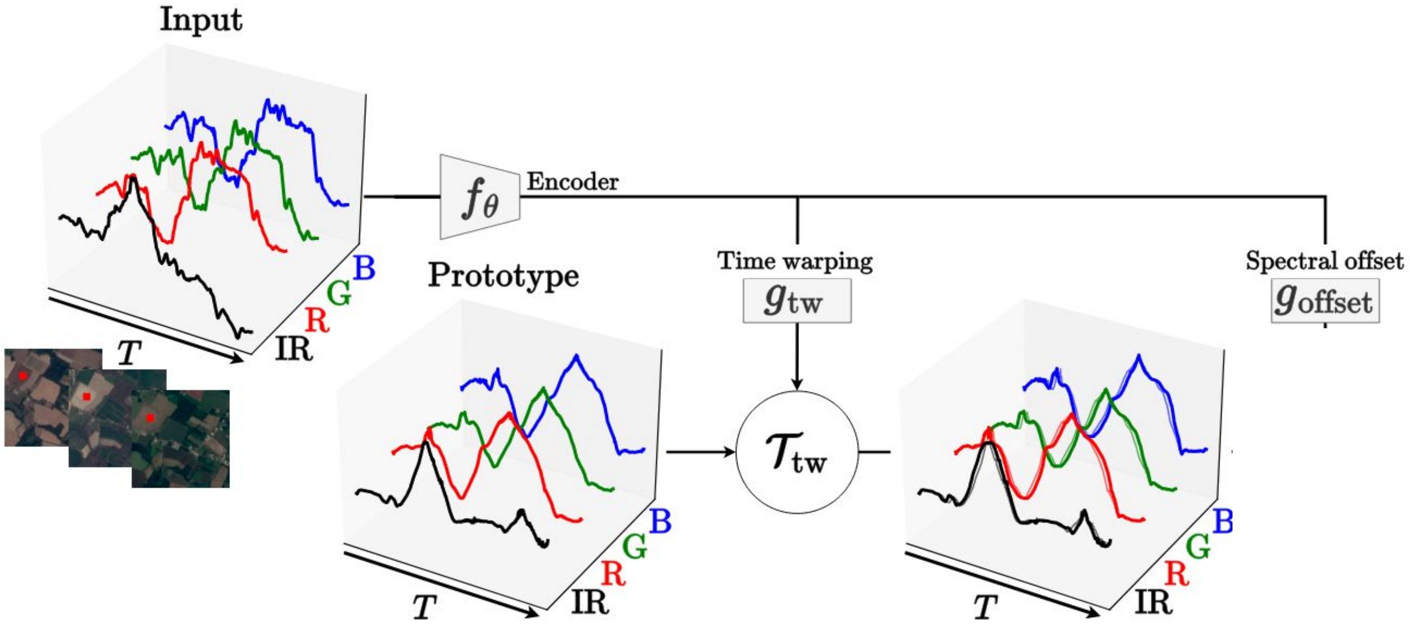
Our Method: Overview



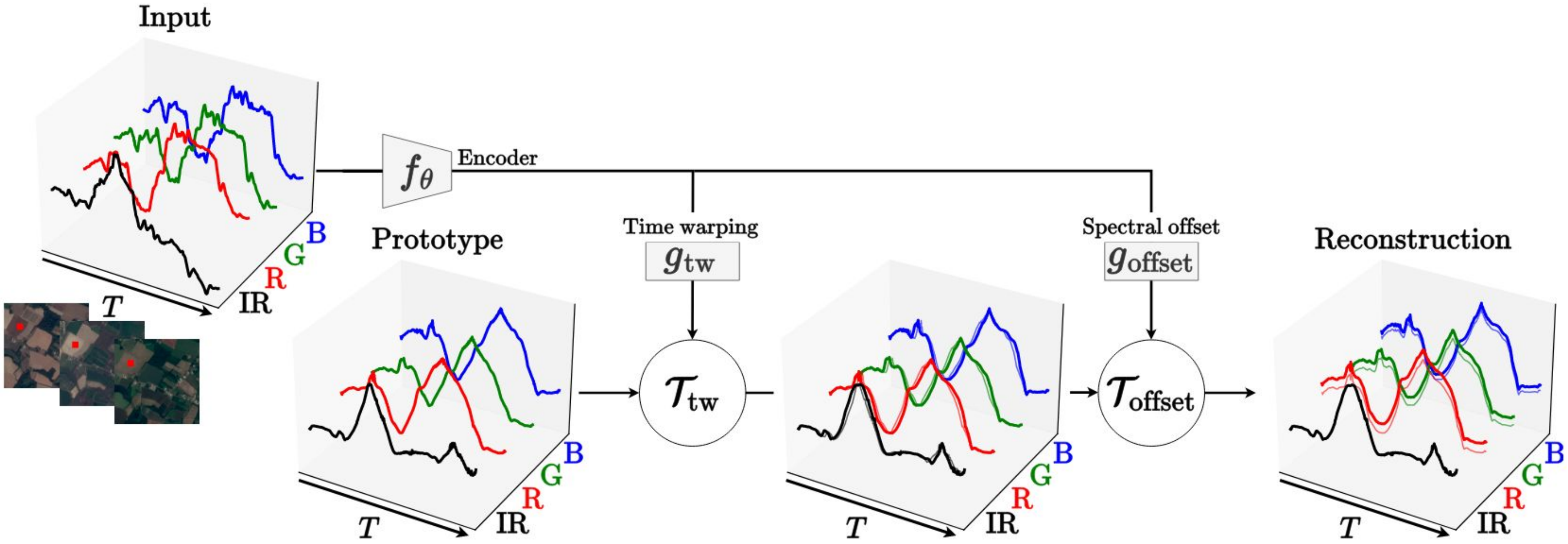
Our Method: Overview



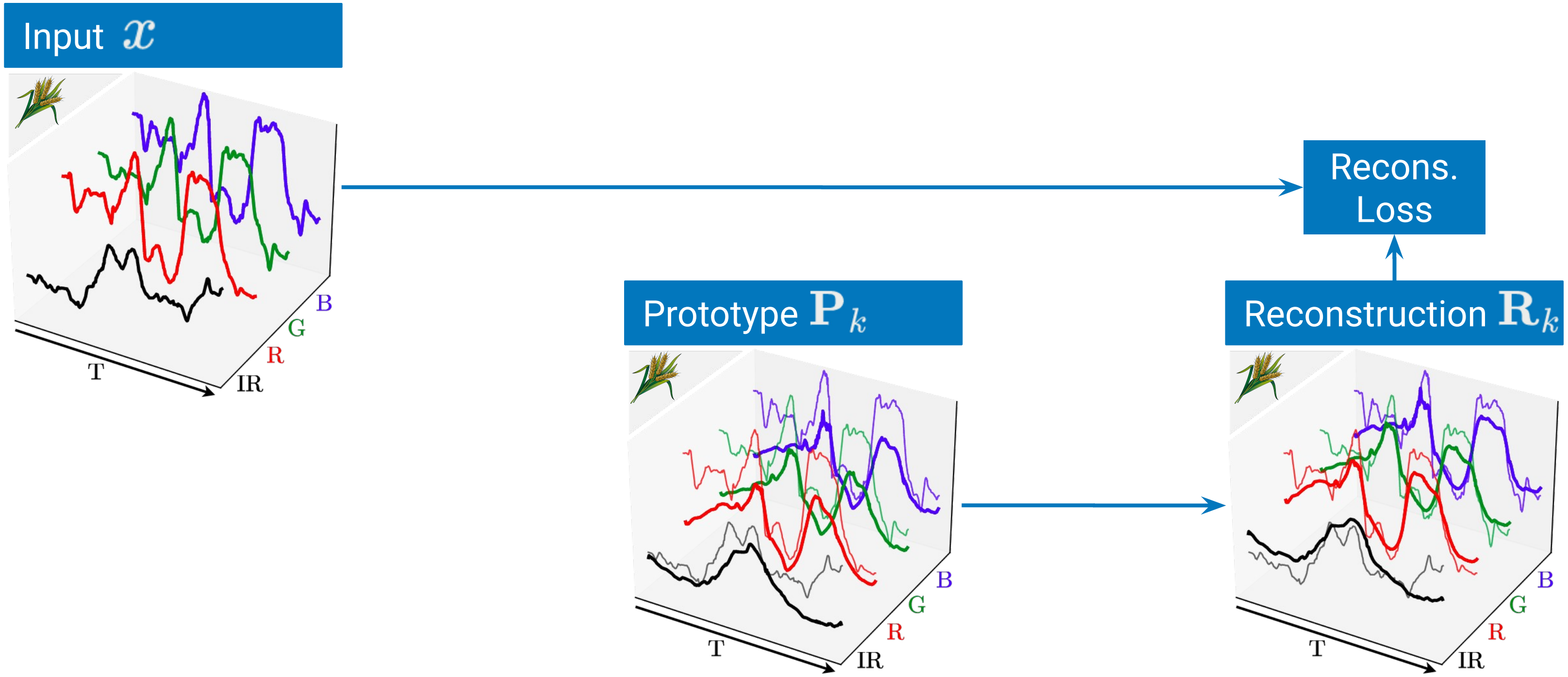
Our Method: Overview



Our Method: Overview



Our Method: Overview

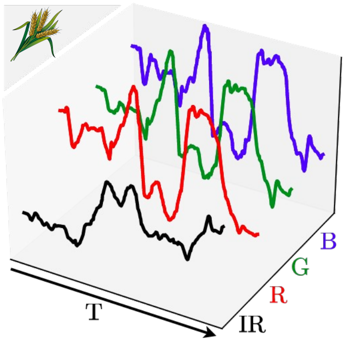


Our Method: Overview

Unsupervised reconstruction loss

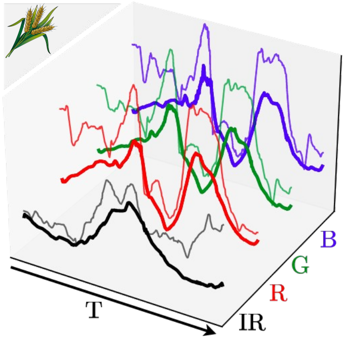
$$\mathcal{L}_{\text{rec}}(\theta, \mathbf{P}) = \sum_{x} \min_{k} \left|\left| x - \mathbf{R}_k(x, \theta, \mathbf{P}_k) \right|\right|^2$$

Input x

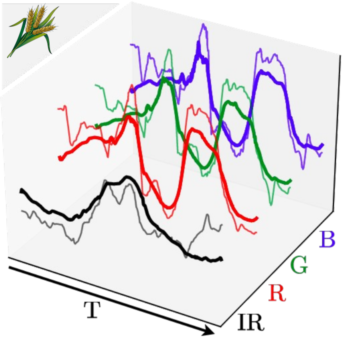


Recons. loss

Prototype $\mathbf{P}_k$



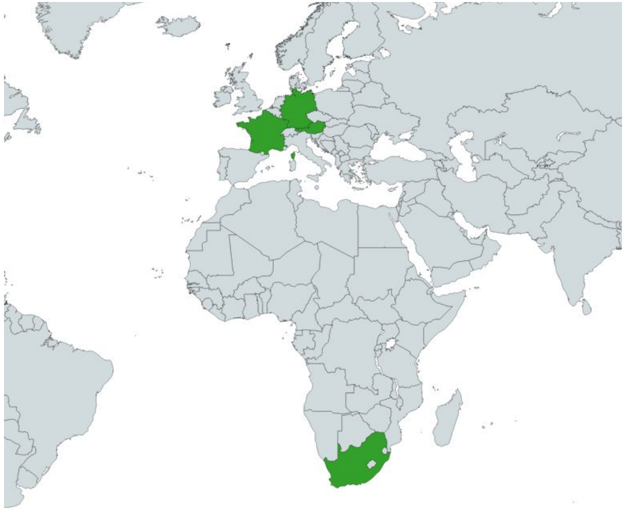
Reconstruction $\mathbf{R}_k$



Datasets

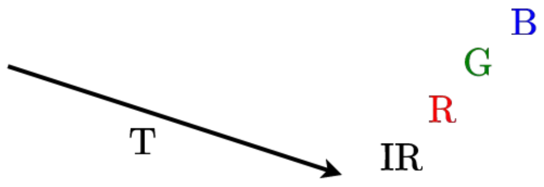
Dataset	Country	T	C	Satellite(s)	Daily	K	Train/Test shift
PASTIS [17]		406	10	Sentinel 2	✗	19	Spat.
TimeSen2Crop [58]		363	9	Sentinel 2	✗	16	Spat.
SA [27]		244	4	PlanetScope	✓	5	Spat.
DENETHOR [28]		365	4	PlanetScope	✓	9	Spat. & Temp.

- 4 different locations



Datasets

Dataset	Country	T	C	Satellite(s)	Daily	K	Train/Test shift
PASTIS [17]	🇫🇷	406	10	Sentinel 2	✗	19	Spat.
TimeSen2Crop [58]	🇦🇹	363	9	Sentinel 2	✗	16	Spat.
SA [27]	🇿🇦	244	4	PlanetScope	✓	5	Spat.
DENETHOR [28]	🇩🇪	365	4	PlanetScope	✓	9	Spat. & Temp.



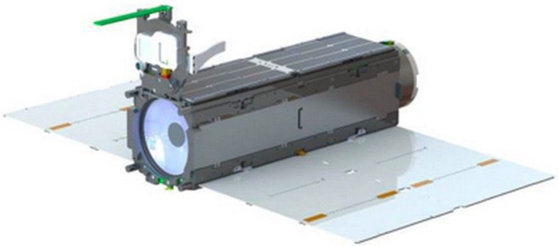
- 4 different locations
- Various sizes on both dimensions

Datasets

Dataset	Country	T	C	Satellite(s)	Daily	K	Train/Test shift
PASTIS [17]	🇫🇷	406	10	Sentinel 2	✗	19	Spat.
TimeSen2Crop [58]	🇦🇹	363	9	Sentinel 2	✗	16	Spat.
SA [27]	🇿🇦	244	4	PlanetScope	✓	5	Spat.
DENETHOR [28]	🇩🇪	365	4	PlanetScope	✓	9	Spat. & Temp.

- 4 different locations
- Various sizes on both dimensions
- Different acquisition parameters

Sentinel 2



PlanetScope

Datasets

Dataset	Country	T	C	Satellite(s)	Daily	K	Train/Test shift
PASTIS [17]	🇫🇷	406	10	Sentinel 2	✗	19	Spat.
TimeSen2Crop [58]	🇦🇹	363	9	Sentinel 2	✗	16	Spat.
SA [27]	🇿🇦	244	4	PlanetScope	✓	5	Spat.
DENETHOR [28]	🇩🇪	365	4	PlanetScope	✓	9	Spat. & Temp.

- 4 different locations
- Various sizes on both dimensions
- Different acquisition parameters
- +/- challenging

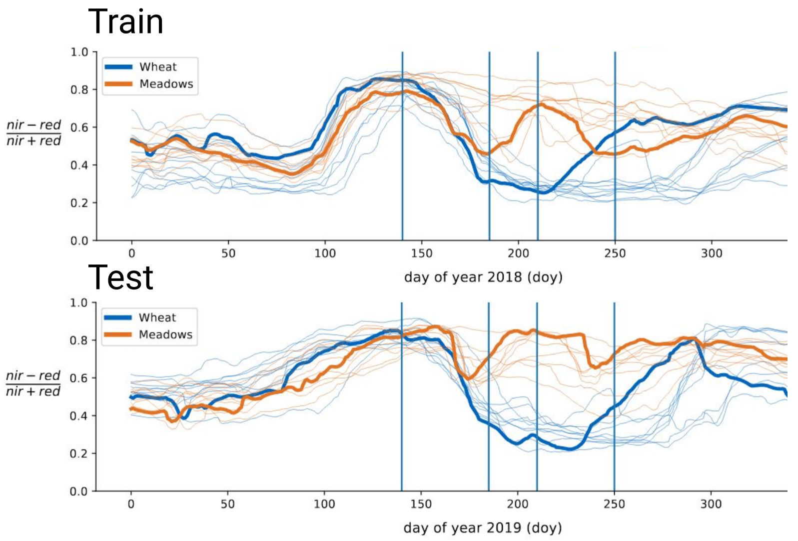
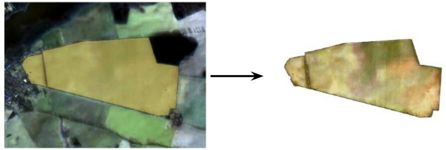


Image from DENETHOR paper

Datasets

Dataset	Country	T	C	Satellite(s)	Daily	K	Train/Test shift
PASTIS [17]		406	10	Sentinel 2	✗	19	Spat.
TimeSen2Crop [58]		363	9	Sentinel 2	✗	16	Spat.
SA [27]		244	4	PlanetScope	✓	5	Spat.
DENETHOR [28]		365	4	PlanetScope	✓	9	Spat. & Temp.

1 SITS = squared satellite acquisition



1 SITS = 1 crop type

Raw MTS (no images)

Training and inference details

1) Train on all Train/Val/Test sets

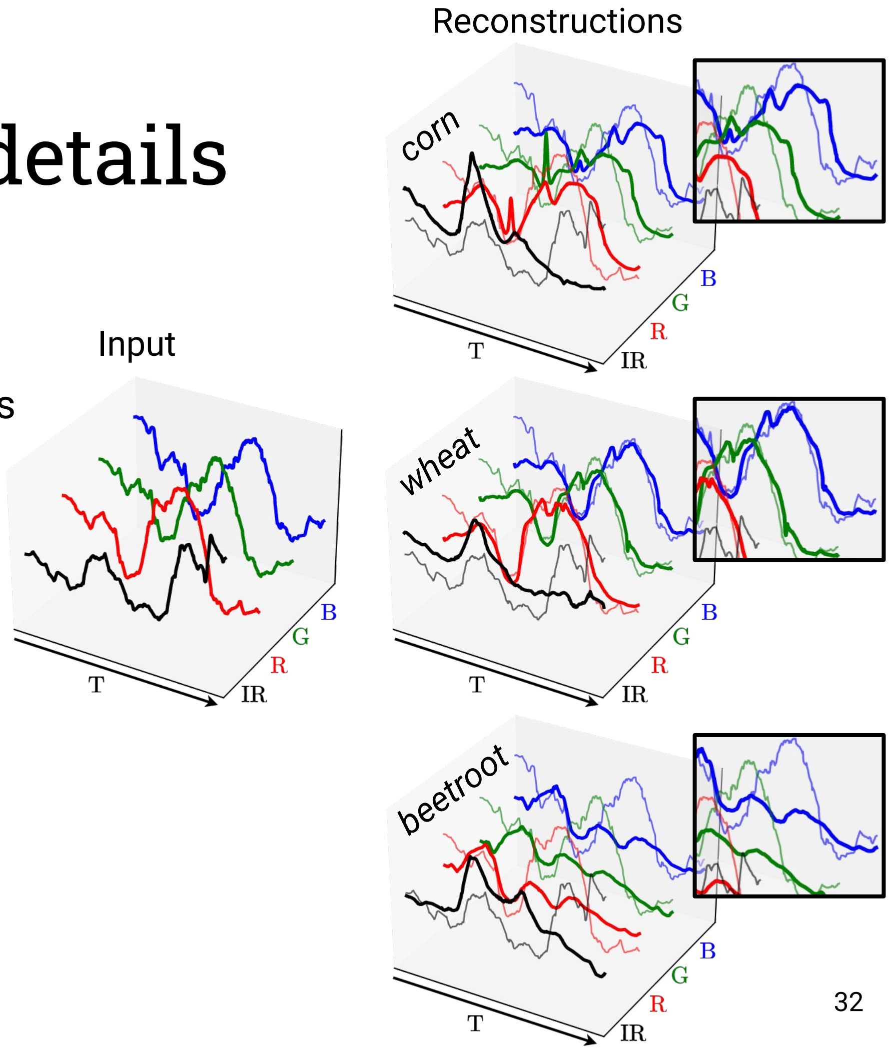
2) Each prototype assigned to the class it represents the most on a sample of Train set

Training and inference details

1) Train on all Train/Val/Test sets

2) Each prototype assigned to the class it represents the most on a sample of Train set

3) Each input TS assigned to the label of its closest transformed prototype

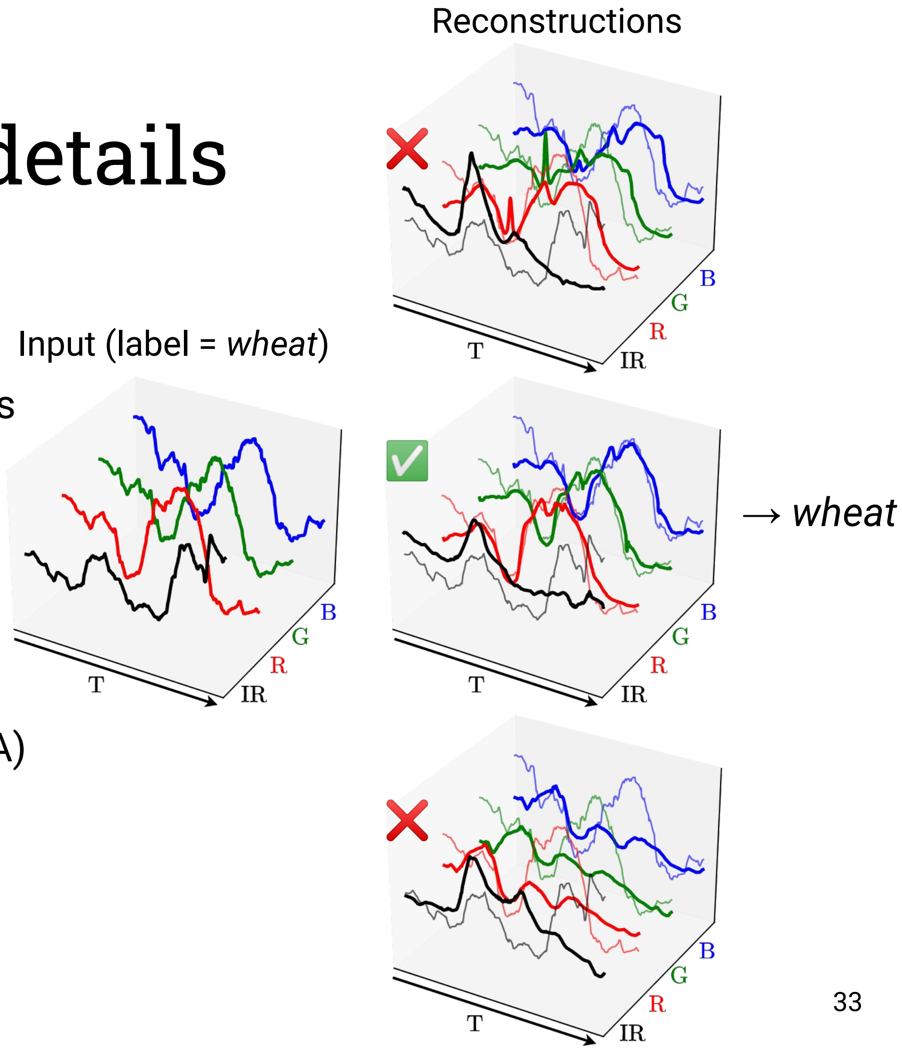


Training and inference details

1) Train on all Train/Val/Test sets

2) Each prototype assigned to the class it represents the most on a sample of Train set

3) Each input TS assigned to the label of its closest transformed prototype on Test set



Metrics: Overall Accuracy (OA), Mean Accuracy (MA)

Quantitative evaluation

Unsupervised

Method	#param (x1000)	PASTIS OA↑	PASTIS MA↑	TimeSen2Crop OA↑	TimeSen2Crop MA↑	SA OA↑	SA MA↑	DENETHOR OA↑	DENETHOR MA↑
K-means-DTW [39]	520	—	—	40.5	26.8	—	—	—	—
USRL [15]+K-means	290	63.9	20.4	34.9	23.6	60.9	48.6	54.0	46.4
DTAN [49]+K-means	646	65.6	21.4	47.7	29.3	60.5	48.6	46.3	36.9
K-means [6]	520	69.0	29.8	49.5	32.5	61.9	47.8	57.2	48.5
Ours + time warping	1 209	69.1	30.4	52.3	36.0	64.1	51.7	57.6	51.1
Ours + offset	1 373	67.7	28.6	52.0	35.5	63.6	50.4	58.5	52.6

Can also be trained under supervision!

Unsupervised

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

Supervised

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

Quantitative evaluation

Supervised

Method	#param (x1000)	PASTIS OA↑	PASTIS MA↑	TimeSen2Crop OA↑	TimeSen2Crop MA↑	SA OA↑	SA MA↑	DENETHOR OA↑	DENETHOR MA↑
UTAE [17]	1 087	83.4	77.7	—	—	—	—	—	—
MLP + LTAE [16]	320	80.6	65.9	88.7	80.9	67.4	63.7	55.6	43.6
OS-CNN [52]	4 729	81.3	68.1	87.9	81.2	64.6	60.3	49.0	39.2
TapNet [61]	1 882	77.4	69.5	83.9	83.0	69.4	62.5	61.5	60.6
MLSTM-FCN [26]	490	44.4	10.9	58.7	44.0	56.1	47.9	58.2	48.3
1NN [10]	0	65.8	40.1	43.9	35.0	60.7	54.9	56.7	48.2
1NN-DTW [48]	0	—	—	32.2	23.0	—	—	—	—
NCC [13]	77	56.5	48.4	57.4	49.5	51.3	46.4	61.3	55.5
Ours + time warping	427	56.2	51.4	59.9	52.3	54.5	49.7	**62.4**	56.4
Ours + offset	451	53.5	53.8	57.3	55.0	60.6	50.0	59.8	**62.9**

Quantitative evaluation

Supervised

Method	#param (x1000)	PASTIS OA↑	PASTIS MA↑	TimeSen2Crop OA↑	TimeSen2Crop MA↑	SA OA↑	SA MA↑	DENETHOR OA↑	DENETHOR MA↑
UTAE [17]	1 087	83.4	77.7	—	—	—	—	—	—
MLP + LTAE [16]	320	80.6	65.9	**88.7**	80.9	67.4	**63.7**	55.6	43.6
OS-CNN [52]	4 729	**81.3**	68.1	87.9	81.2	64.6	60.3	49.0	39.2
TapNet [61]	1 882	77.4	**69.5**	83.9	**83.0**	**69.4**	62.5	61.5	60.6
MLSTM-FCN [26]	490	44.4	10.9	58.7	44.0	56.1	47.9	58.2	48.3
1NN [10]	0	65.8	40.1	43.9	35.0	60.7	54.9	56.7	48.2
1NN-DTW [48]	0	—	—	32.2	23.0	—	—	—	—
NCC [13]	77	56.5	48.4	57.4	49.5	51.3	46.4	61.3	55.5
Ours + time warping	427	56.2	51.4	59.9	52.3	54.5	49.7	62.4	56.4
Ours + offset	451	53.5	53.8	57.3	55.0	60.6	50.0	59.8	62.9

Quantitative evaluation

Supervised

Method	#param (x1000)	PASTIS OA↑	PASTIS MA↑	TimeSen2Crop OA↑	TimeSen2Crop MA↑	SA OA↑	SA MA↑	DENETHOR OA↑	DENETHOR MA↑
UTAE [17]	1 087	83.4	77.7	—	—	—	—	—	—
MLP + LTAE [16]	320	80.6	65.9	**88.7**	80.9	67.4	**63.7**	55.6	43.6
OS-CNN [52]	4 729	**81.3**	68.1	87.9	81.2	64.6	60.3	49.0	39.2
TapNet [61]	1 882	77.4	**69.5**	83.9	**83.0**	**69.4**	62.5	61.5	60.6
MLSTM-FCN [26]	490	44.4	10.9	58.7	44.0	56.1	47.9	58.2	48.3
1NN [10]	0	65.8	40.1	43.9	35.0	60.7	54.9	56.7	48.2
1NN-DTW [48]	0	—	—	32.2	23.0	—	—	—	—
NCC [13]	77	56.5	48.4	57.4	49.5	51.3	46.4	61.3	55.5
Ours + time warping	427	56.2	51.4	59.9	52.3	54.5	49.7	**62.4**	56.4
Ours + offset	451	53.5	53.8	57.3	55.0	60.6	50.0	59.8	**62.9**

Conclusion

- Unsupervised: **our** prototype-based method outperforms
  - all current methods for unsupervised clustering of SITS on all 4 datasets

- Supervised: **our** prototype-based method
  - outperforms all supervised MTSC methods on the challenging DENETHOR
  - is less prone to overfitting

- Advantage of prototype-based methods:
  - Unsupervised and supervised training
  - Interpretability

Thank you!

https://arxiv.org/abs/2303.12533

https://github.com/ElliotVincent/AgriITSC

Elliot Vincent - 31/03/2023
– IGN Reading group –