Boosting Crop Classification by Hierarchically Fusing Satellite, Rotational, and Contextual Data

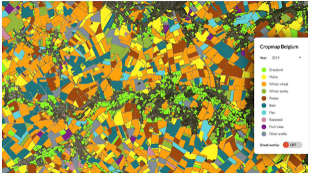
Valentin Barriere, Martin Claverie, Maja Schneider, Guido Lemoine, Raphaël d'Andrimont

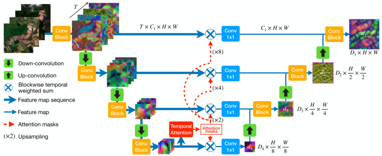
Universidad de Chile, European Commission's Joint Research Center, Technical University of Munich

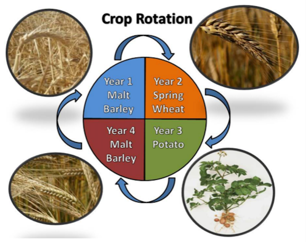
EO4AGRI 2024, Published in Remote Sensing of Environment

Intro I

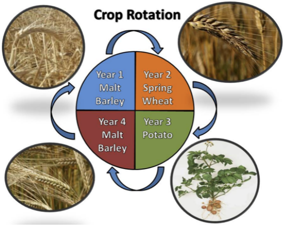
- Many countries monitor and forecast their crop production for Food Security reasons
- However the existing methods are modeling only using the remote sensing data
- We propose a model **fusing different modalities**







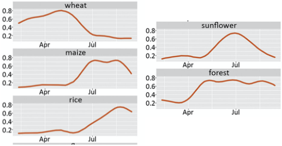
Crop rotation



- Crop rotation is a crop management technique which consists in adopting a series of different types of crops across the years for the same field
- Using crop type mapping of the past years, we obtain a sequence of crop type that we model in order to predict the current year

2014	2015	2016	2017	2018	2019	2020	2021
Barley	Wheat	potato	Barley	Barley	Wheat	potato	???

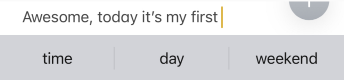
- In addition to the Crop rotation, we use the remote sensing information as an additional source of information



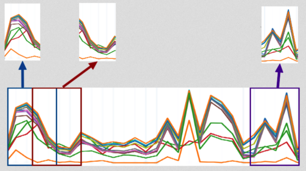
Multimodal Language Model

We modeled our problem as an Multimodal Natural Language task:

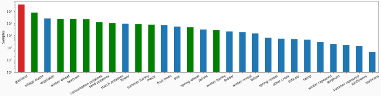
- **The crop types** were modeled as words like in a language model



- **The satellite signals** were modeled as an acoustic signal



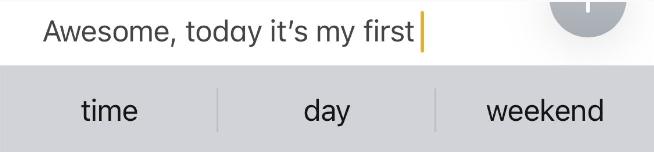
- **The crop distribution** acted as a speaker-specific vocabulary distribution



Multimodal Language Model I

We modeled our problem as a Multimodal NLP task:

The crop types were modeled as words like in a language model:

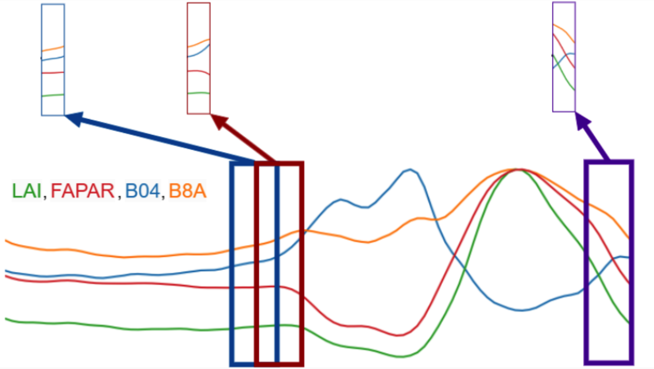


$$P(c_{t+1}|c_t, ..., c_1)$$

Multimodal Language Model II

We modeled our problem as an Multimodal NLP task:

The satellite signals were modeled as an acoustic signal, using a sliding window and statistical functionals to temporally aggregate the signal



Sliding window of size 30 days, with a step of 15 days

Multimodal Language Model III

We modeled our problem as a Multimodal NLP task:

The crop distribution acted as a speaker-specific vocabulary distribution:

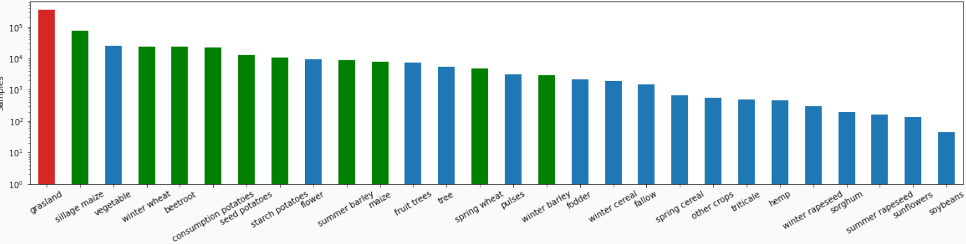
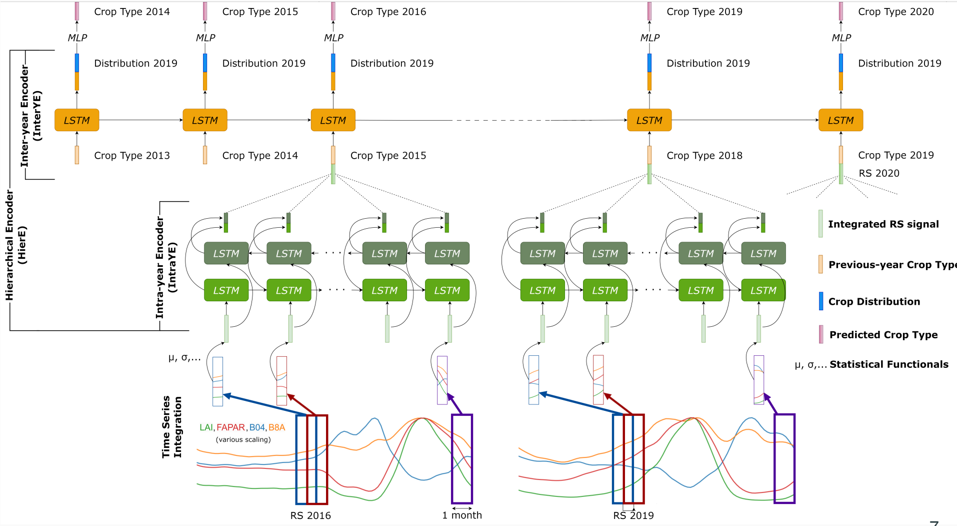
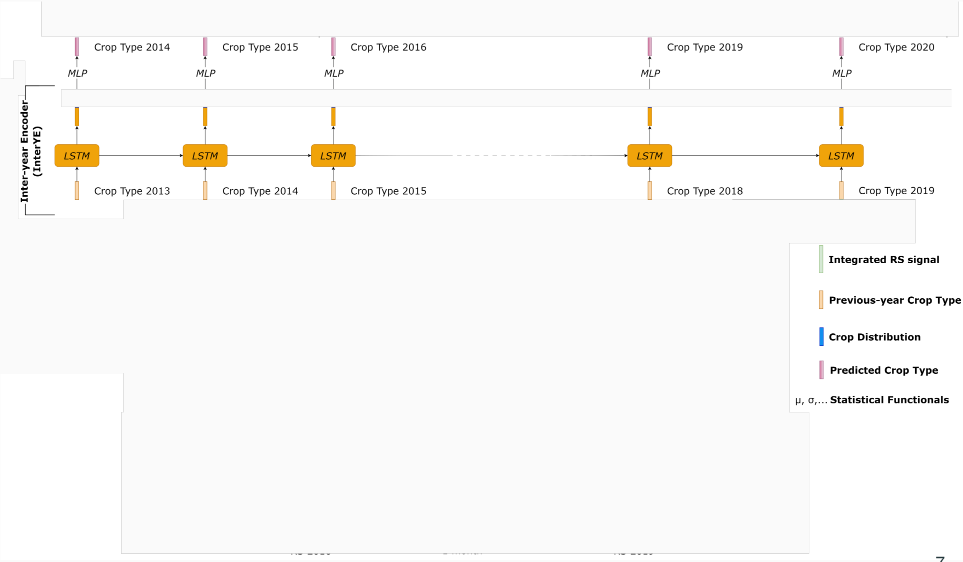


Figure 1: Distributions of the crop types in the dataset. Green crops are the remaining crop of interests used for the 10-class evaluation.

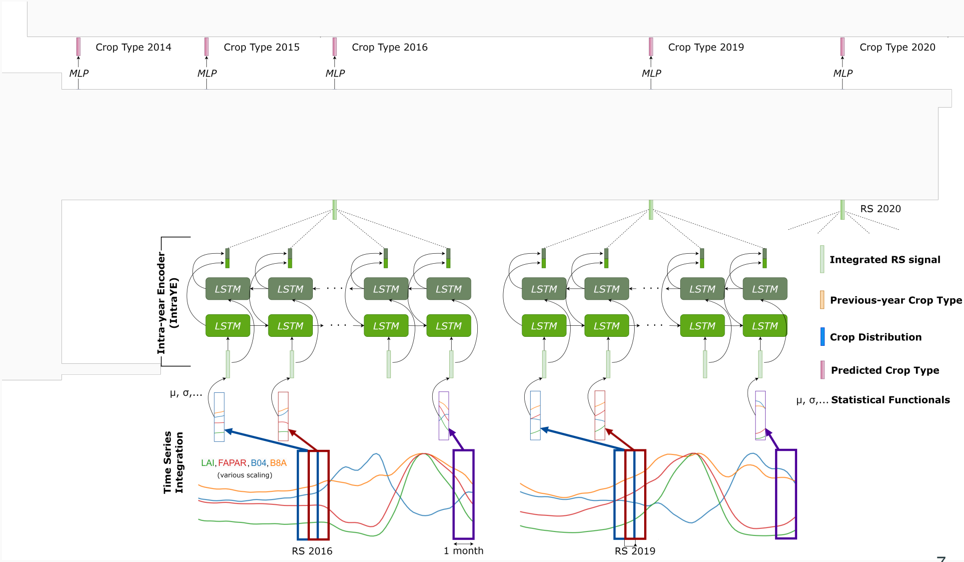
Hierarchical model I



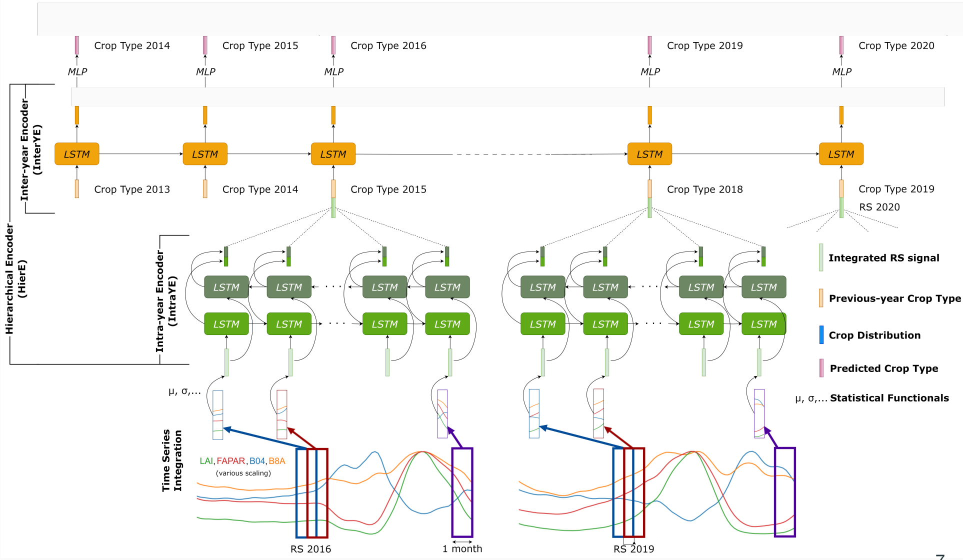
Hierarchical model I – IntraYear Encoder



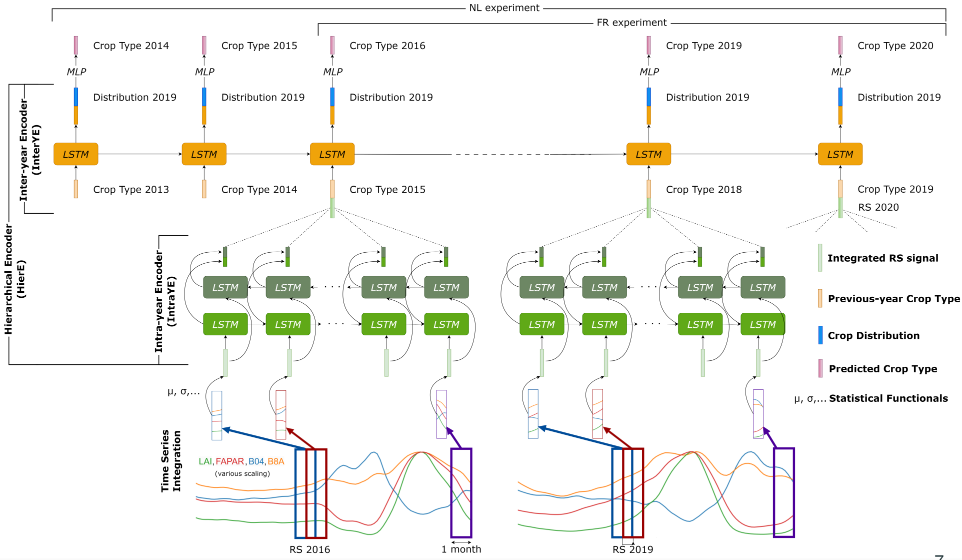
Hierarchical model I – InterYear Encoder



Hierarchical model I – Hierarchical Encoder



Hierarchical model I – Hierarchical Encoder



Hierarchical model II

Table 1: Summary of the different models used in this work, using Crop Rotations (CR), Remote Sensing (RS), and Crop Distribution (CD).

Models	CR	RS	CD	Modelisation-level Within season	Modelisation-level Between seasons	Hierarchical
IntraYE$_{RS}$	✗	✓	✗	✓	✗	✗
IntraYE$_{MM}$	✓	✓	✗	✓	✗	✗
InterYE$_{Crop}$	✓	✗	✗	✗	✓	✗
InterYE$_{RS}$	✗	✓	✗	✗	✓	✗
InterYE$_{MM}$	✓	✓	✗	✗	✓	✗
HierE$_{RS}$	✗	✓	✗	✓	✓	✓
HierE$_{MM}$	✓	✓	✗	✓	✓	✓
HierE$_{final}$	✓	✓	✓	✓	✓	✓

Experiments

Crop Types

141/151 Crop types (NL/FR) over 5 years, harmonized with EuroCrop taxonomy

Sentinel2 EO-based features

Time-series of B4 (red band) Surface Reflectance, b8a (near infrared band) Surface Reflectance, Leaf Area Index and Fraction of Absorbed Photosynthetically Active Radiation

Local features: Crop Distribution

141/151-dimension vector representing the distribution of the crops in the surrounding of the parcel (10km radius)

Validation: How to smartly aggregate classes?

Aggregated in 24/32 or 10/14 general classes regarding *(i)* the distribution of the crops in the country (**data-driven**) and *(ii)* EuroCrop taxonomy (**knowledge-driven**).

Number of parcels: 596k (NL) and 6.49M (FR)

Results NL – End of season classification

Labels / Model	# Modalities	141-class P	R	F1	Acc	10-class P	R	F1	Acc	8-class P	R	F1	m-F1
InterYE$_{Crop}$	1 (C)	36.0	25.5	27.4	76.2	51.8	43.0	43.5	77.7	43.3	35.5	34.9	53.6
IntraYE$_{RS}$ [2]	1 (RS)	27.4	20.9	20.4	89.8	78.8	75.9	74.5	92.9	76.1	72.6	70.8	87.8
InterYE$_{RS}$	1 (RS)	22.8	17.7	17.1	89.1	71.2	73.4	72.0	92.1	67.0	69.6	68.0	85.6
HierE$_{RS}$	1 (RS)	20.7	17.5	16.7	90.2	80.5	74.4	74.3	93.5	78.0	70.4	70.3	88.3
IntraYE$_{MM}$ [1]	2 (RS+BoC)	55.6	39.7	43.2	92.8	83.0	80.5	80.9	94.7	80.2	77.9	78.0	90.0
InterYE$_{MM}$	2 (RS+C)	41.1	33.0	33.6	92.2	82.2	79.7	80.4	94.5	80.2	76.3	77.5	89.5
HierE$_{MM}$	2 (RS+C)	47.3	38.7	39.7	93.3	85.2	81.9	83.1	95.2	83.6	78.8	80.6	91.1
HierE$_{final}$	3 (All)	47.1	39.3	40.2	**93.6**	86.7	81.9	83.6	**95.5**	85.3	78.7	81.1	**91.6**

Table 2: Results over Netherlands of the end-of-season classification models with different modalities: Remote Sensing (RS), Crop Rotations as embeddings (C) or BoC, and Spatial Crop Distribution.

Macro- Precision (P), Recall (R) and F1 score, and accuracy or micro-F1 score (m-F1).

Results FR – End of season classification

Labels / Model	# Modalities	151-class P	151-class R	151-class F1	151-class Acc	14-class P	14-class R	14-class F1	14-class Acc	12-class P	12-class R	12-class F1	12-class m-F1
InterYE$_{Crop}$	1 (C)	35.6	31.0	31.7	66.0	38.9	34.3	31.7	69.1	30.9	26.4	23.0	42.7
IntraYE$_{RS}$ [2]	1 (RS)	22.9	15.7	15.2	64.0	69.8	62.2	64.7	75.7	69.3	59.7	63.1	74.6
InterYE$_{RS}$	1 (RS)	21.3	13.2	12.6	54.9	63.9	59.6	60.2	72.2	62.7	57.4	58.5	71.2
HierE$_{RS}$	1 (RS)	25.3	19.0	18.8	66.3	72.5	65.5	67.7	76.9	71.9	63.2	66.1	76.5
IntraYE$_{MM}$ [1]	2 (RS+BoC)	52.7	32.4	35.9	82.7	78.1	68.7	71.0	86.6	76.2	65.6	68.0	80.3
InterYE$_{MM}$	2 (RS+C)	45.9	35.2	36.4	82.4	72.7	67.4	69.2	86.1	70.0	63.6	65.8	77.5
HierE$_{MM}$	2 (RS+C)	50.2	41.9	43.2	84.8	77.0	73.4	74.9	88.4	75.0	70.2	72.3	81.8
HierE$_{final}$	3 (All)	45.1	37.3	38.1	**85.4**	79.8	76.1	77.6	**89.1**	78.1	73.5	75.4	**83.6**

Table 3: Results over France of the end-of-season classification models with different modalities: Remote Sensing (RS), Crop Rotations as embeddings (C) or Bag-of-Crops (BoC), and Spatial Crop Distribution.

Macro- Precision (P), Recall (R) and F1 score, and accuracy or micro-F1 score (m-F1).

Few-shot/Domain Adaptation experiments

Labels Pre-train.	N	141-class P	R	F1	Acc	24-class P	R	F1	Acc	10-class P	R	F1	Acc	8-class P	R	F1	m-F1
	0	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅
	16	5.8	5.1	4.8	70.8	23.7	21.4	20.4	71.1	38.5	37.4	36.3	73.6	38.5	37.4	36.3	45.3
✗	64	2.7	2.5	2.2	69.2	17.1	13.1	12.5	69.4	27.3	25.7	23.3	69.6	27.3	25.7	23.3	34.7
	256	4.2	4.8	2.9	66.5	18.2	16.9	14.1	66.8	25.0	23.2	20.5	68.1	25.0	23.2	20.5	20.4
	1024	19.6	13.3	12.4	80.8	53.6	39.8	37.2	80.3	69.7	60.4	61.5	84.0	69.7	60.4	61.5	76.3
	0	5.7	4.8	4.2	47.3	14.7	15.1	11.1	46.6	20.6	19.7	16.6	46.9	12.3	7.4	8.4	24.5
	16	12.2	7.8	7.6	70.3	30.5	23.8	24.5	70.4	37.9	33.9	34.0	72.3	37.9	33.9	34.0	45.2
✓	64	16.7	13.6	13.5	74.7	41.9	38.7	38.1	75.0	51.6	45.4	46.6	76.4	51.6	45.4	46.6	54.4
	256	25.8	21.4	20.8	82.5	55.6	51.1	50.6	82.7	67.3	58.0	60.1	84.6	67.3	58.0	60.1	69.2
	1024	32.7	27.3	26.0	84.9	61.3	57.3	54.3	84.9	73.8	72.0	71.6	87.0	73.8	72.0	71.6	80.9
✗	All	47.1	39.2	**40.2**	**93.7**	76.6	75.8	**75.8**	94.0	86.7	81.9	83.6	95.5	85.3	78.7	81.1	91.6
✓	All	42.5	35.3	36.0	92.8	67.3	53.4	55.9	**94.2**	89.9	82.2	**85.3**	**95.7**	88.8	77.6	**82.3**	**91.8**

Table 4: Results over Netherlands of the few-shot final classification models, with or without pre-training over France.

- Pre-training helps when a few samples are available
- The PT model gets better results for the aggregated distribution: so it overfits less the target domain data distribution

Splitting experiments

Labels / Dataset	Split	141/151-class				24/32-class				10/14-class				8/12-class			
		P	R	F1	Acc	P	R	F1	Acc	P	R	F1	Acc	P	R	F1	m-F1
NL	T	47.2	41.9	42.7	93.7	77.2	75.9	76.0	94.1	87.0	82.1	83.8	95.6	85.7	78.9	81.4	91.7
	T+S	47.3	42.5	42.7	93.3	75.8	74.8	74.6	93.7	86.6	81.2	82.9	95.3	85.3	77.8	80.4	91.0
FR	T	44.4	38.7	39.4	85.4	72.0	68.6	69.1	85.6	79.5	75.9	77.4	89.1	77.8	73.2	75.1	83.5
	T+S	44.8	38.4	38.4	85.1	71.6	68.0	68.8	85.4	79.1	74.4	76.3	88.8	77.3	71.5	73.9	82.7

Table 5: Results over the Netherlands and France of the best end-of-season classification architecture using different splits, for the test season 2020.

- Trained over the year 201x-2019
- Tested on 10% of the parcels (year 2020)
- One split purely temporal (T): trained over all the dataset
- One split temporal and spatial (T+S): trained over 90% of the dataset

Feature experiments

Labels / Model	Features	141-class P	141-class R	141-class F1	141-class Acc	10-class P	10-class R	10-class F1	10-class Acc	8-class P	8-class R	8-class F1	8-class Acc
IntraYE$_{RS}$	Set1	27.4	20.9	20.4	89.8	78.8	75.9	74.5	92.9	76.1	72.6	70.8	87.8
IntraYE$_{MM}$		55.6	39.7	43.2	92.8	83.0	80.5	80.9	94.7	80.2	77.9	78.0	90.0
InterYE$_{MM}$		41.1	33.0	33.6	92.2	82.2	79.7	80.4	94.5	80.2	76.3	77.5	89.5
HierE$_{MM}$		47.3	38.7	39.7	93.3	85.2	81.9	83.1	95.2	83.6	78.8	80.6	91.1
HierE$_{final}$		47.1	39.3	40.2	**93.6**	86.7	81.9	83.6	**95.5**	85.3	78.7	81.1	**91.6**
IntraYE$_{RS}$	Set2	36.0	27.4	27.4	92.5	86.5	82.3	82.2	95.2	84.7	79.9	79.6	91.7
IntraYE$_{MM}$		61.0	45.6	49.0	94.3	88.0	85.9	86.3	96.1	85.9	84.2	84.2	92.7
InterYE$_{MM}$		47.1	35.8	37.9	94.0	88.6	84.5	86.0	96.1	87.3	82.2	84.0	92.8
HierE$_{MM}$		48.6	40.1	41.0	95.0	90.8	87.7	88.9	96.8	89.7	85.8	87.3	94.2
HierE$_{final}$		52.5	46.3	46.7	**95.3**	90.6	87.9	89.0	**96.9**	8.96	85.7	87.3	**94.2**
IntraYE$_{RS}$	Set3	26.7	17.4	18.4	88.6	81.5	72.1	73.2	91.9	80.0	68.6	69.8	86.3
IntraYE$_{MM}$		55.2	41.8	44.6	92.8	85.2	81.9	82.8	94.9	83.0	79.5	80.4	90.4
InterYE$_{MM}$		40.5	30.6	31.8	90.8	80.6	78.7	79.1	93.2	78.1	75.4	76.0	86.2
HierE$_{MM}$		48.5	39.3	40.5	**93.5**	88.0	83.7	85.4	**95.5**	86.9	81.0	83.4	91.8
HierE$_{final}$		49.6	41.1	42.1	**93.5**	88.8	83.7	85.8	**95.5**	88.3	80.7	84.0	**91.9**

Set1: B4, B8A, FAPAR and LAI.

Set2: Set1+B2, B3, B8, B11 and B12.

Set3: 10m bands of Sentinel2, B2/3/4/8

Contributions

- This architecture can be used with any backbone
- It outperforms from far classical SOTA networks
- A new method to aggregate classes for validation
- New abilities for early-season and few-shot learning!

Contributions

- This architecture can be used with any backbone
- It outperforms from far classical SOTA networks
- A new method to aggregate classes for validation
- New abilities for early-season and few-shot learning!
- And more in the paper...

ELSEVIER

Remote Sensing of Environment

Volume 305, 1 May 2024, 114110

Boosting crop classification by hierarchically fusing satellite, rotational, and contextual data

Valentin Barriere [a b c 1], Martin Claverie [c 1], Maja Schneider [d], Guido Lemoine [c], Raphaël d'Andrimont [c]

Thanks for listening!

References i

F. Quinton and L. Landrieu.
Crop rotation modeling for deep learning-based parcel classification from satellite time series.
Remote Sensing, 13(22), 2021.

M. Rußwurm, S. Lefèvre, and M. Körner.
Breizhcrops: a satellite time series dataset for crop type identification.
Time Series Workshop of the 36th ICML, 2019.