TERRITOIRE ENVIRONNEMENT TELEDETECTION INFORMATION SPATIALE

Combining AI and Earth Observation data to deal with land cover mapping

Dino Ienco, PhD, HDR
INRAE Researcher @UMR TETIS
dino.ienco@inrae.fr

Joint collaboration with:
- Alessandro Michele Censi (INRAE)
- Yawogan Jean-Eudes Gbodjo (INRAE)
- Raffaele Gaetano (CIRAD)
- Roberto Interdonato (CIRAD)
- Ruggero G. Pensa (UniTo)

#DigitAg

My Institue & My Lab

INRAE is a public research institute working for the coherent and sustainable development of agriculture, food and the environment.

Joint Research Unit with **INRAE**, **CIRAD**, **AgroParisTech** and **CNRS**

My Institue & My Lab

INRAE is a public research institute working for the coherent and sustainable development of agriculture, food and the environment.

TERRITOIRE ENVIRONNEMENT TELEDETECTION INFORMATION SPATIALE

tetis

Joint Research Unit with **INRAE**, **CIRAD**, **AgroParisTech** and **CNRS**




Outline

Introduction / Context / Background

SITS and Object-based Image Analysis

Reunion Island Case Study

TASSEL: Manage intra-object heterogeneity for SITS analysis

STARCANE: How much spatial context matters for SITS analysis

Conclusions

Introduction

Nowadays, many earth observation satellite missions exist:
- Sentinel [Senti]
- LandSat-8 [LandSat]
- SPOT 6/7[Spot]
- …

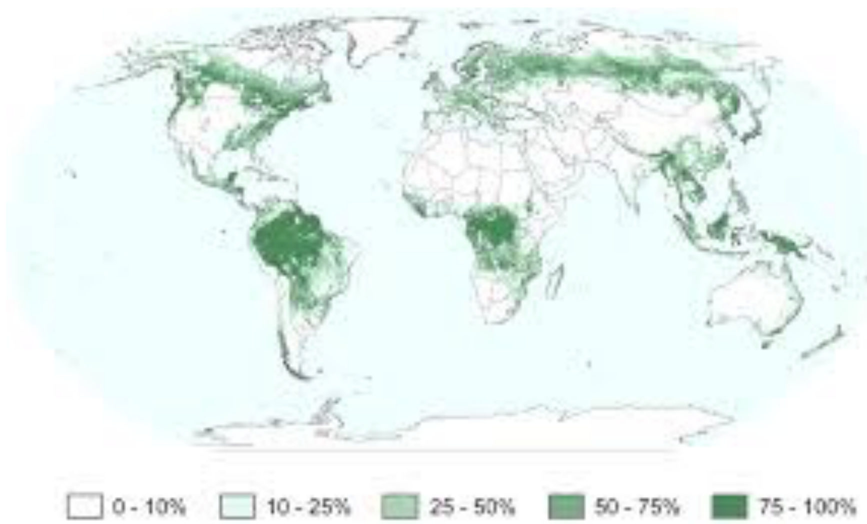
Acquired images have different:
- spatial resolution (0.5 – 30 meters)
- radiometric content (spectral bands)
- temporal resolution (every 5 – 365 days)

HUGE quantity of Satellite Images Describing Earth Phenomena at different scales

Why EOD is an Opportunity

Earth Observation Data can have practical influence on different domains:

Continental Surface analysis



Sustainable Agriculture

Climate Changes Analysis

Biodiversity Monitoring

EOD to support SDG

In 2015, a collection of 17 interlinked global goals for a "a better and more sustainable future for all" were defined by United Nations with the objective to be achieved by 2030

EOD to support SDG

In 2015, a collection of 17 interlinked global goals for a "a better and more sustainable future for all" were defined by United Nations with the objective to be achieved by 2030

EOD has been also highlighted as an actionable tool to support actions towards many Sustainable Development Goals (SDG), like :

- 1 - No Poverty
- 2 - Zero Hunger
- 11 - Sustainable City
- 13 - Climate Action
- 15 - Life on land
- ….

EOD to support SDG

In 2015, a collection of 17 interlinked global goals for a "a better and more sustainable future for all" were defined by United Nations with the objective to be achieved by 2030

EOD has been also highlighted as an actionable tool to support actions towards many Sustainable Development Goals (SDG), like :

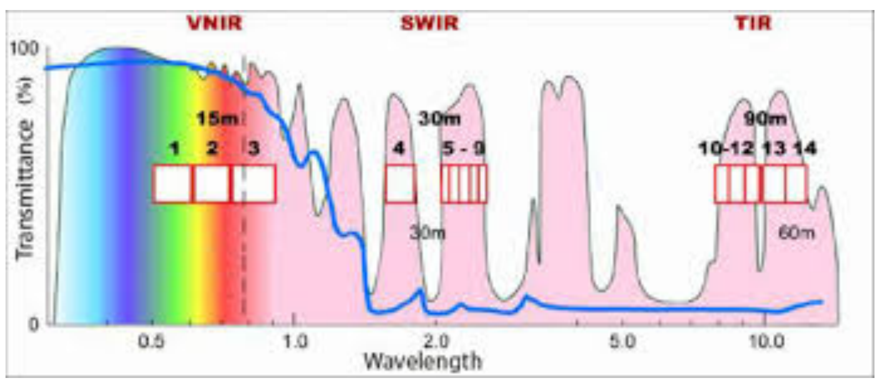
- 1 - No Poverty
- 2 - Zero Hunger
- 11 - Sustainable City
- 13 - Climate Action
- 15 - Life on land
- ….

For instance, improve agricultural monitoring systems (via EOD) is one of the way to promote sustainable agriculture thus, supporting the achievement of Zero Hunger SDG.

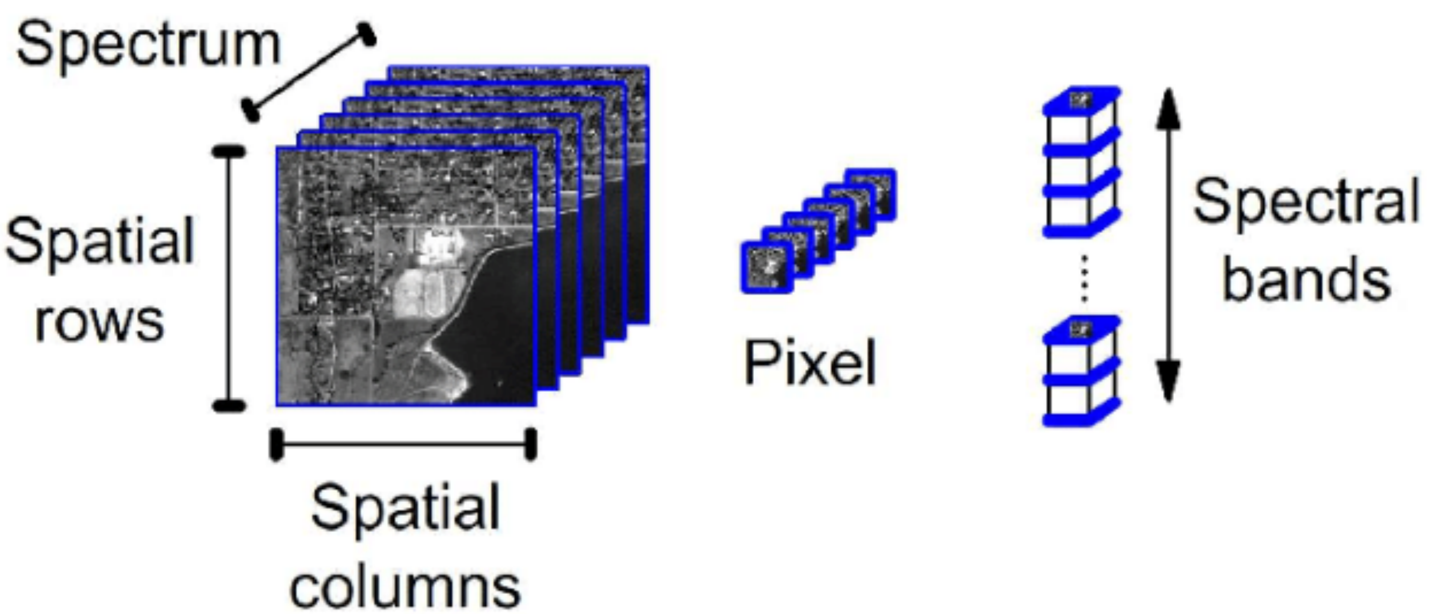
Context

A Satellite Image:

A data cube that describes a spatial area by means of several spectral bands



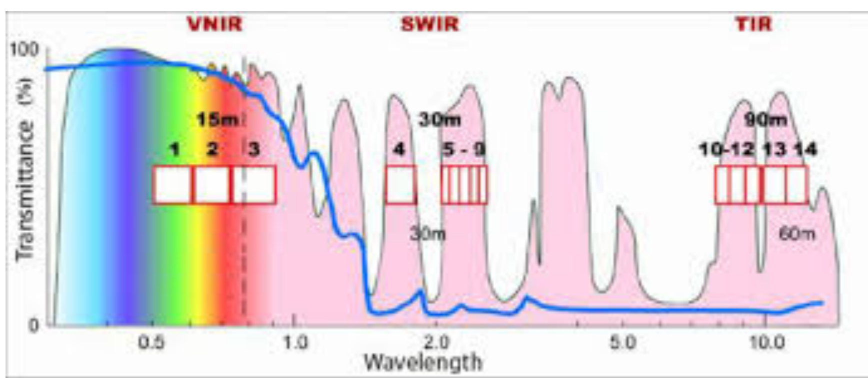
Spectral Bands



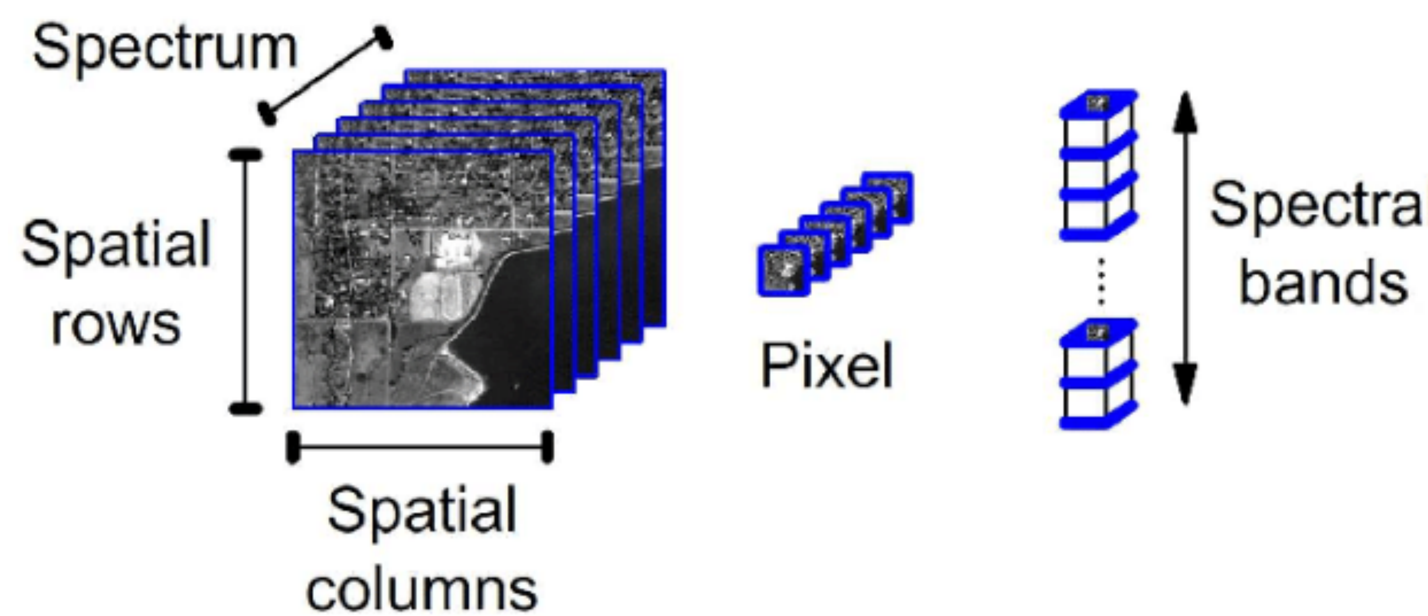
Context

A Satellite Image:

A data cube that describes a spatial area by means of several spectral bands



Spectral Bands



Type of information:

- Optical Images (Multi-Spectral / Hyperspectral)
- Radar Images (phase, amplitude, etc…)
- LIDAR (point clouds)
- Etc…

Context

EOD allows to collect Very High Resolution Images (VHR) i.e. Spot6/7 (at 1.5m), Pléiades (.5m), WorldView3 (.3m) at Low Temporal Frequency (once or twice per year).

VHSR data are useful to obtain fine resolution information to characterise spatial pattern and spatial texture

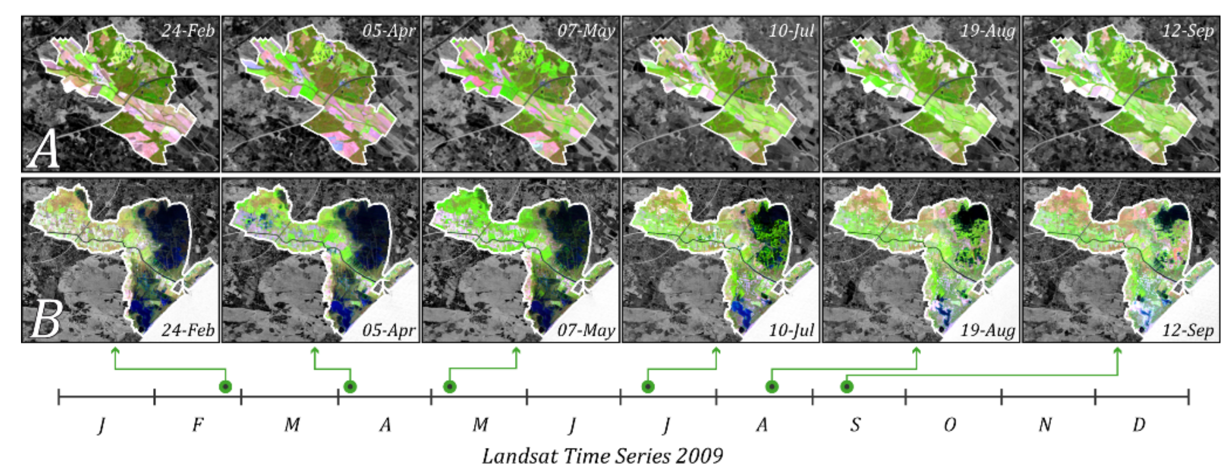
Context

EOD allows to collect Very High Resolution Images (VHR) i.e. Spot6/7 (at 1.5m), Pléiades (.5m), WorldView3 (.3m) at Low Temporal Frequency (once or twice per year).

VHSR data are useful to obtain fine resolution information to characterise spatial pattern and spatial texture

EOD allows to collect Satellite Image Time Series (SITS) at High Spatial Resolution (Sentinel ~10m) and High Temporal Frequency (every 5/10 days)

The same geographical area is observed



SITS data are useful to analyze spatio-temporal phenomena (trends and changes) over the time

Satellite Image Time Series

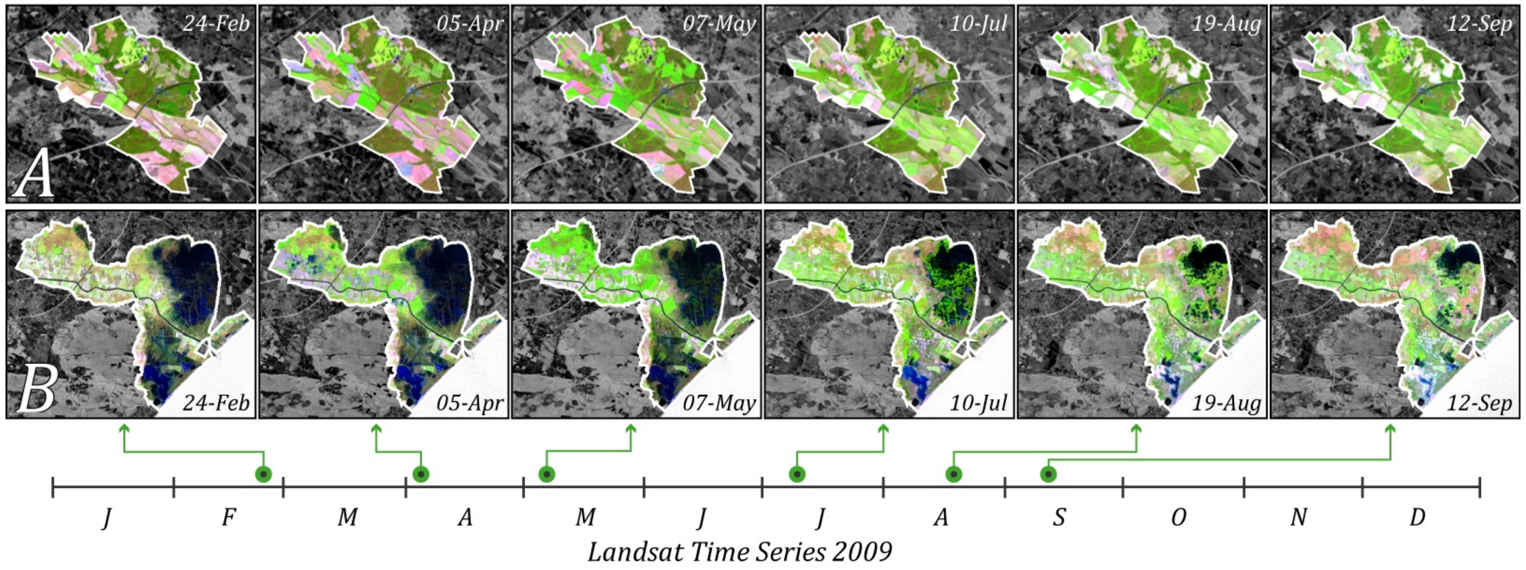
Among all the opportunities, the possibility to collect **multiple satellite images** (SITS: Satellite Image Time Series), **on the same area**, with **high revisit period** and **high spatial resolution** is paving the way to new applications (especially in agricultural land monitoring)

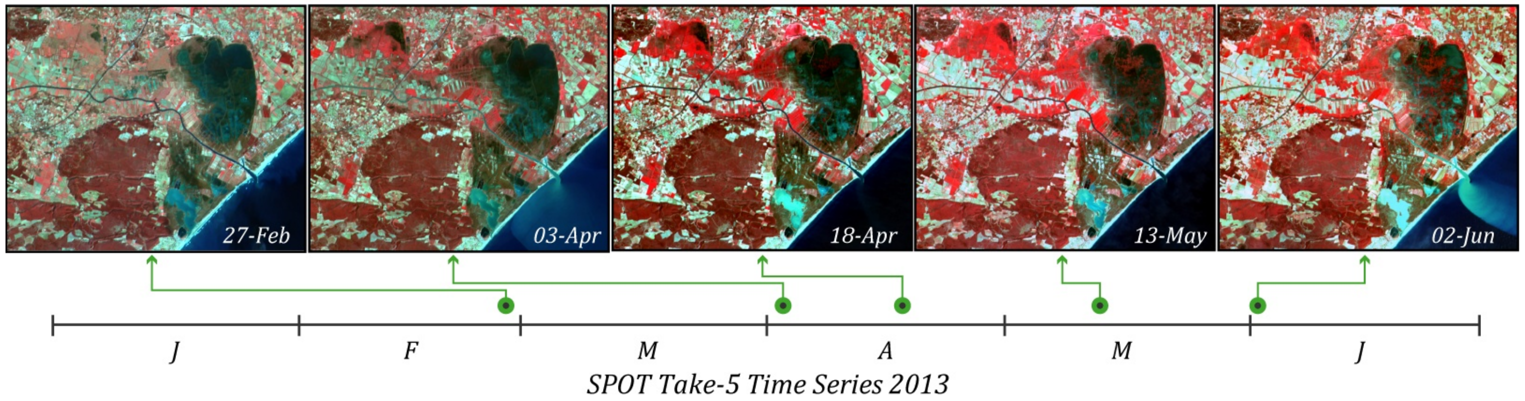
Satellite Image Time Series

Among all the opportunities, the possibility to collect **multiple satellite images** (SITS: Satellite Image Time Series), **on the same area**, with **high revisit period** and **high spatial resolution** is paving the way to new applications (especially in agricultural land monitoring)

<u>In the context of agriculture:</u>

- SITS allows to **distinguish between** different crops
- SITS captures **phenological cycle**
- SITS supports **change detection analysis**
- SITS helps to monitor **spatio-temporal phenomena**



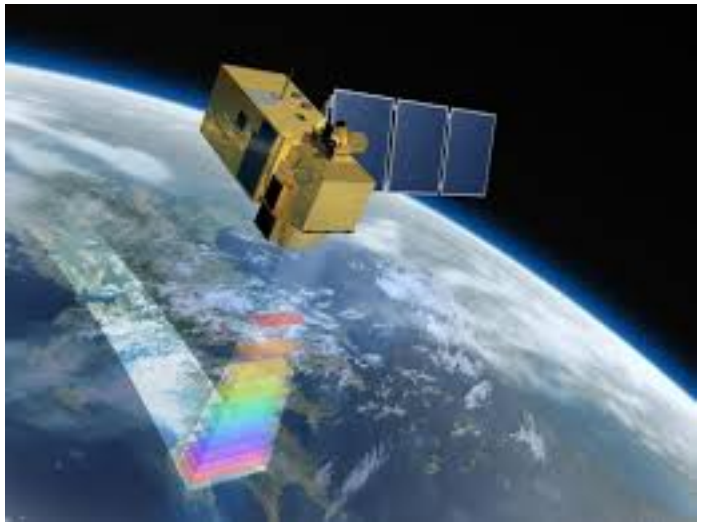


Satellite Image Time Series

Sentinel Missions belong to the **Copernicus Programme**

Copernicus Programme is provided by the **ESA** (European Space Agency)

Provide Remote Sensing data at **High Spatial/Temporal Resolution** of the **Earth Surface**



Satellite Image Time Series

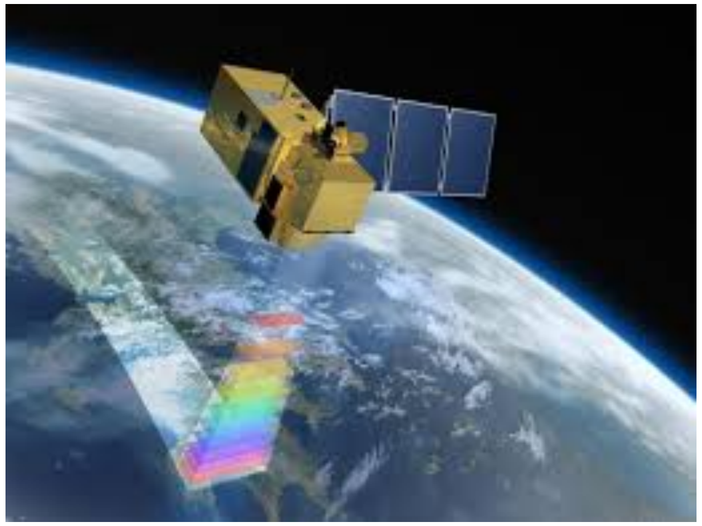
Sentinel Missions belong to the **Copernicus Programme**

Copernicus Programme is provided by the **ESA** (European Space Agency)

Provide Remote Sensing data at **High Spatial/Temporal Resolution** of the **Earth Surface**

Sentinel 1 :

- Two satellites supplying C-band synthetic aperture radar imaging,
- Revisit time period between 5 and 10 days with a spatial resolution of 10m.
- Especially useful to monitor soil and structural properties (i.e. rugosity and humidity).

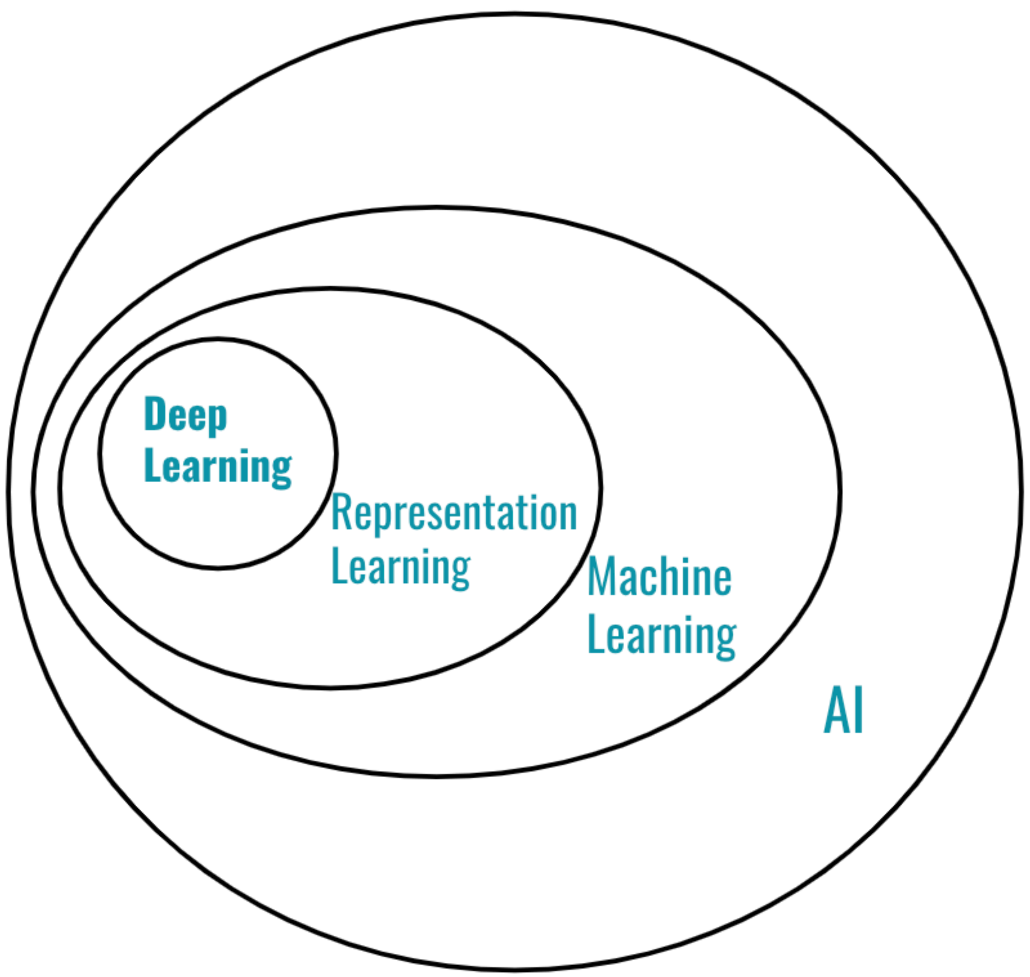


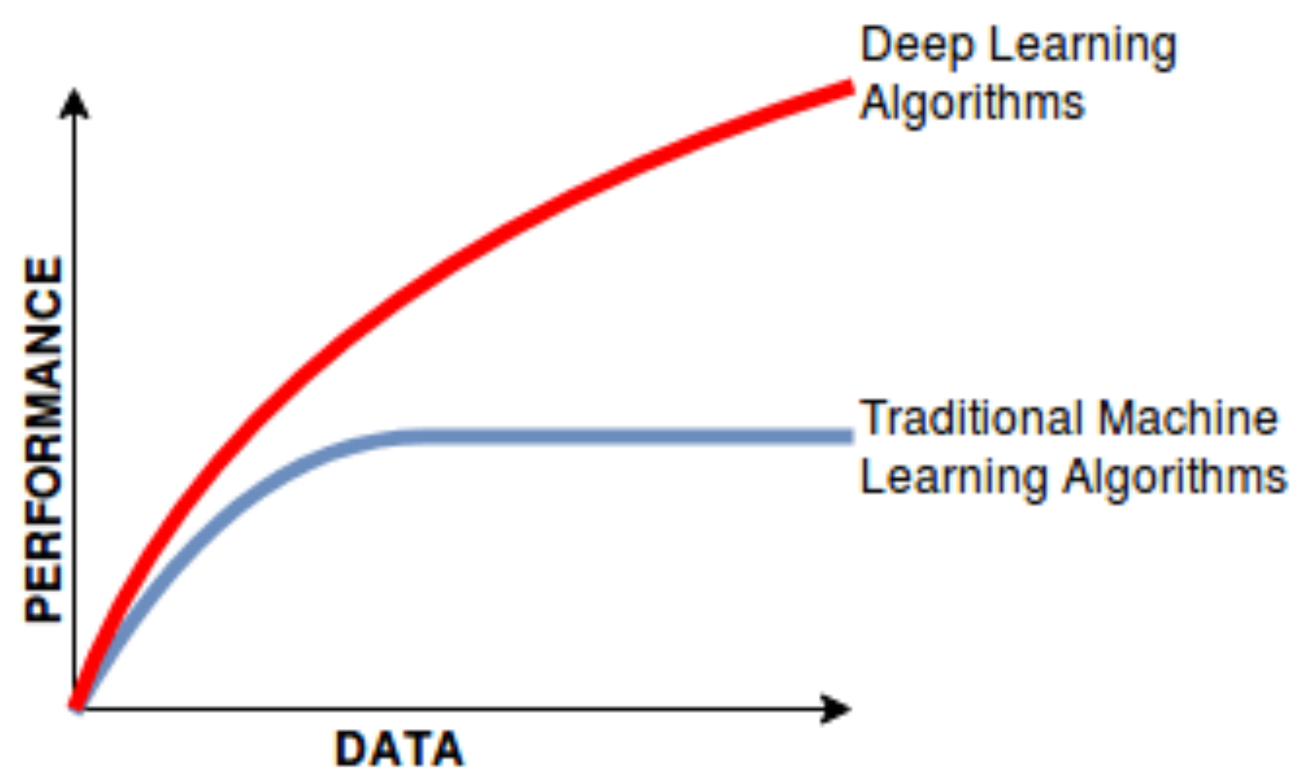
Sentinel 2:

- Two satellites supplying optical information.
- Revisit time period between 5 and 10 days till with a spatial resolution between 10 and 20m.
- Especially useful to monitor surface reflectance (i.e. land cover).

Machine Learning

- Increasing application of Machine Learning approaches on signal data
- Deep Learning, Neural Networks
- Deep Learning is a subfield of Machine Learning





Source: Blog Datacamp

Deep Learning
Learning representation

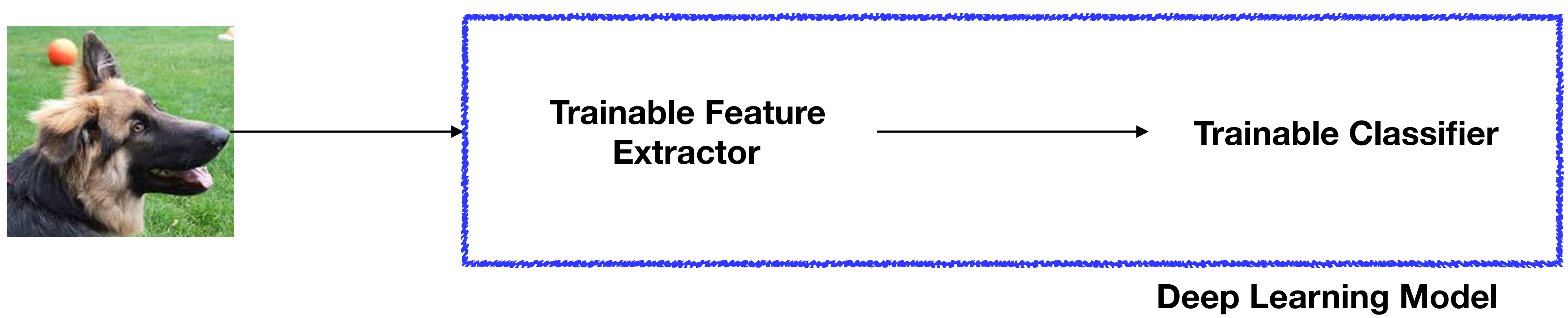
Traditional Machine Learning systems leverage **feature engineering** to represent the data:
- Text Analysis: Bag of Words
- Image Analysis: Hog (Histogram of Oriented gradient), SIFT (Scale Invariant Feature Transform)

Hand-Crafted Features → **Simple Trainable Classifier (SVM, RF, NB, …)**

Deep Learning
Learning representation

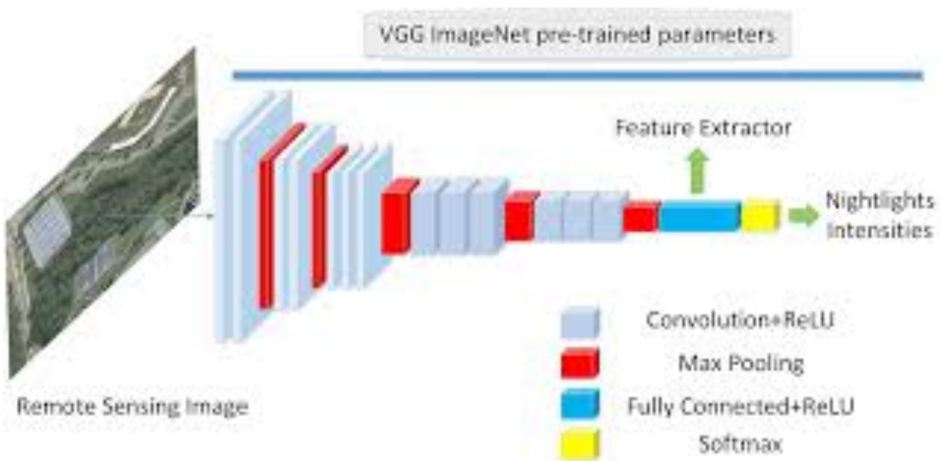
Traditional Machine Learning systems leverage **feature engineering** to represent the data:
- Text Analysis: Bag of Words
- Image Analysis: Hog (Histogram of Oriented gradient), SIFT (Scale Invariant Feature Transform)

Deep Learning approaches **learn internal representations (new features)** without necessity of hand-crafted features

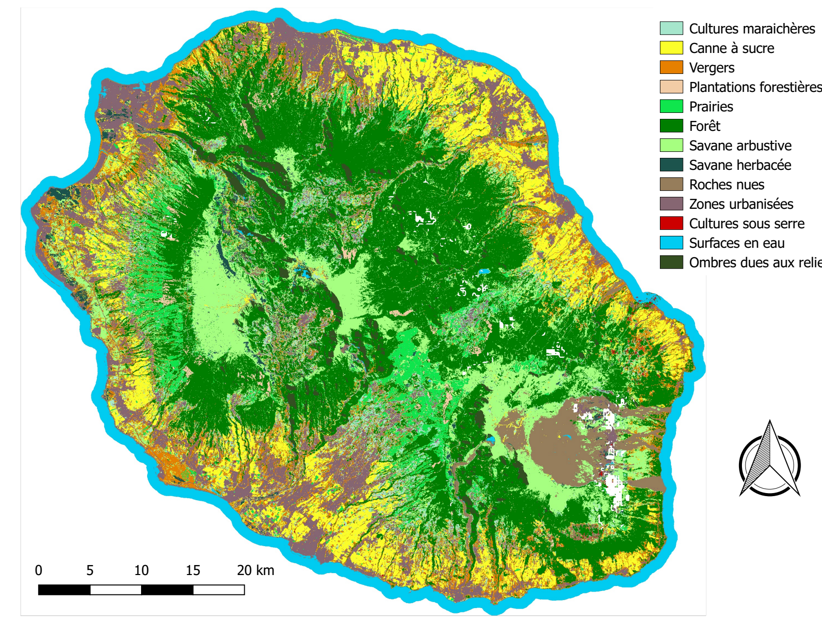


Deep Learning & EO data applications

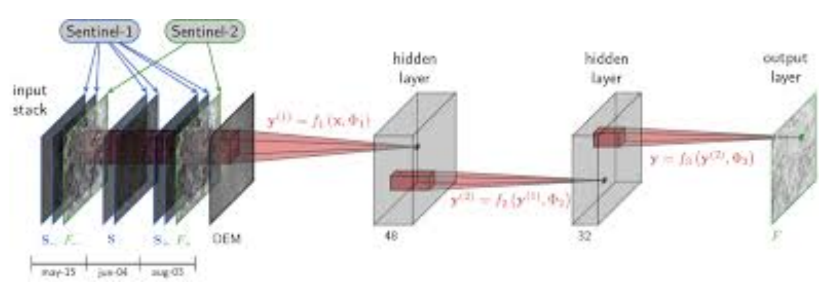
Scene Classification



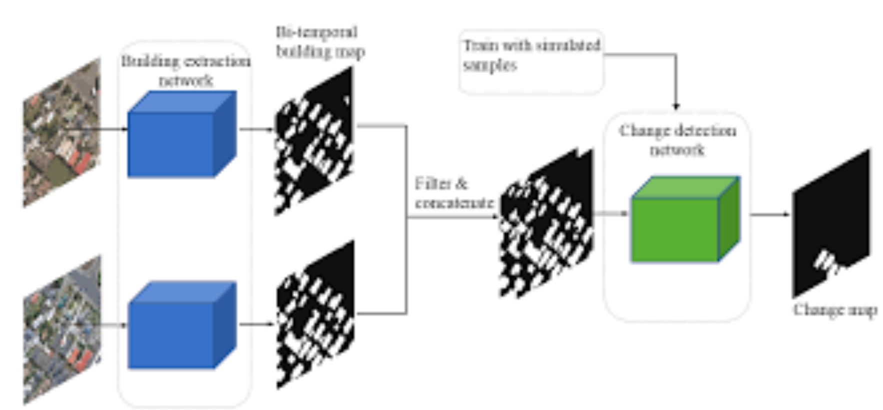
Land Cover Mapping



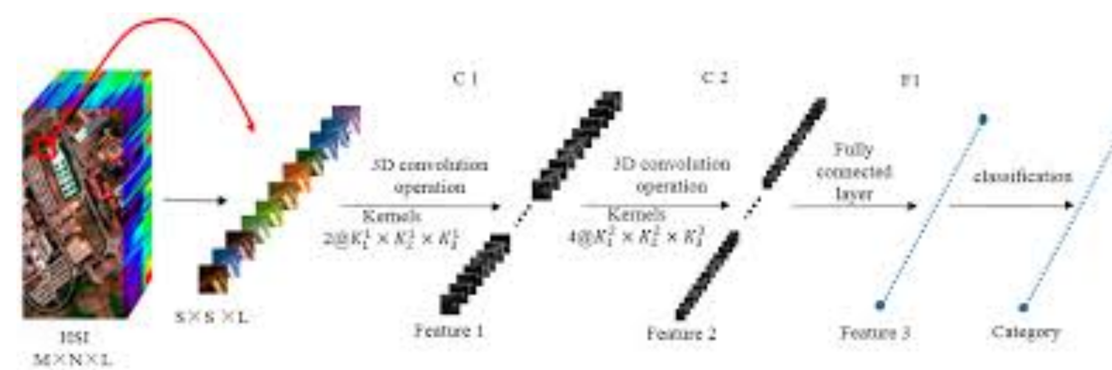
Satellite Image Time Series



Bi-temporal Change Detection

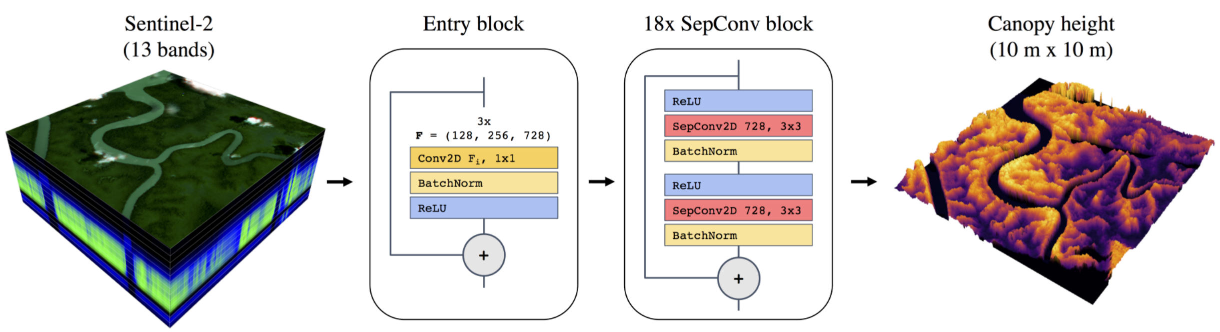


Hyperspectral Classification and Retrieval

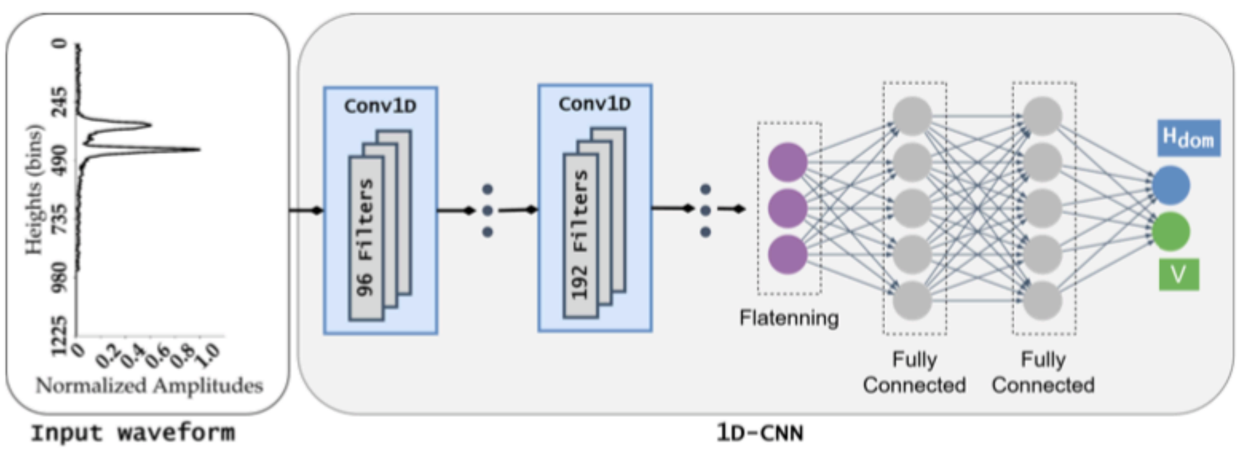


Deep Learning & EO data analysis

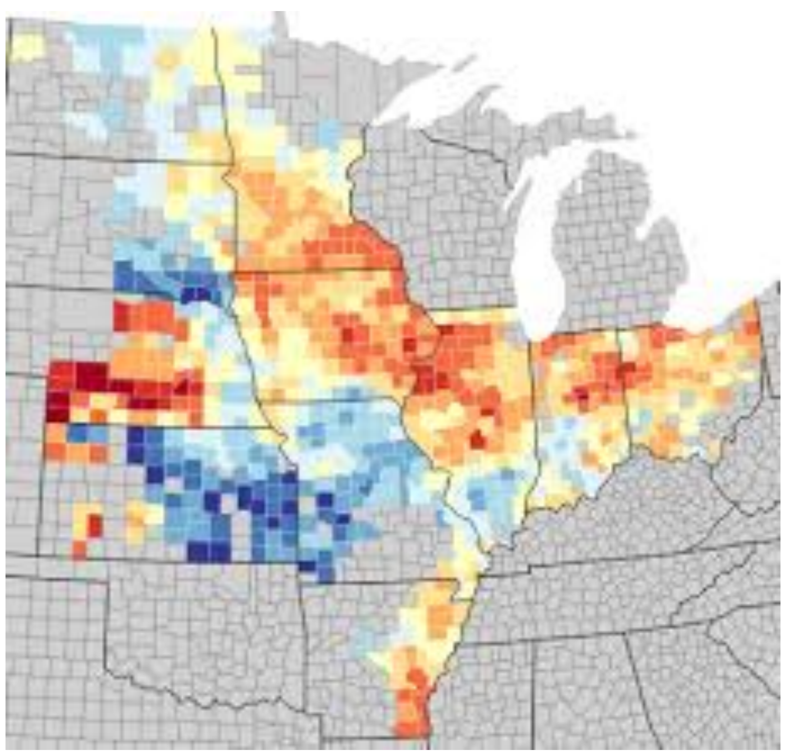
Canopy Height Estimation



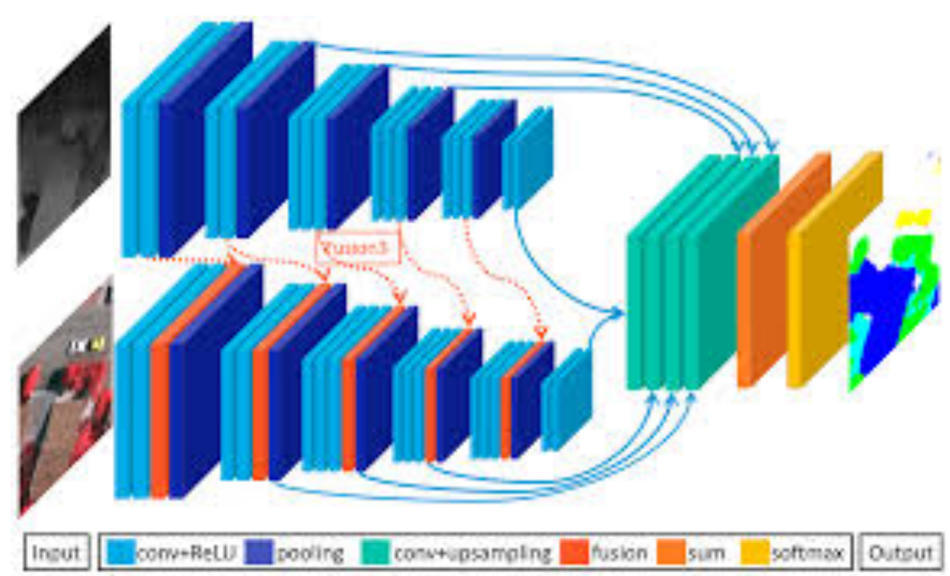
Forest Biomass Estimation



Crop Yield Prediction



Remote Sensing Data Fusion



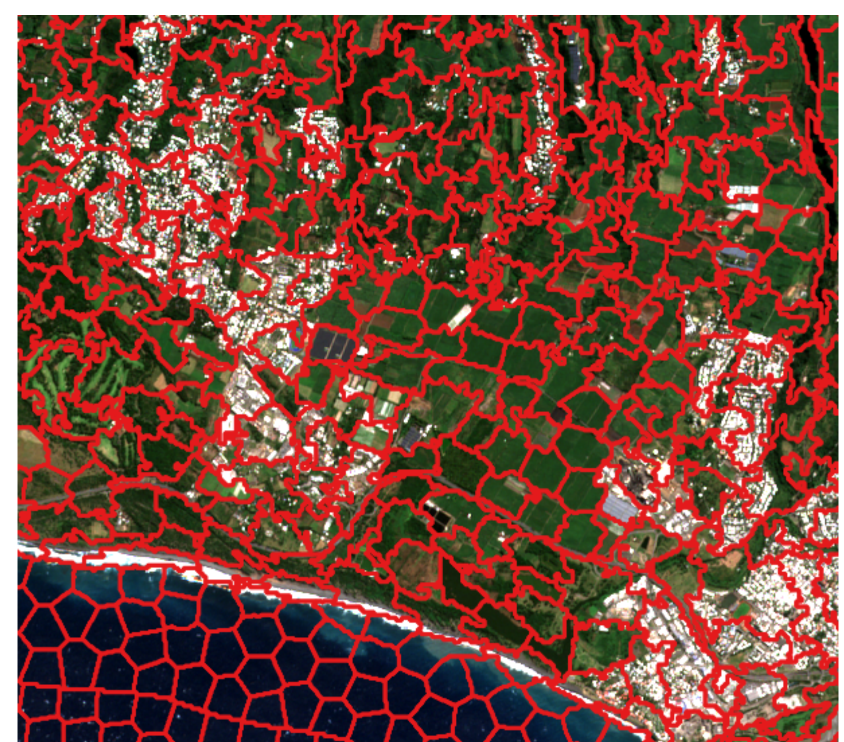
Pixel vs. Object analysis

When working on Earth Observation data, two different levels of granularity:

- Pixel: the base unit of image analysis
- Object: group of pixel (land unit) with an high level of semantic
  - Needs of a preprocessing step to extract object (segmentation)

Pixel

Object



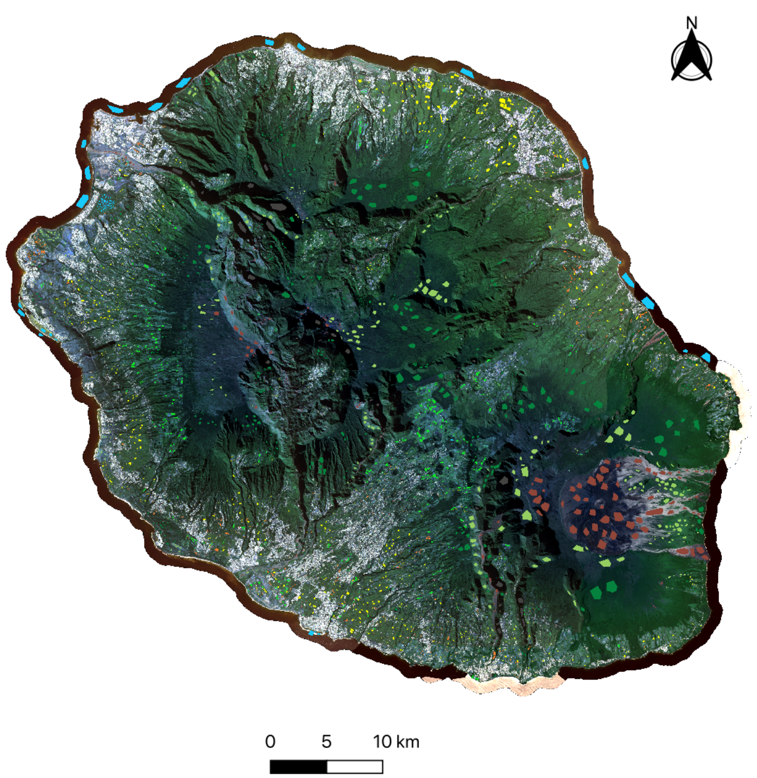
Land cover mapping task

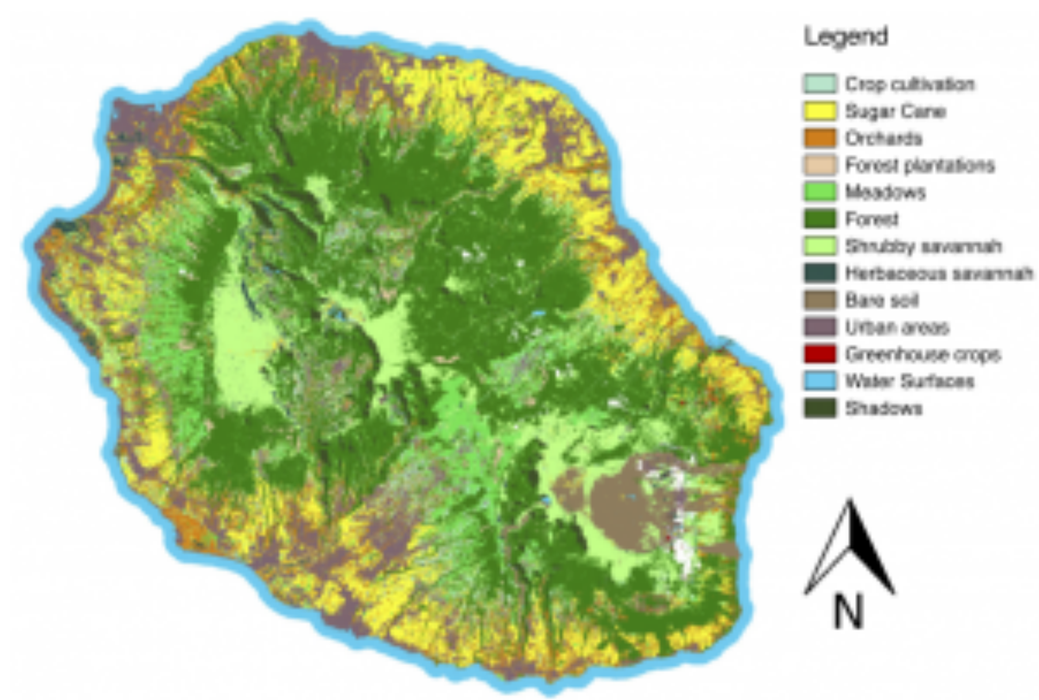
Task:

Given EO data + a limited number of reference data, the goal is to map each pixel (or object) to the corresponding land cover class

Common approach:

- Land cover mapping is addressed via **Machine Learning** methods.
- A ML method is **calibrated/trained on reference data** to classify the rest of pixels or objects (unlabelled data) that belongs to the same study area.





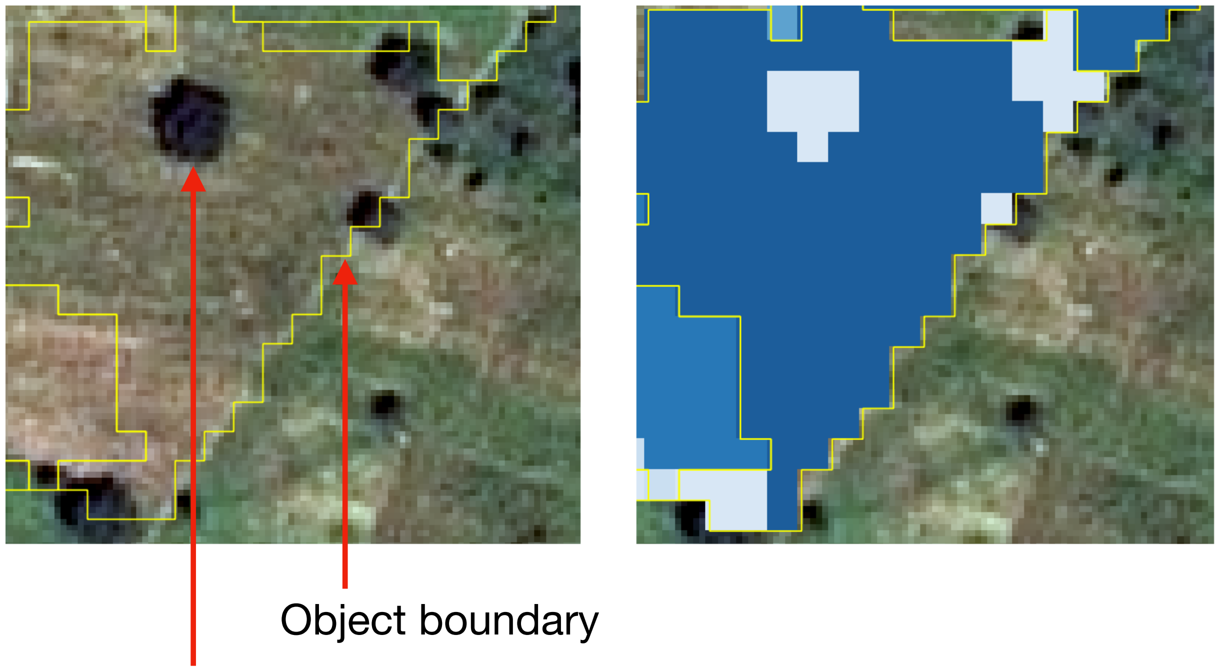
Two methodological challenges in object-based land cover mapping

An **Object** should be an homogeneous group of pixels but it can:

- Represent complex land unit (i.e. urban areas: built-up, garden, street, etc…)
- Be approximate or contain noise components that are unrelated with the major land cover class

Problem (1): intra-object heterogeneity

Agricultural Field



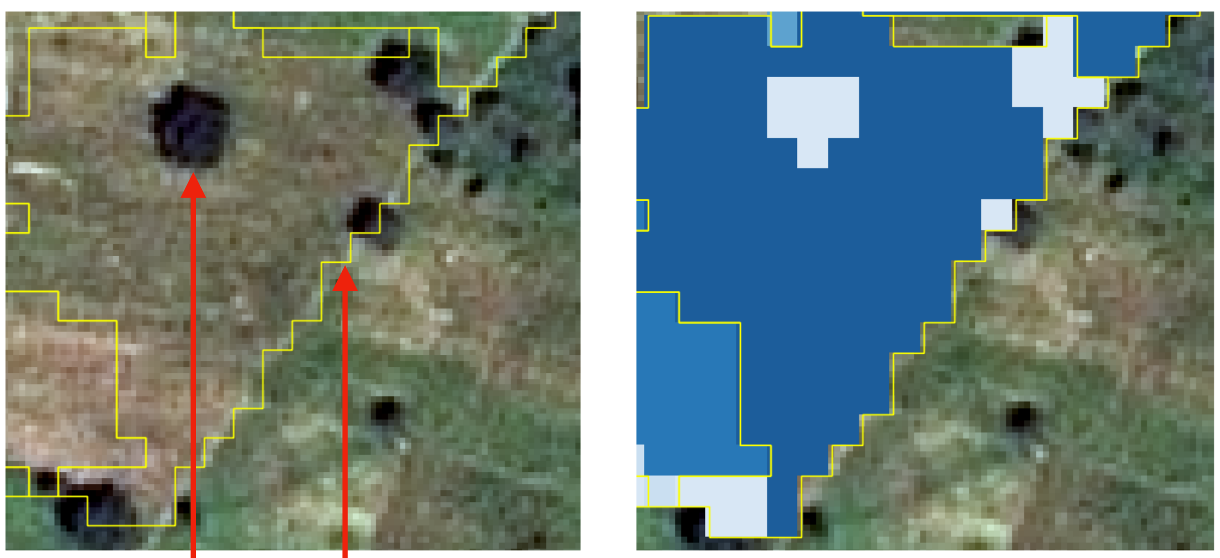
Two methodological challenges in object-based land cover mapping

An **Object** should be an homogeneous group of pixels but it can:

- Represent complex land unit (i.e. urban areas: built-up, garden, street, etc…)
- Be approximate or contain noise components that are unrelated with the major land cover class

Problem (1): intra-object heterogeneity

Agricultural Field

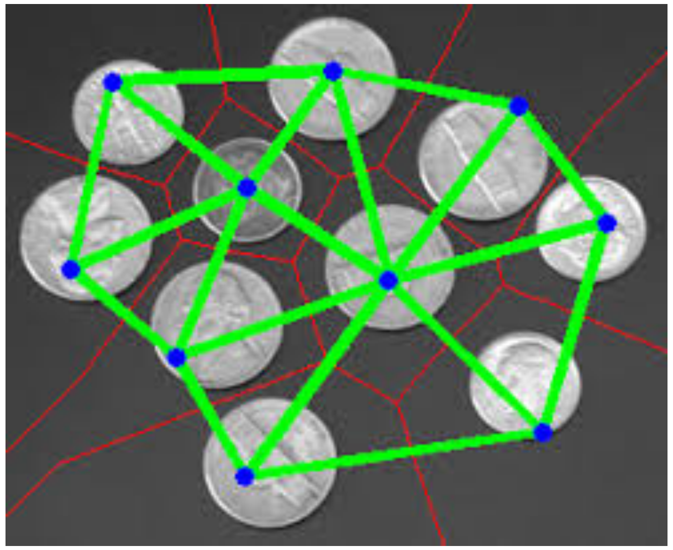


Object boundary

Noise components in the object

An **Object** is embedded in a landscape (spatial context):

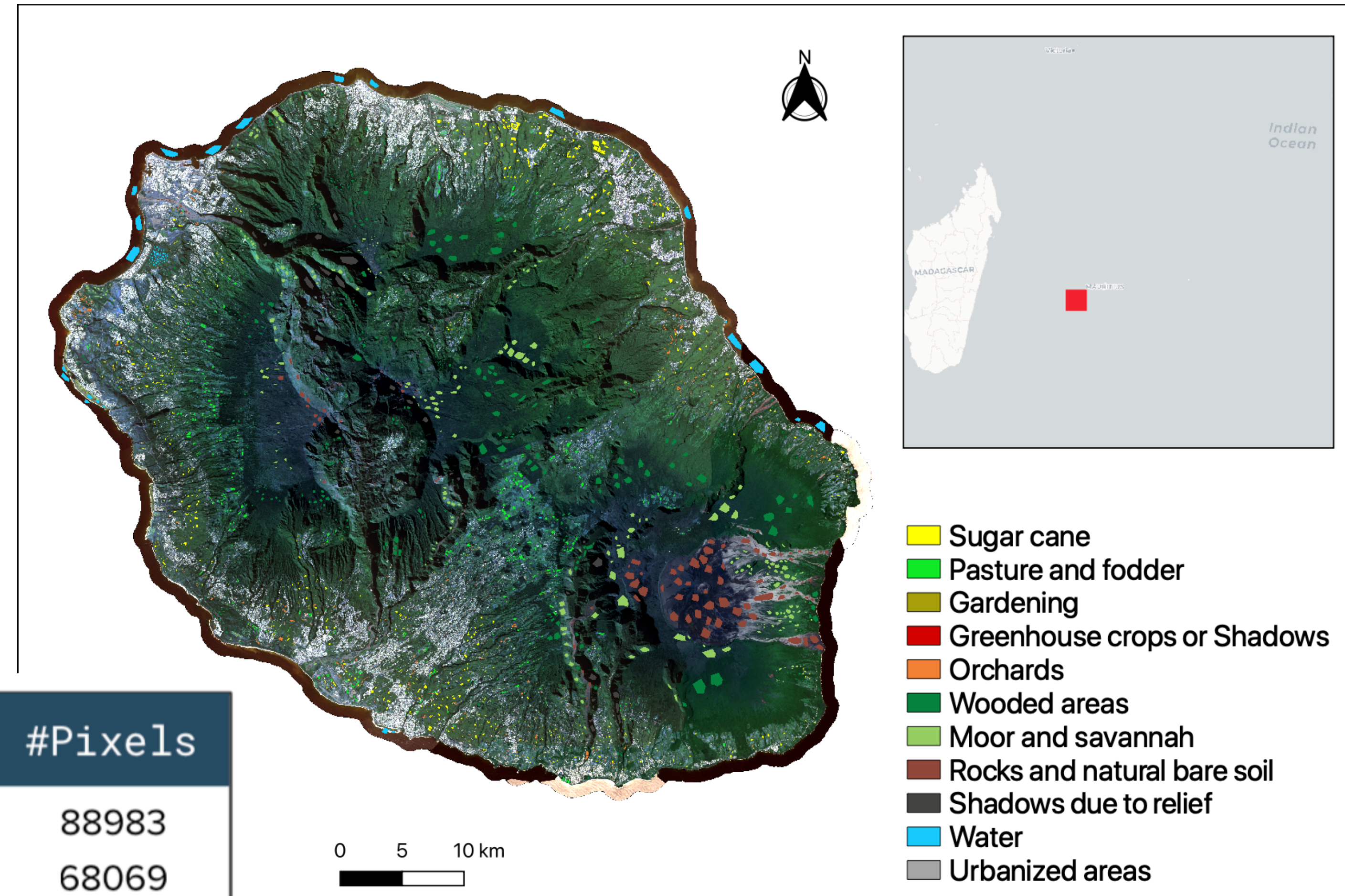
- It is usually neglected
- Difficult to manage due to the irregular neighbourhood (different number of neighbour segments)



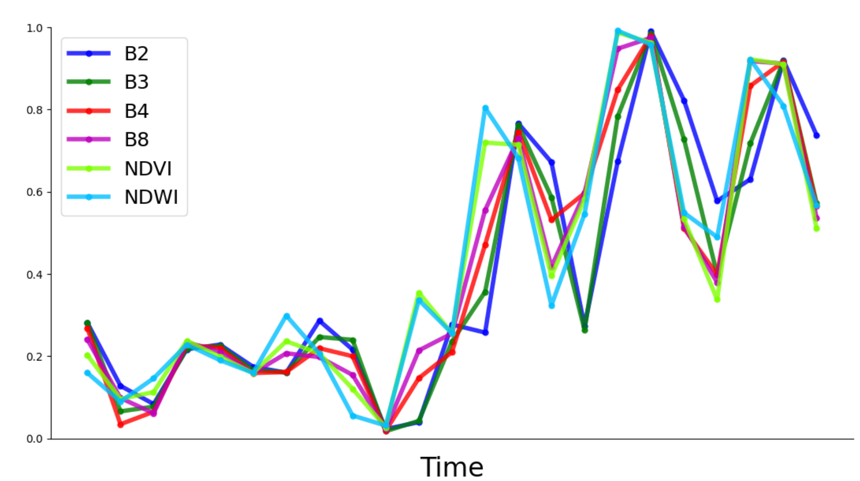
Problem (2): How to integrate the spatial context

Reunion Island case study

Surface: around 3000km²
Sentinel-2: 21 images
Image size: 6656 x 5913
 # Bands: 6
 # LC classes: 11
Amount of data: 19Gb



Class	Label	#Polygons	#Objects	#Pixels
0	Sugar Cane	869	1466	88983
1	Pasture and fodder	582	1042	68069
2	Market Gardening	758	1038	17574
3	Green. Crops	260	308	1928
4	Orchards	767	1174	33694
5	Wooded areas	570	1467	205050
6	Moor and Savannah	506	1172	155229
7	Rocks and bare soil	299	845	154283
8	Relief shadows	81	248	54308
9	Water	177	458	82547
10	Urbanized areas	1396	1360	19004
Total		**6265**	**10578**	**880669**



TASSEL

Does intra-object variability/heterogeneity affect Satellite Image Time Series based land cover mapping?

D. Ienco, Y. J. E. Gbodjo, R. Gaetano, R. Interdonato: **Weakly Supervised Learning for Land Cover Mapping of Satellite Image Time Series via Attention-Based CNN**. IEEE Access 8: 179547-179560 (2020)

TASSEL

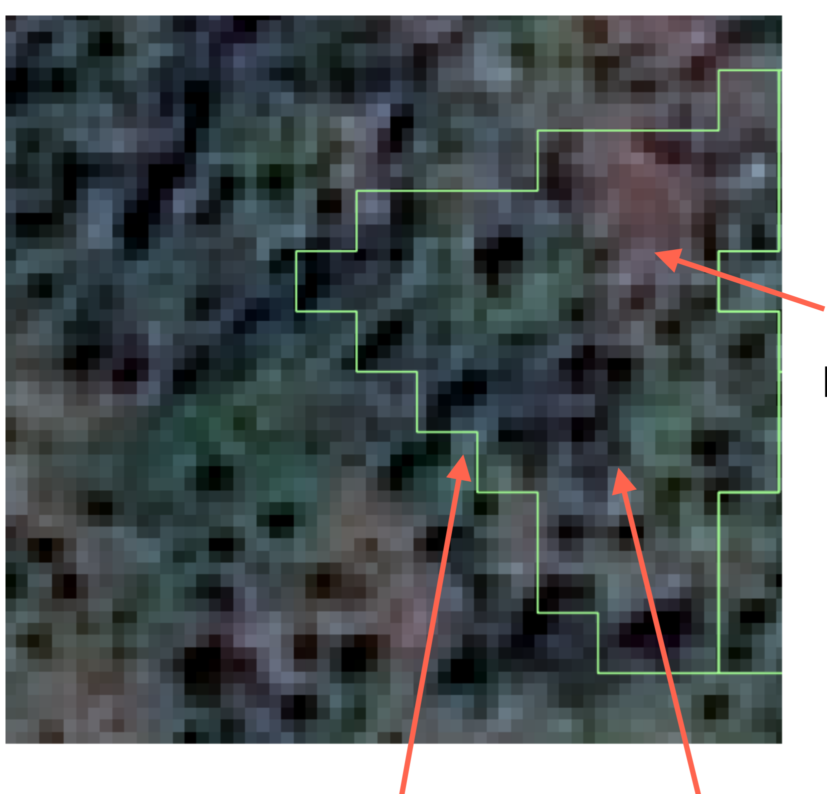
How to manage intra-object heterogeneity

Explicitly take into account:

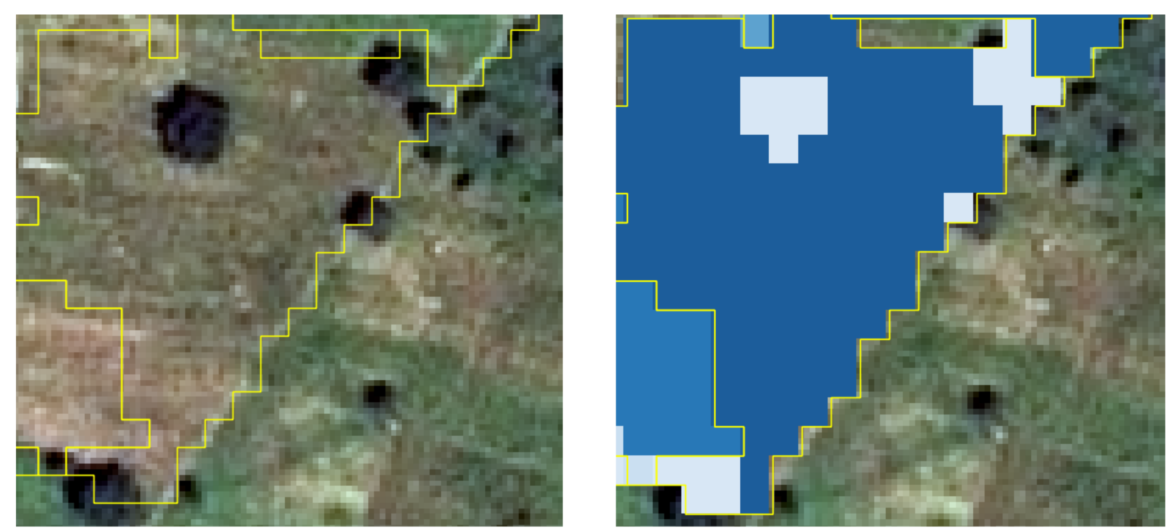
- The intra-object heterogeneity
- Problem related to approximate or inexact annotation
- Land-unit involving multifaceted information

<u>Manage object as a set of components</u>

Forest Object



Crop Object



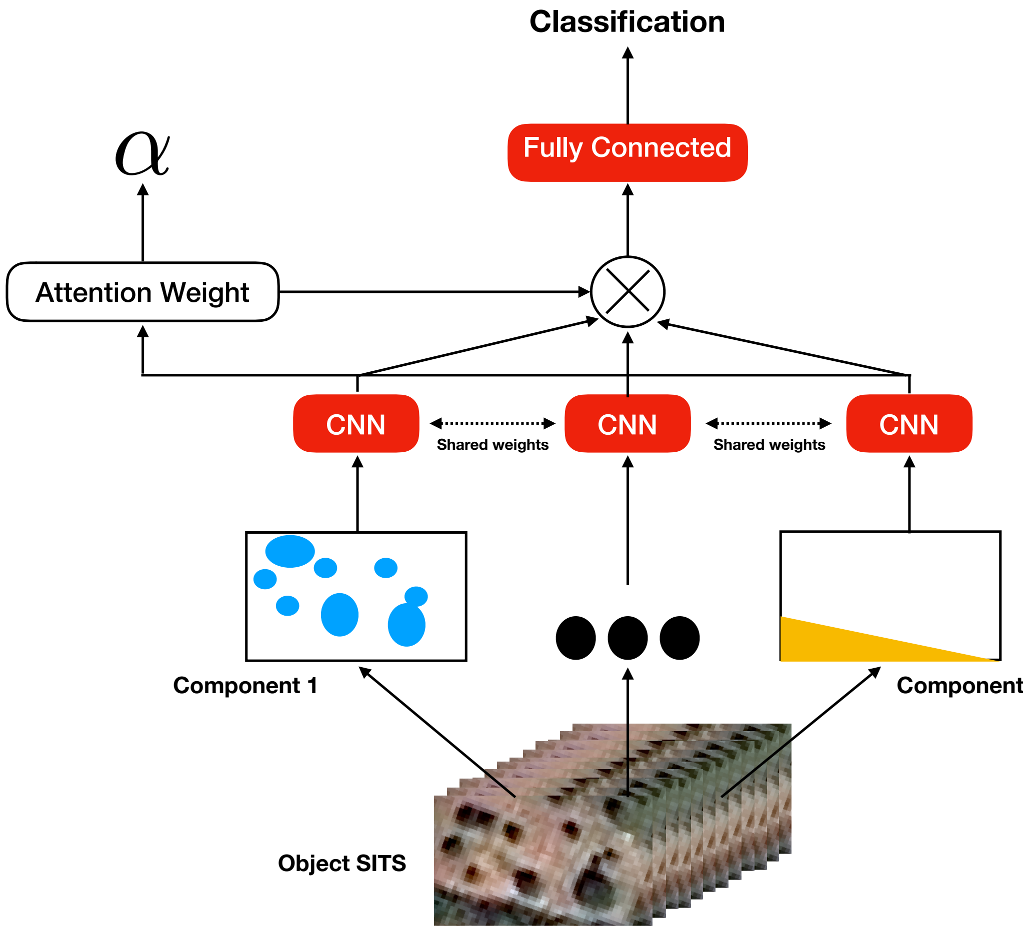
Component contribution to the final decision

TASSEL

How to manage intra-object heterogeneity

Method Description

- Identify components for each object (K-means)
- Use Convolutional Neural Networks (CNN1D) to manage per-component information
- Aggregate per-component representation to take the final decision



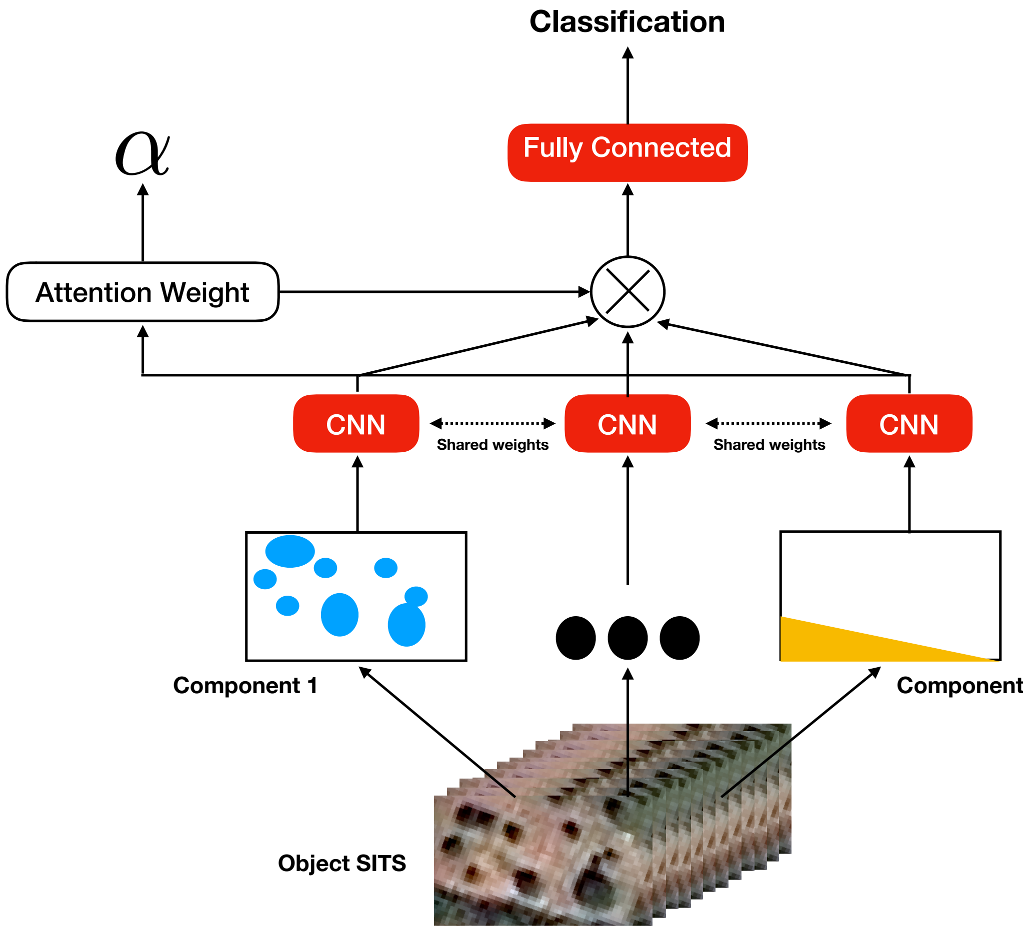
TASSEL

How to manage intra-object heterogeneity

Method

Method Description

- Identify components for each object (K-means)
- Use Convolutional Neural Networks (CNN1D) to manage per-component information
- Aggregate per-component representation to take the final decision



The output of TASSEL is twofold:

- A **classification** for each object Satellite Image Time Series
- An **attention weight** in the range [0,1] associated to each component that can be interpreted as the contribution of that component to the decision process

TASSEL
How to manage intra-object heterogeneity

Method

The attention mechanism

Given $H = \{h_1, ..., h_l\}$ the set of all the components representations

$$\tilde{h} = \sum_{l=1}^{L} \alpha_l \cdot h_l$$

Where

$$\alpha_l = \frac{exp(v_a^{\intercal}\ tanh(W_a\, h_l + b_a))}{\sum_{l'=1}^{L} exp(v_a^{\intercal}\ tanh(W_a\, h_{l'} + b_a))}$$

TASSEL

How to manage intra-object heterogeneity

Method

The attention mechanism

Given $H = \{h_1, ..., h_l\}$ the set of all the components representations

$$\tilde{h} = \sum_{l=1}^{L} \alpha_l \cdot h_l$$

Where

$$\alpha_l = \frac{exp(v_a^\intercal \; tanh(W_a \, h_l + b_a))}{\sum_{l'=1}^{L} exp(v_a^\intercal \; tanh(W_a \, h_{l'} + b_a))}$$

The purpose of this procedure is to learn a set of weights $(\alpha_1, \ldots, \alpha_L)$ to estimate the contribution of each component

TASSEL

How to manage intra-object heterogeneity

Results

Experimental Settings:

- We compare TASSEL w.r.t. standard competitors: RF, LSTM, MLP, CNN
- We employ standard evaluation measures: F1-score, Kappa and Accuracy
- We divided the dataset in training/validation/test (50%/30%/20%) and repeat 5 times

Results

Experimental Settings:

- We compare TASSEL w.r.t. standard competitors: RF, LSTM, MLP, CNN
- We employ standard evaluation measures: F1-score, Kappa and Accuracy
- We divided the dataset in training/validation/test (50%/30%/20%) and repeat 5 times

Competitors work **on the average object representation** without considering object components

	F1 Score	*Kappa*	*Accuracy*
RF	81.74 ± 0.47	0.7991 ± 0.0052	82.13 ± 0.46
LSTM	82.91 ± 0.66	0.8098 ± 0.0078	83.06 ± 0.69
MLP	85.81 ± 0.60	0.8423 ± 0.0074	85.94 ± 0.66
CNN	87.11 ± 0.61	0.8565 ± 0.0068	87.20 ± 0.61
TASSEL	$\mathbf{89.13} \pm 0.62$	$\mathbf{0.8797} \pm 0.0072$	$\mathbf{89.28} \pm 0.63$

Results

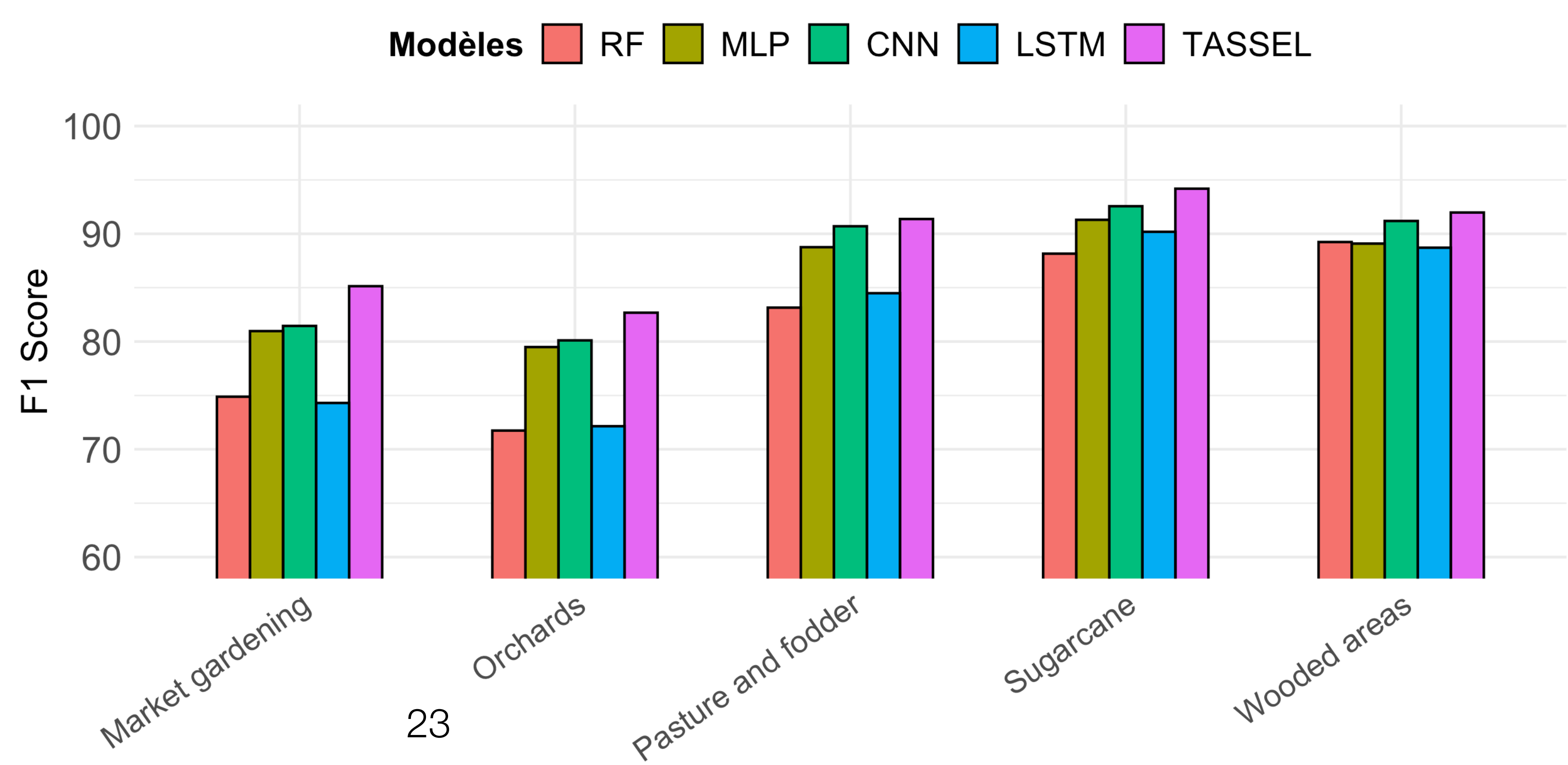
Experimental Settings:

- We compare TASSEL w.r.t. standard competitors: RF, LSTM, MLP, CNN
- We employ standard evaluation measures: F1-score, Kappa and Accuracy
- We divided the dataset in training/validation/test (50%/30%/20%) and repeat 5 times

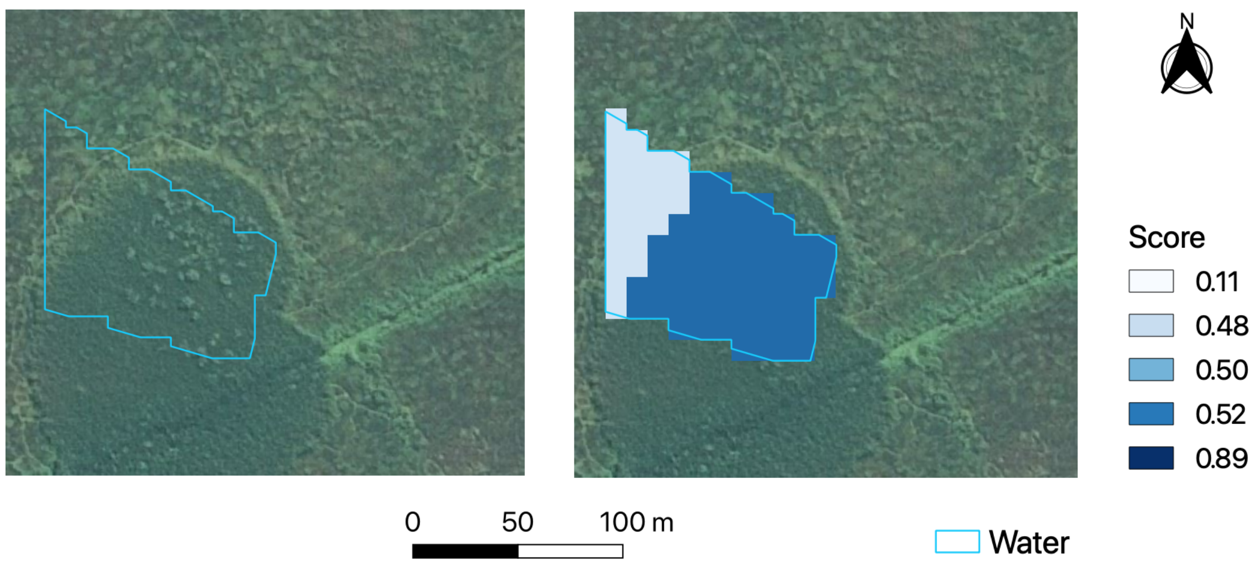
Competitors work **on the average object representation** without considering object components

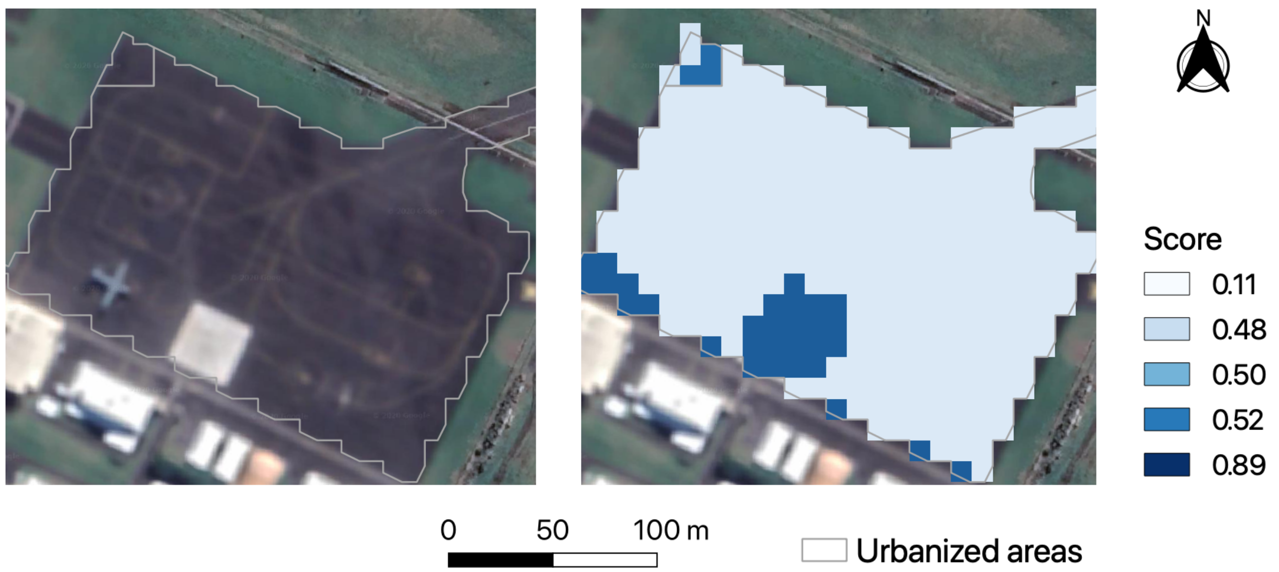
	F1 Score	*Kappa*	*Accuracy*
RF	81.74 ± 0.47	0.7991 ± 0.0052	82.13 ± 0.46
LSTM	82.91 ± 0.66	0.8098 ± 0.0078	83.06 ± 0.69
MLP	85.81 ± 0.60	0.8423 ± 0.0074	85.94 ± 0.66
CNN	87.11 ± 0.61	0.8565 ± 0.0068	87.20 ± 0.61
TASSEL	**89.13** ± 0.62	**0.8797** ± 0.0072	**89.28** ± 0.63

We observe relative improvement on all the natural/agricultural classes.



Interpret model decision by attention weight on the object components



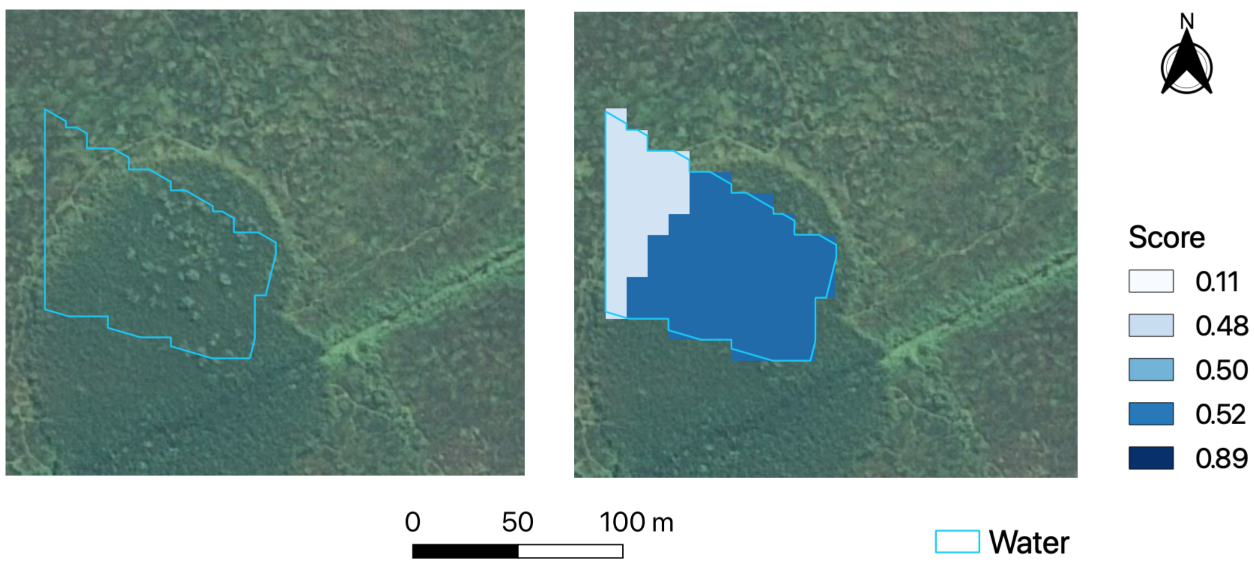


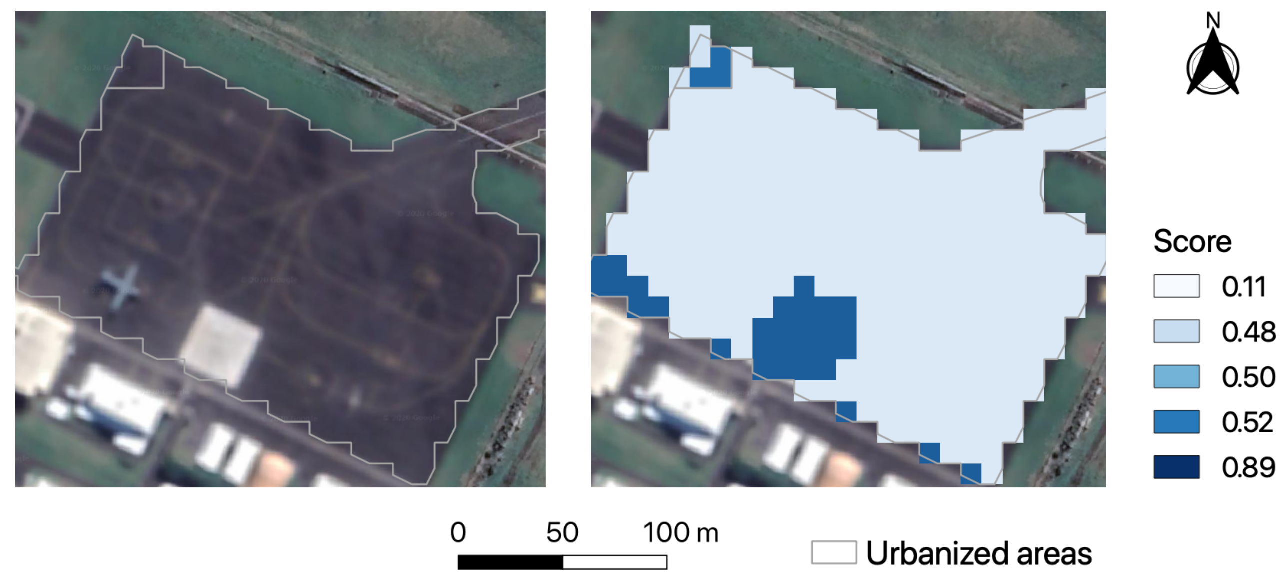
TASSEL

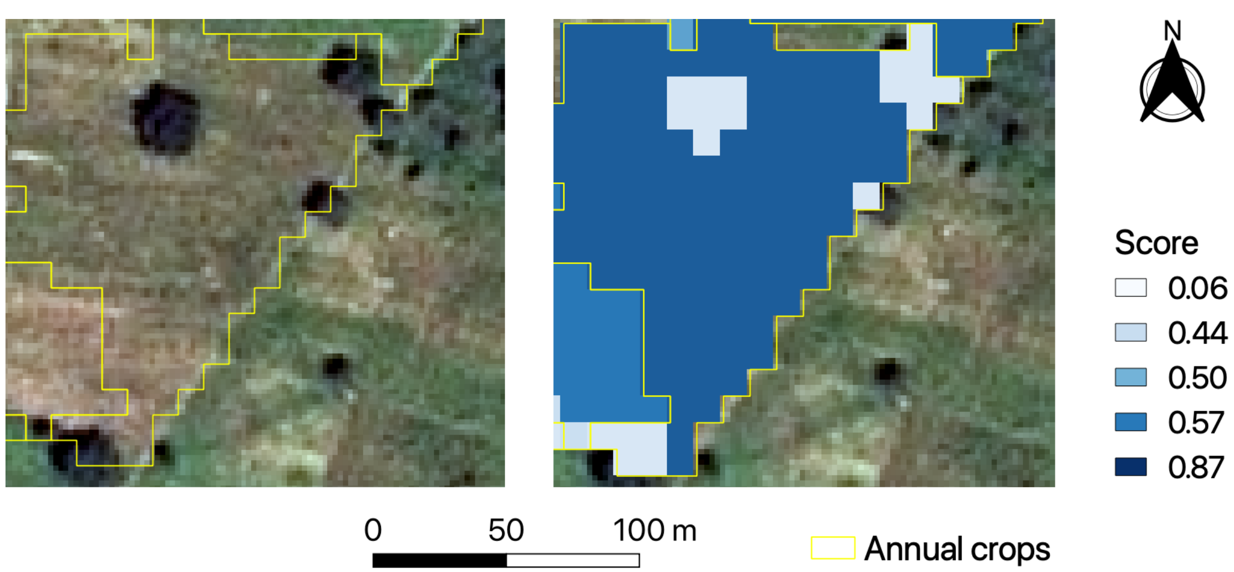
How to manage intra-object heterogeneity

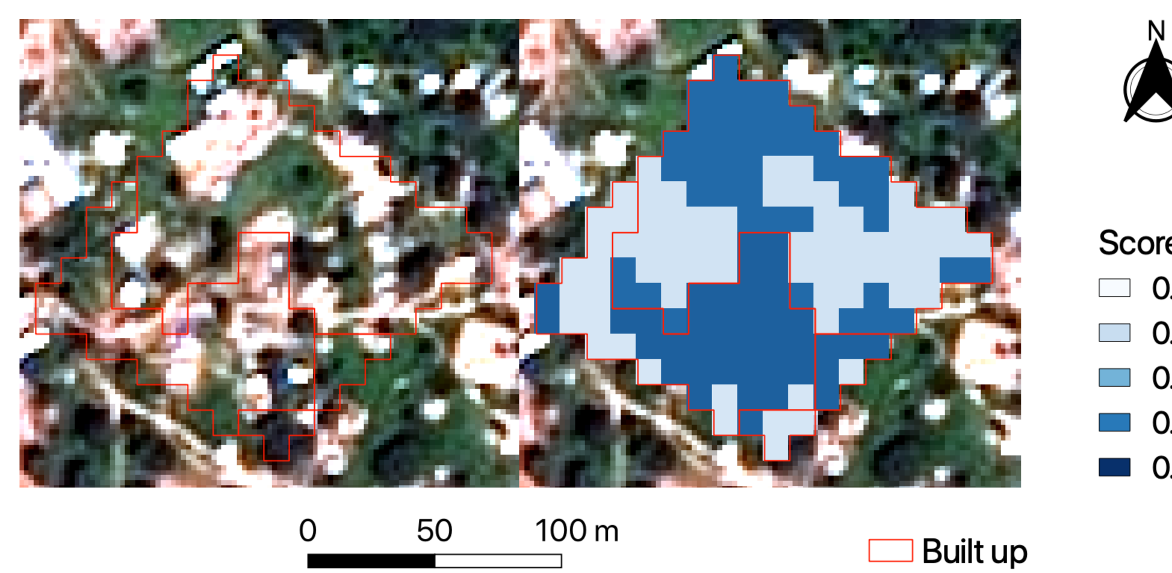
Results

Interpret model decision by attention weight on the object components









Example coming from another study sites (Bourkina Faso)

TASSEL

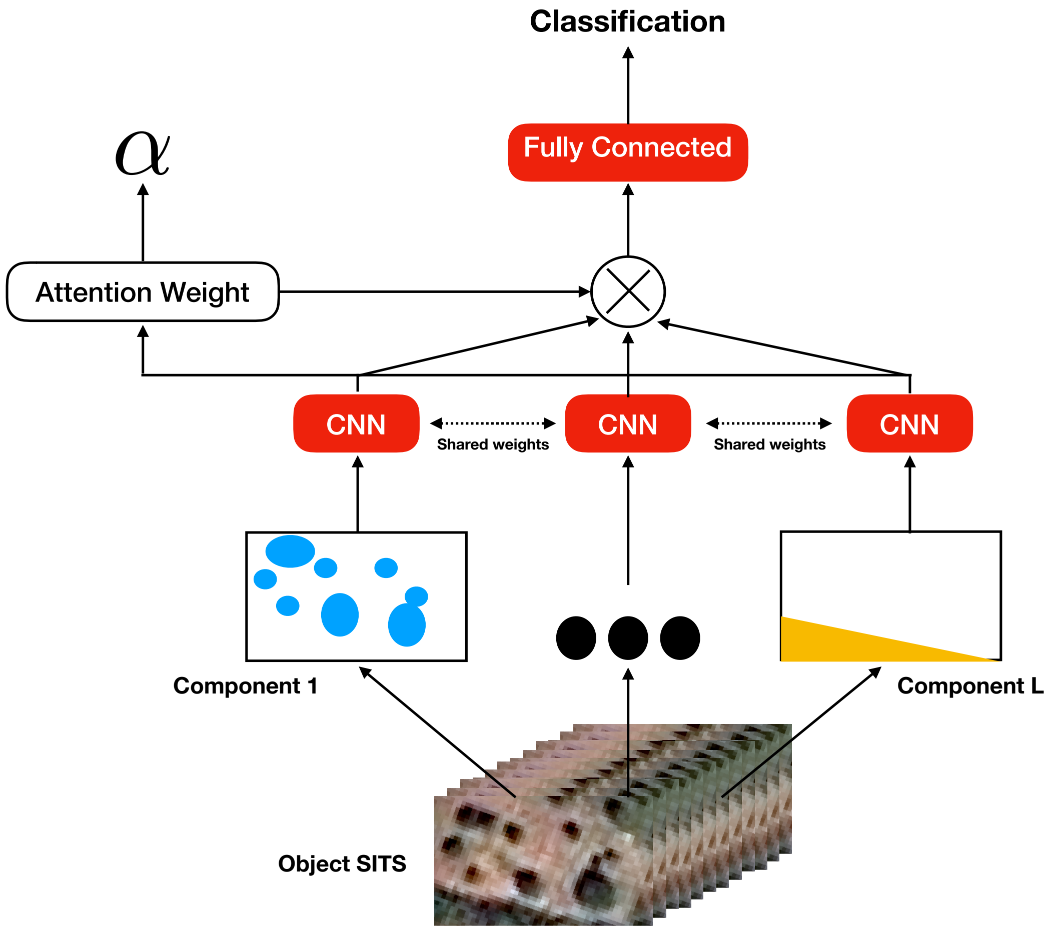
How to manage intra-object heterogeneity

Conclusions

Experimental results support the intuition to **explicitly manage intra-object heterogeneity**

The TASSEL model also supplies **“a kind of” interpretation** about its decision

The main gain are obtained considering **agricultural land cover classes** that exhibits mixed or complex spatial patterns



STARCANE

Does the spatial context matter for land cover mapping
via Satellite Image Time Series data?

A. M. Censi, D. Ienco, Y. J. E. Gbodjo, R. G. Pensa, R. Interdonato, R. Gaetano: **Attentive Spatial Temporal Graph CNN for Land Cover Mapping From Multi Temporal Remote Sensing Data**. IEEE Access 9: 23070-23082 (2021)

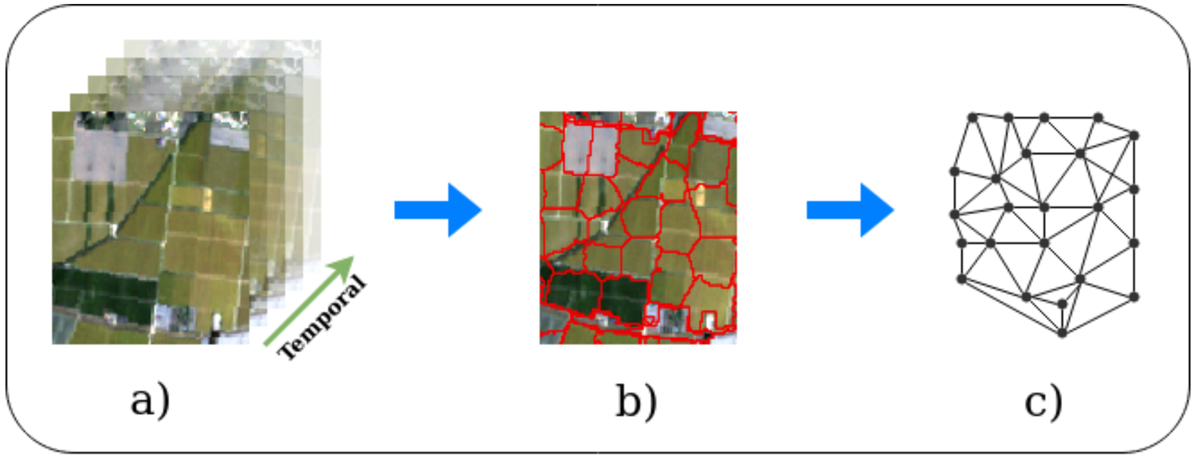
Integrate the landscape (spatial context) in which an object is embedded

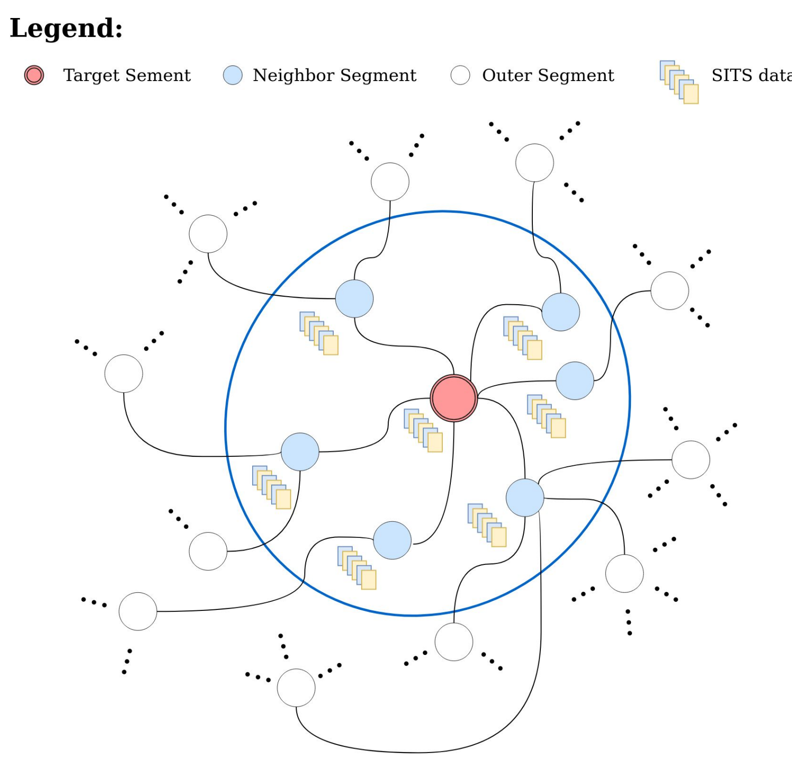
Method Description

- From the segmentation we derive a **Region Adjacency Graph**

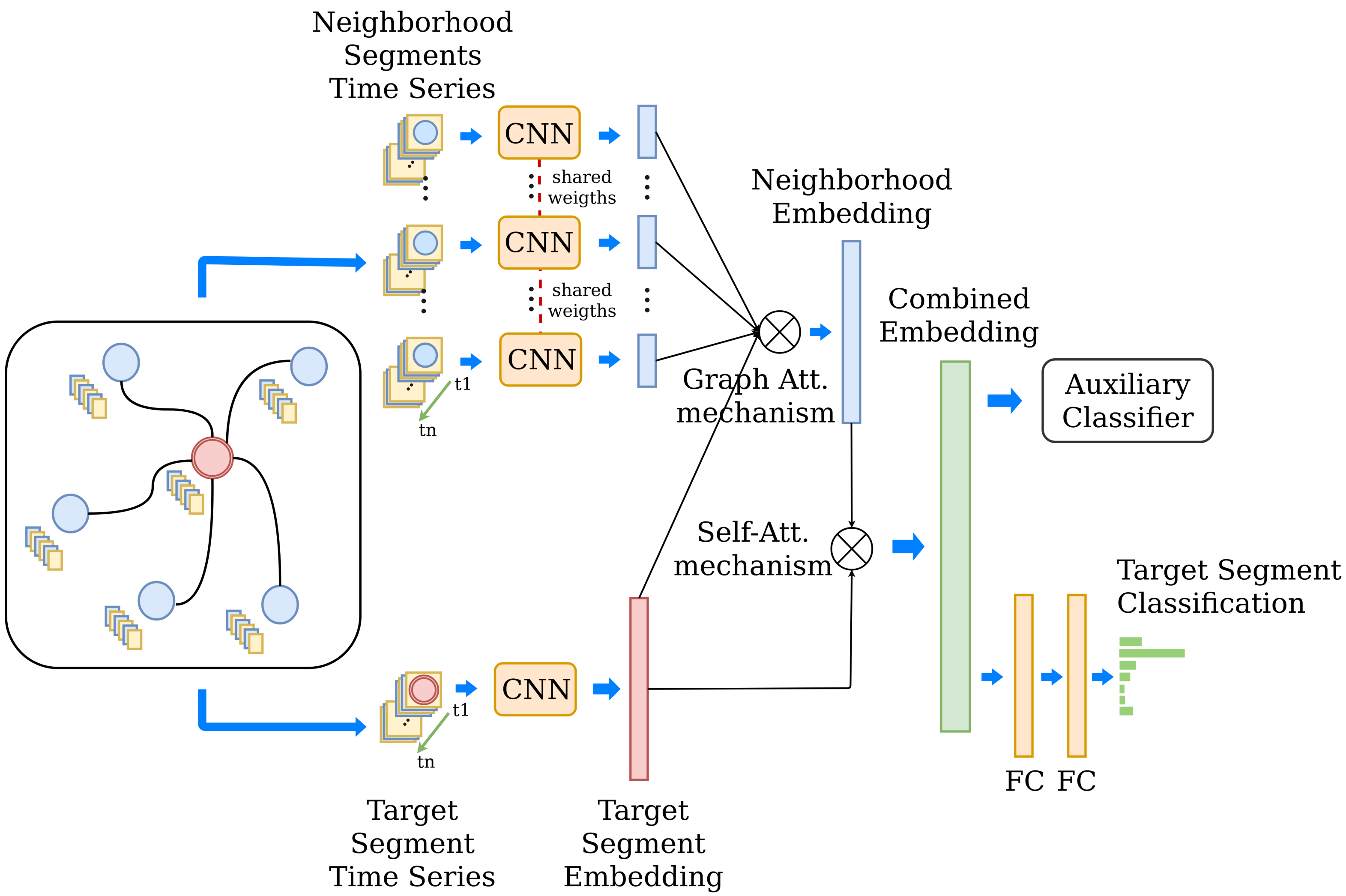
- **Spatio-Temporal Graph Convolutional Neural Network** to manage, simultaneously, the target SITS object as well as the neigh. SITS objects information

- **Automatically weight** the **neigh. objects contribution** belonging to the spatial context w.r.t. the target node





- A (1D) **CNN architecture** is employed to manage object-level time series.
- For the neighbourhood set, another (1D) **CNN (with shared weights)** is employed.
- An **attention mechanism** is employed to **weight differently the contribution of each neighbour** in the aggregation.
- Finally, the classification is obtained after **combining together the embedding of the target object** and **the one obtained by the neighbours.**



Attention Mechanism

$$h^i_{neigh} = |N(v_i)| \cdot \sum_{v_j \in N(v_i)} \alpha_{ij} \cdot h_{v_j}$$

$$\alpha_{ij} = \frac{exp(LReLU(a^T[Wh_{v_i}||Wh_{v_j}]))}{\sum_{v_k \in N(v_i)} exp(LReLU(a^T[Wh_{v_i}||Wh_{v_k}]))}$$

Results

We compare STARCANE w.r.t. standard competitors: RF, LSTM, MLP, CNN that not consider spatial context
We employ standard evaluation measures: F1-score, Kappa and Accuracy
We divided the dataset in training/validation/test (50%/30%/20%) and repeat 5 times

STARCANE

Does spatial context matter?

Results

We compare STARCANE w.r.t. standard competitors: RF, LSTM, MLP, CNN that not consider spatial context
We employ standard evaluation measures: F1-score, Kappa and Accuracy
We divided the dataset in training/validation/test (50%/30%/20%) and repeat 5 times

The competing approaches does not (cannot) use the spatial context information

	F1 Score	*Kappa*	*Accuracy*
RF	82.43 ± 0.15	80.65 ± 0.17	82.79 ± 0.15
MLP	80.78 ± 0.53	78.60 ± 0.52	80.96 ± 0.46
CNN	84.40 ± 0.37	82.73 ± 0.45	84.62 ± 0.41
LSTM	83.36 ± 0.57	81.41 ± 0.71	83.44 ± 0.64
STARCANE	**90.50** ± 0.1	**89.37** ± 0.08	**90.52** ± 0.08

STARCANE

Does spatial context matter?

Results

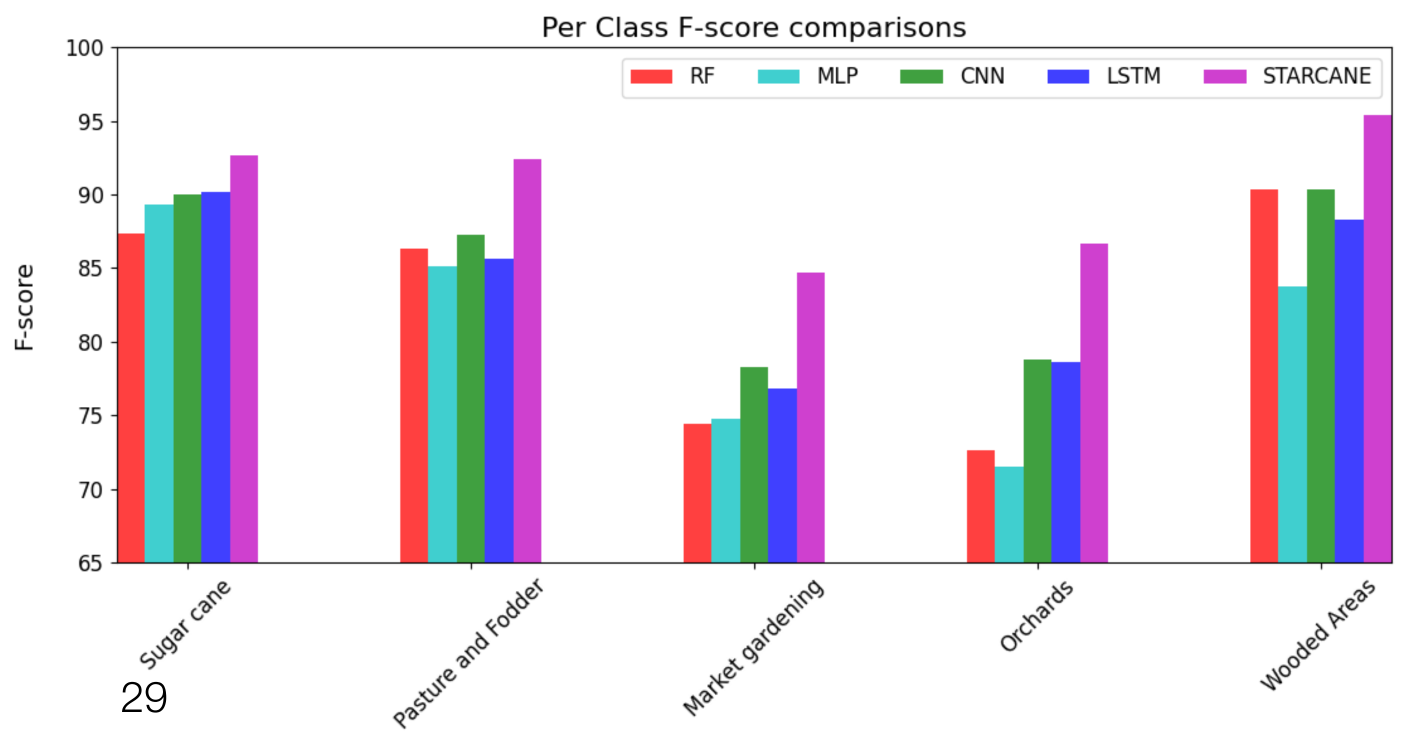
We compare STARCANE w.r.t. standard competitors: RF, LSTM, MLP, CNN that not consider spatial context

We employ standard evaluation measures: F1-score, Kappa and Accuracy

We divided the dataset in training/validation/test (50%/30%/20%) and repeat 5 times

The competing approaches does not (cannot) use the spatial context information

	F1 Score	*Kappa*	*Accuracy*
RF	82.43 ± 0.15	80.65 ± 0.17	82.79 ± 0.15
MLP	80.78 ± 0.53	78.60 ± 0.52	80.96 ± 0.46
CNN	84.40 ± 0.37	82.73 ± 0.45	84.62 ± 0.41
LSTM	83.36 ± 0.57	81.41 ± 0.71	83.44 ± 0.64
STARCANE	$\mathbf{90.50} \pm 0.1$	$\mathbf{89.37} \pm 0.08$	$\mathbf{90.52} \pm 0.08$



STARCANE

Does spatial context matter?

Results

We compare STARCANE w.r.t. standard competitors: RF, LSTM, MLP, CNN that not consider spatial context

We employ standard evaluation measures: F1-score, Kappa and Accuracy

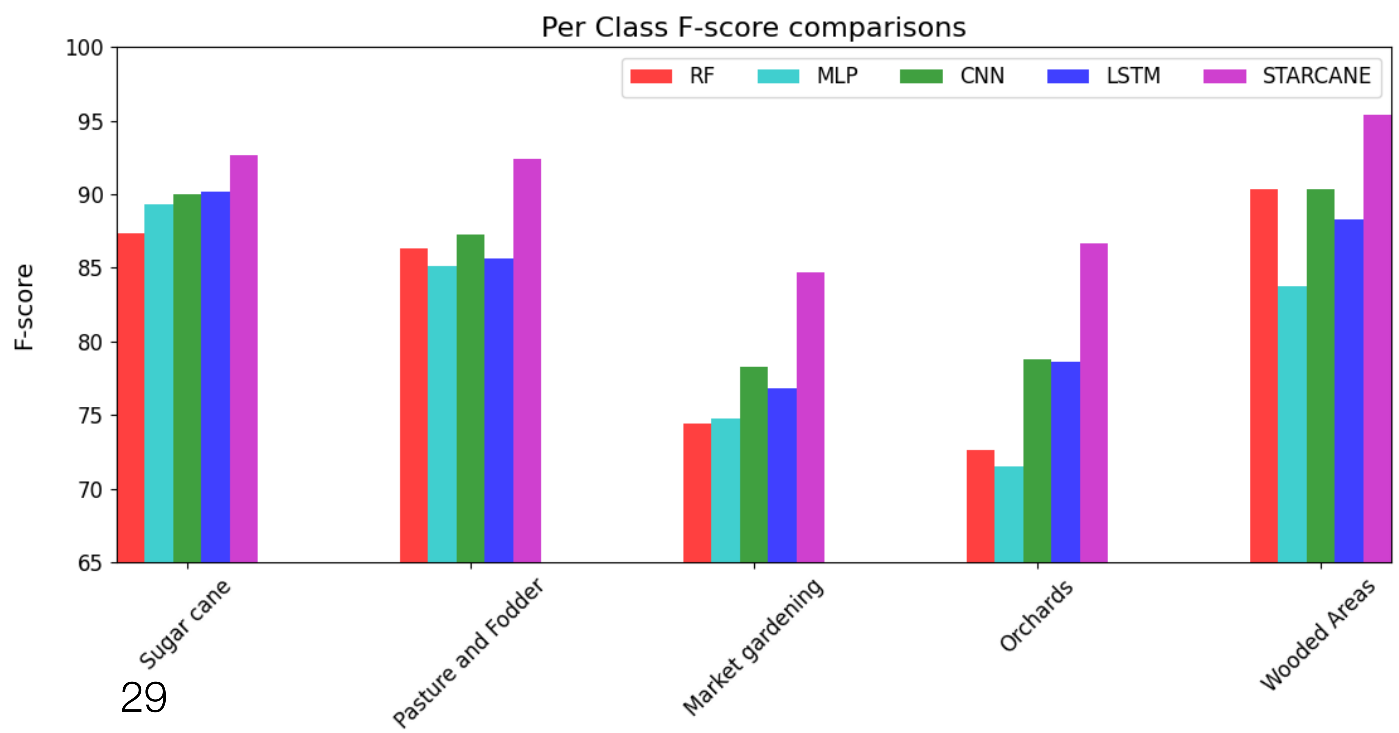
We divided the dataset in training/validation/test (50%/30%/20%) and repeat 5 times

The competing approaches does not (cannot) use the spatial context information

	F1 Score	*Kappa*	*Accuracy*
RF	82.43 ± 0.15	80.65 ± 0.17	82.79 ± 0.15
MLP	80.78 ± 0.53	78.60 ± 0.52	80.96 ± 0.46
CNN	84.40 ± 0.37	82.73 ± 0.45	84.62 ± 0.41
LSTM	83.36 ± 0.57	81.41 ± 0.71	83.44 ± 0.64
STARCANE	**90.50** ± 0.1	**89.37** ± 0.08	**90.52** ± 0.08

Gain can be observed considering **all the LC classes.**

Regarding agricultural and natural LC, STARCANE has notable improvement d**ue to the use of spatial context.**



STARCANE

Does spatial context matter?

Results

Due to the ability of STARCANE to weight the contribution of neigh. objects:

- For a land cover class, **we analyse the spatial (pattern) co-occurrence** of the land cover classes in the surrounding
- We can **sort the objects in the spatial context** considering the attention/contrib. weight

STARCANE

Does spatial context matter?

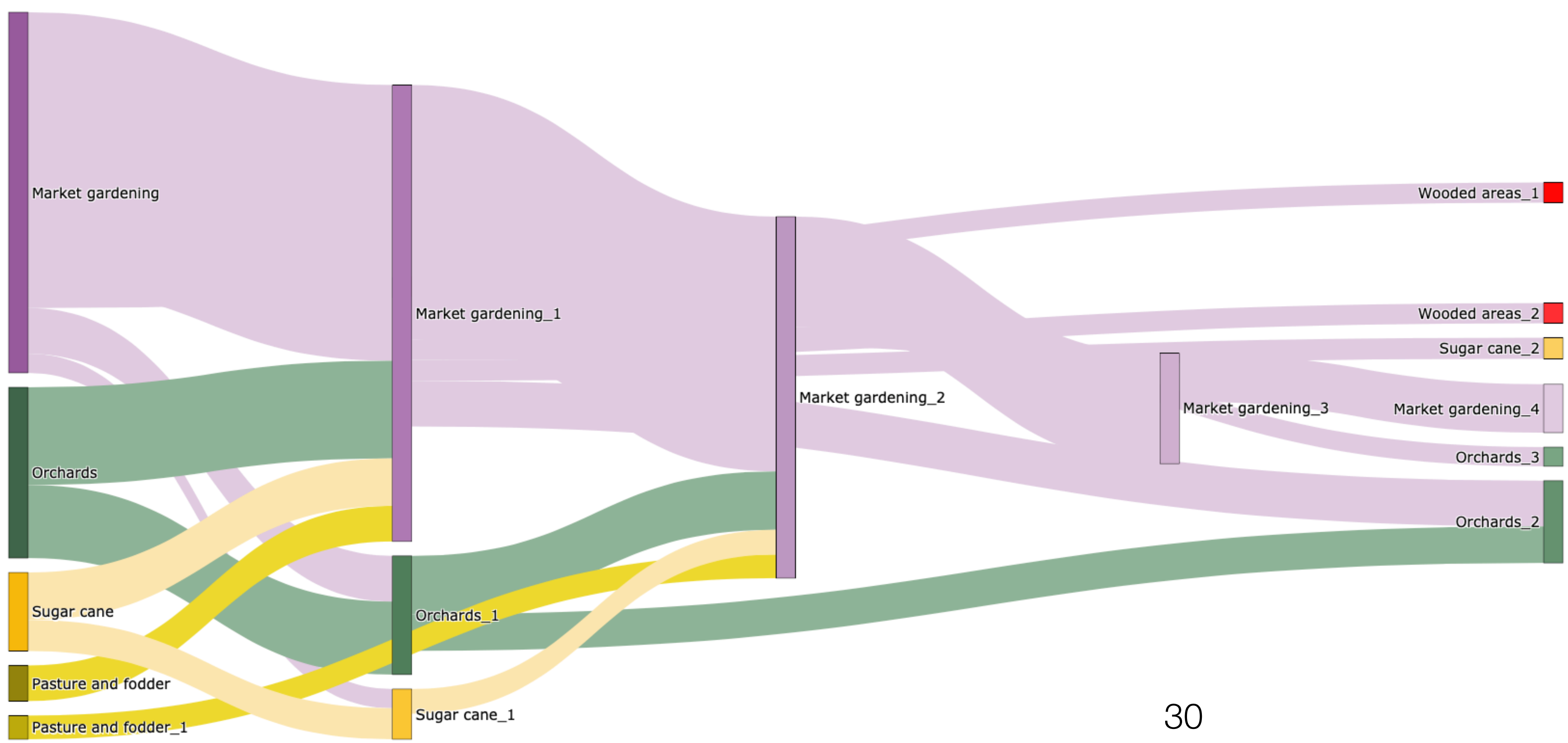
Results

Due to the ability of STARCANE to weight the contribution of neigh. objects:

- For a land cover class, **we analyse the spatial (pattern) co-occurrence** of the land cover classes in the surrounding
- We can **sort the objects in the spatial context** considering the attention/contrib. weight

From the most important neighbour to the least important

Spatial context related to objects classified as **Market Gardening**

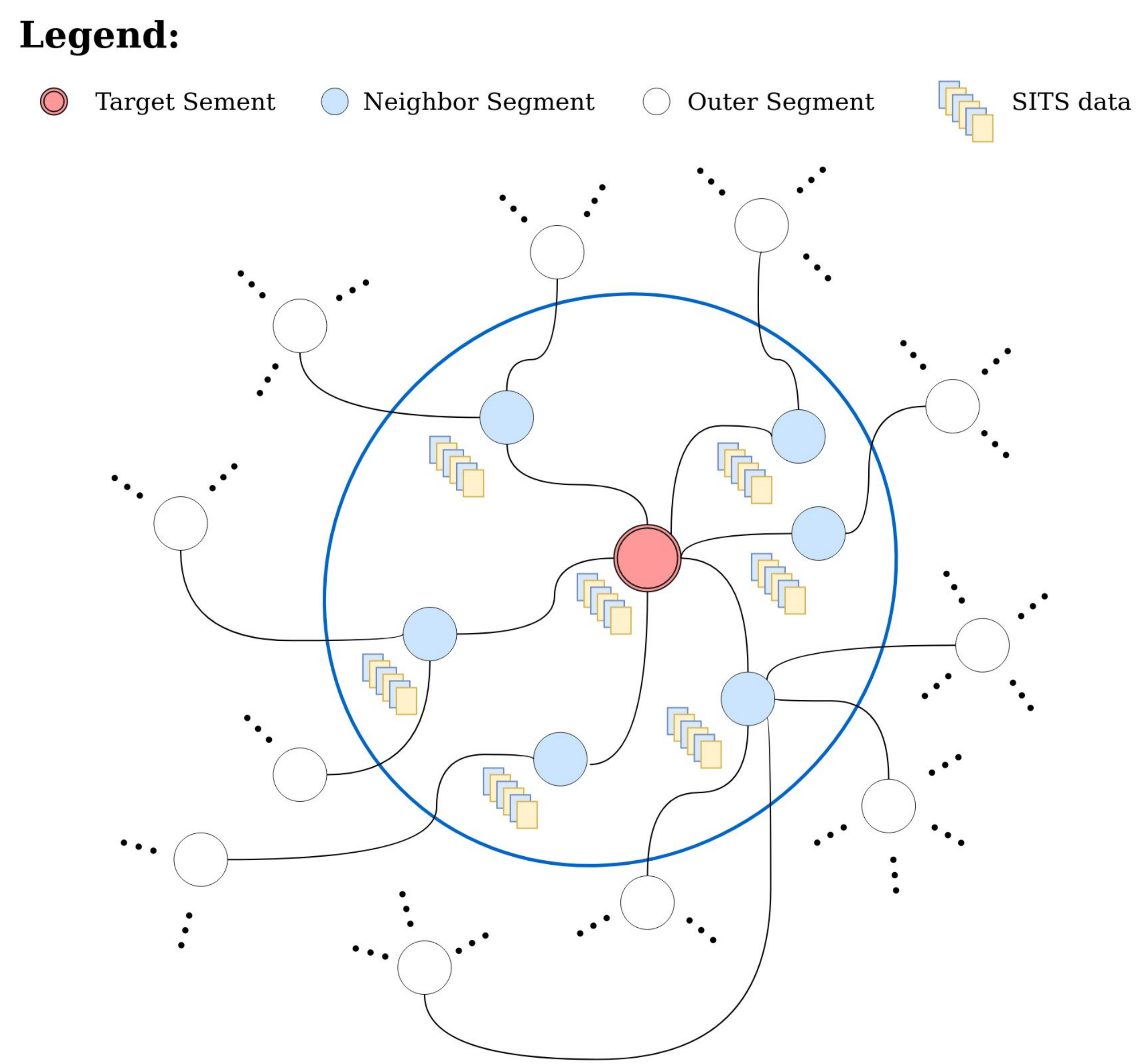


Conclusions

Experimental results support the intuition that **spatial context matters in land cover mapping** through SITS data

The STARCANE model provides **information** about the **neighbourhood importance in its decision**

Gain are systematically obtained on all the land cover classes. The spatial context allows to **reduce ambiguity**, in particular, on agricultural classes.



To wrap up

Earth Observation data is **a valuable information source** to support agricultural monitoring systems at medium and large scale:

- Support **public policy**
- Map **natural resources**

Among all the EO data, Satellite Image Time Series offer new opportunities to monitor the Earth Surface evolution and provide insights in agricultural productions.

To wrap up

Earth Observation data is **a valuable information source** to support agricultural monitoring systems at medium and large scale:

- Support **public policy**
- Map **natural resources**

Among all the EO data, Satellite Image Time Series offer new opportunities to monitor the Earth Surface evolution and provide insights in agricultural productions.

In the context of EO data, **Machine Learning/DL tools** seem adequate to get the most of EO data but:

- It is mainly data-driven (some efforts are starting to combine data-driven and knowledge-based approaches).
- ML/DL is not yet fully consolidated in the contest of EO analysis and further research is still necessary.
- Necessity to extract additional information that can support the model decision (explainability).

Perspectives

In the context of Object-based analysis, **combine** TASSEL and STARCANE principles.

Extend approach to leverage **heterogeneous EO sources** (Sentinel-2, Sentinel-1, Very High Spatial Resolution, etc..).

Towards **limited reference data** to train the model.

Spatial and Temporal model transfer: from an area to another area, from a time period to another time period.

Combine EO data with insitu (or proxy detection) data to combine information at extreme different scales.

Thank You for your attention

INRAE

TERRITOIRE ENVIRONNEMENT TELEDETECTION INFORMATION SPATIALE
tetis

#DigitAg

Thank You for your attention

INRAE

TERRITOIRE ENVIRONNEMENT TELEDETECTION INFORMATION SPATIALE

tetis

#DigitAg

Questions