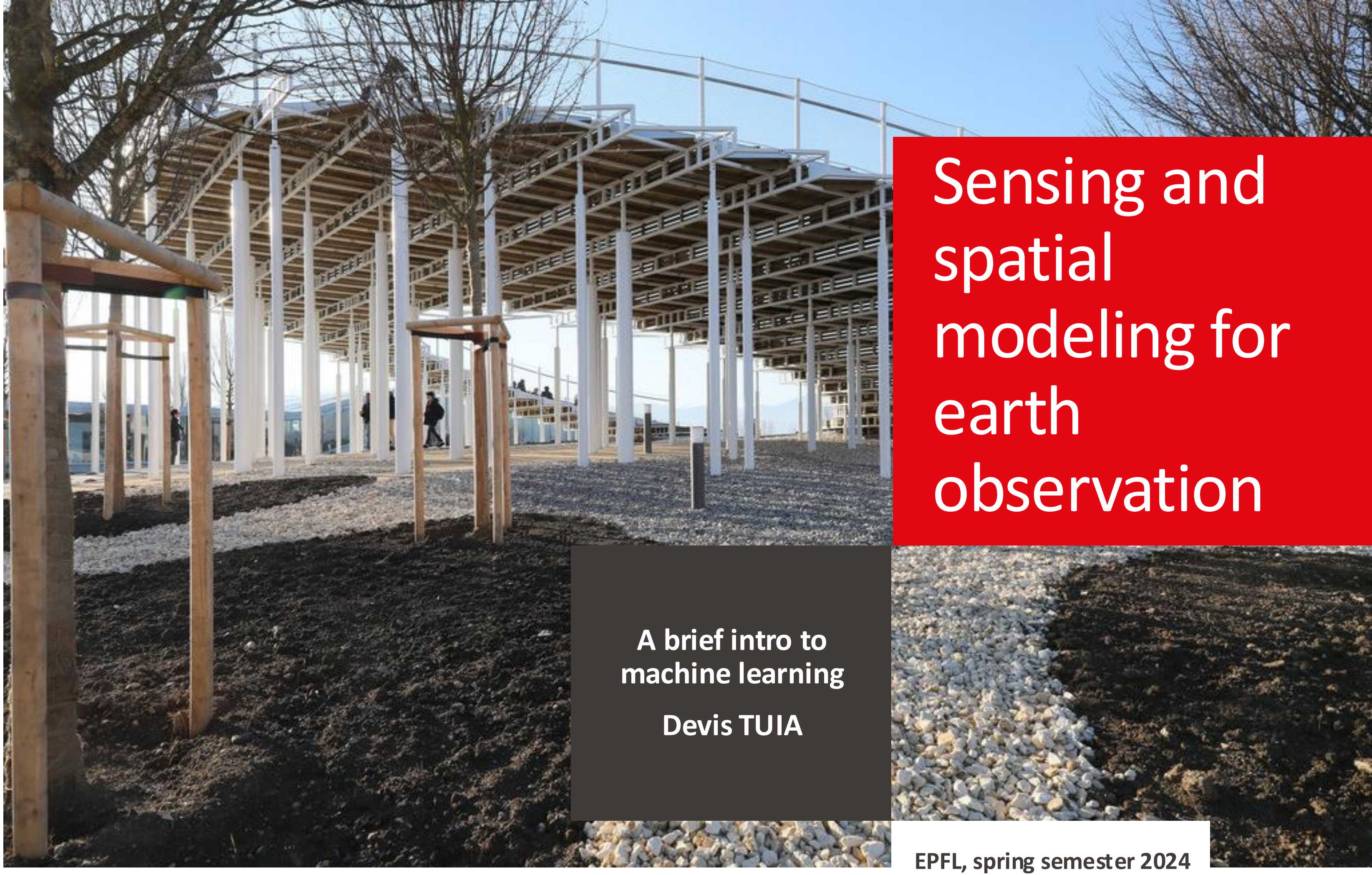
EPFL



# Sensing and spatial modeling for earth observation

A brief intro to machine learning

Devis TUIA

EPFL, spring semester 2024

École polytechnique fédérale de Lausanne

# So far in ENV-440…

- Image acqusition and formation
- Keypoint detection and matching
- Image orientation
- Creating orthos
- Extracting some features from elevation models

Now let's move into process prediction.

# Process prediction





- We want to estimate a quantity of interest given some input data.
- In our case, the input data will be
  - the DEM
  - the variables of interest we extracted from it (see last course)
- We will consider methods from two families
  - Machine learning (this and next course)
  - Geostatistics (following courses by Alexis Berne)


# Machine learning

What is ML for you?

# Let’s hear it from you!

# Machine learning?



- Let's stick to the basics: ML has been defined as "learning from data".

- And I am pretty sure you've done some so far.

You maybe did not know it

(but probably you did)

# Machine learning?

- Let’s stick to the basics: ML has been defined as “learning from data”.
- Learning what?

| All kinds of desirable outputs | Given some relevant variables (inputs) |
|---|---|
| - What price will my computer have in 3 years ? | *Color, brand, processor, RAM, …* |
| - Will it rain tomorrow ? | *Temperature, pressure, 10years of rain measurements, …* |
| - What animal is in this picture ? | *Pictures of animals, colors, …* |
| - What is the shortest route to go to Chailly ? | *A street network, cell phone data, TCS/TL websites, …* |

# Machine learning?

- Let's stick to the basics: ML has been defined as "learning from data".
- Learning what?

| Our kinds of desirable outputs | Given some relevant variables (inputs) |
|---|---|
| - What is the runoff of the river `there'?<br>- What is the chlorophyll in this plant?<br>- Which portions of that region have been flooded ?<br>- What was the bird doing during its flight ?<br>- What is the content (tag) of this image ? | *YOU name them!* |

# Machine learning?

- Let's stick to the basics: ML has been defined as "learning from data".

- **Learning how?**
- With generic algorithms

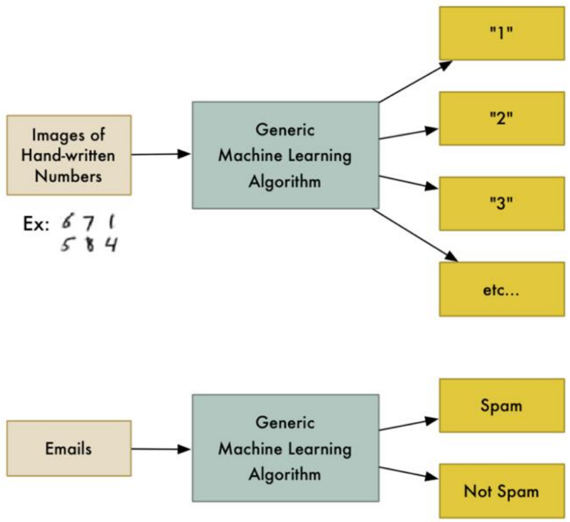
- We don't want to write specific code
- We want to feed data to the generic algorithm
- We leave the algorithm build its own logic linking inputs and the output

(…. and then improve it with specific knowledge)

# Machine learning?

- Let’s stick to the basics: ML has been defined as “learning from data”.

- **Learning how?**
- With generic algorithms




Source: https://medium.com/@ageitgey/machine-learning-is-fun-80ea3ec3c471#.7x1l6rhtr

# Machine learning?





- Let’s stick to the basics: ML has been defined as “learning from data”.
- **In what is it different to physical models?**

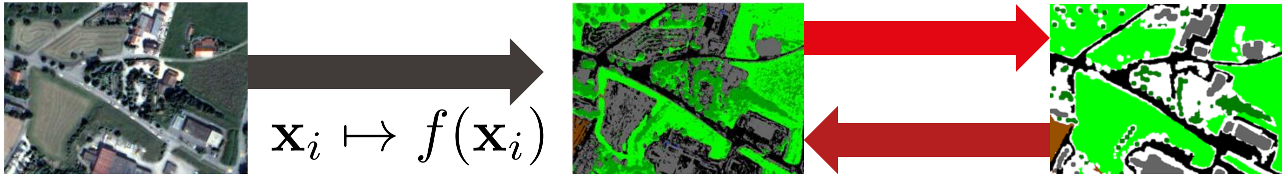
| | Physical | Machine learning |
|---|---|---|
| Ground | Physics<br>= 1 problem, 1 model | Learn from data<br>= many problems, 1 model |
| Training | No need for training | Need to be trained |
| Comp. | Generally slow | Very fast |
| General | Depends strongly on boundary conditions | Cannot learn other things than those seen in training |

# So how does it work?

- ML algorihtms have generally two phases
  - training phase: the model learng the input/output relations from a set of known data
  - inference (or testing phase): the trained model predicts unseen data using the knowledge gathered during training.

# So how does it work?

- Machine learning: statistical models learning patterns from observational data



$$\mathbf{x}_i \mapsto f(\mathbf{x}_i)$$

$$\hat{y}_i = f(\mathbf{x}_i) \qquad y_i$$

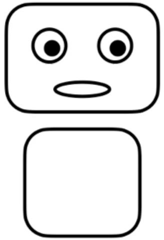
# So how does it work?

- ML algorihtms have generally two phases
  - training phase: the model learng the input/output relations from a set of known data
  - inference (or testing phase): the trained model predicts unseen data using the knowledge gathered during training.

# In a nutshell







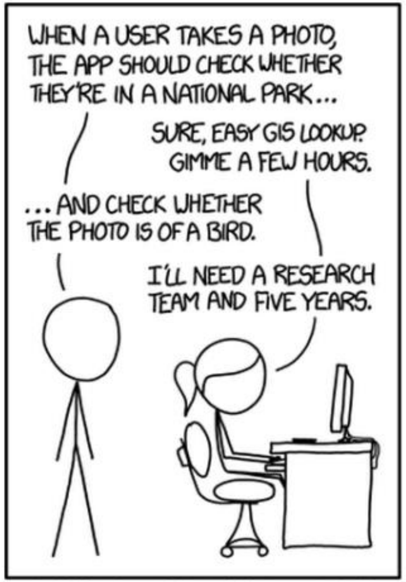


Animation by B. Kellenberger, 2021 © EPFL ECEO

# Machine learning?

- Does it work all the time?
- No. it is not magic.

- It works if we have
- the right inputs
- the right learning machine
- sufficient training data



IN CS, IT CAN BE HARD TO EXPLAIN THE DIFFERENCE BETWEEN THE EASY AND THE VIRTUALLY IMPOSSIBLE.

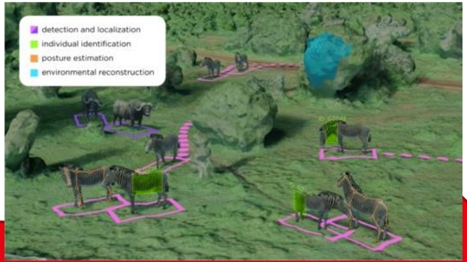
Source xkcd

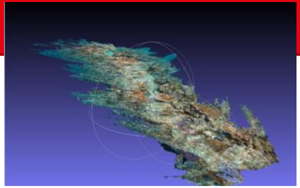
# Why now : statistical and computational models are good enough...

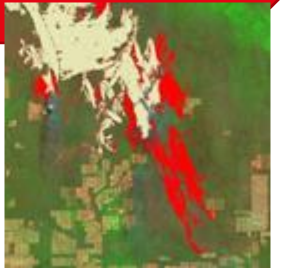
- Machine learning has reached a certain maturity... and percolated in many fields of science.

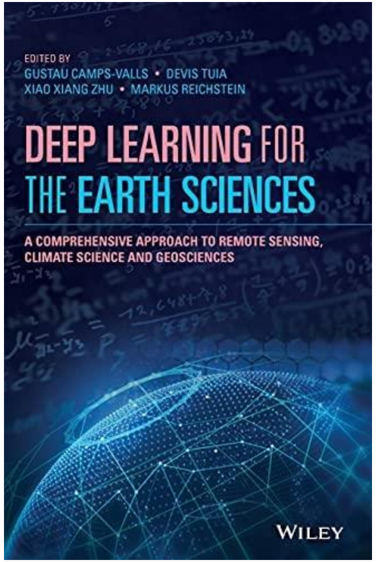
2022










2015

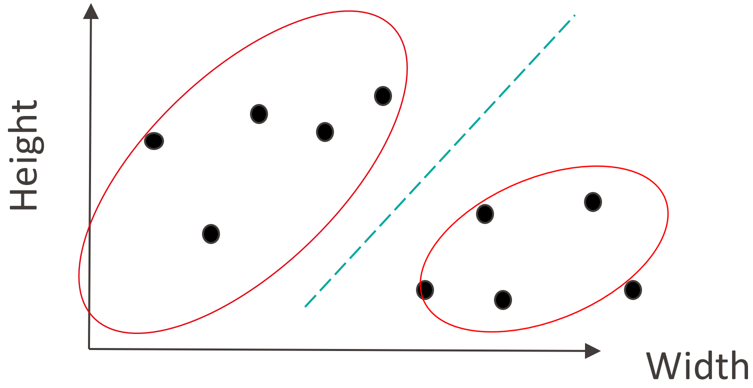

# A taxonomy of ML algorithms with selected examples

# A taxonomy of machine learning methods





- Unsupervised learning
  - Clustering
  - Structure



# Unsupervised learning - clustering

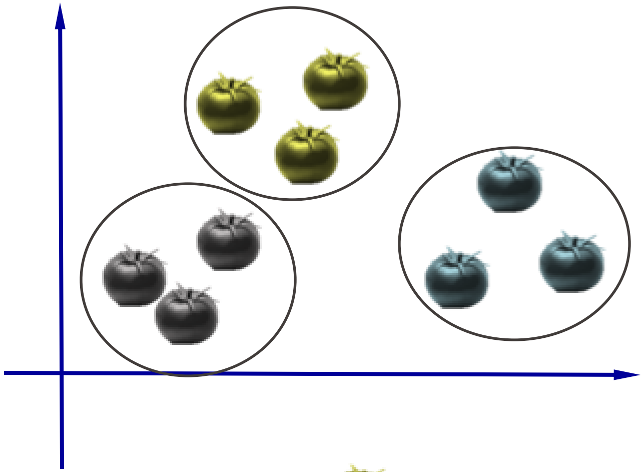




- Unsupervised learning
  - Learning the relationship $f$ between inputs $\mathrm{X}$ based on training data $\{(x_1\ ), \ldots (x_n\ )\}$
  - Input $\mathrm{X}$ is called predictor, independent variable, feature
  - There is no output to predict, you learn the structure of the data.
  - For example, when grouping data, one predicts a cluster
- A cluster is the ID of a group of data points that are similar to each other

# A clustering example

## Clustering



Cluster 1

Cluster 2

Cluster 3

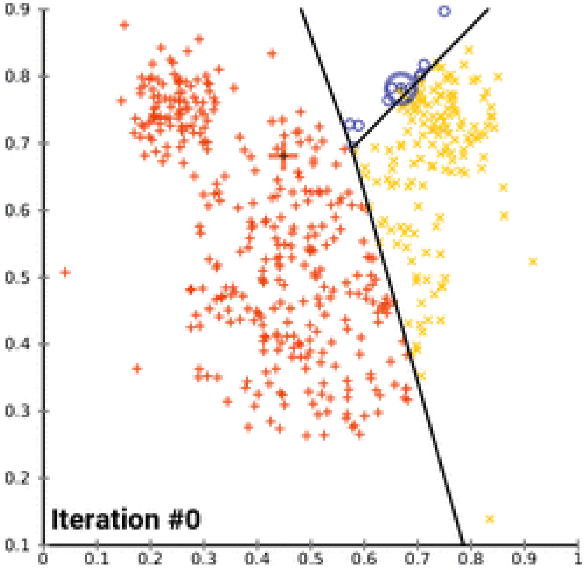
# Example of unsupervised method: k-means

*K*-means is an iterative method

1. It starts with a guess of the cluster centers (often random)
2. Computes the assignment by minimizing variance of each cluster
3. Updates the centers as the means of the samples assigned to each center
4. Repeats 2.-3. until stability is reached.

Stability can be:

- a fixed number of iterations
- when the centroids do not move anymore



# Clusters to classes

- Once we have a decent clustering result, where each cluster can be linked to a single class
- We proceed to a manual class assignation
  - Cluster 1 → Grass
  - Cluster 2 → water
  - Cluster 3 → Grass
  - Cluster 4 → Built
  - …

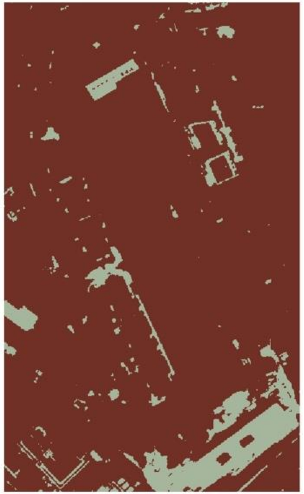
Note that **many clusters can be assigned to the same class!**

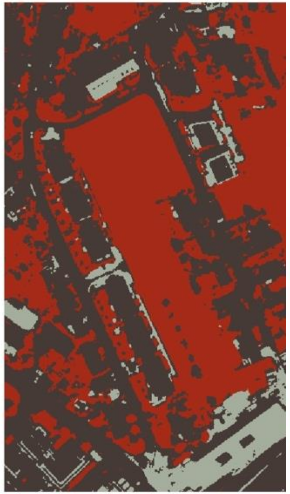
and it does happen!!!

# Example on a very high resolution image

## Quantization

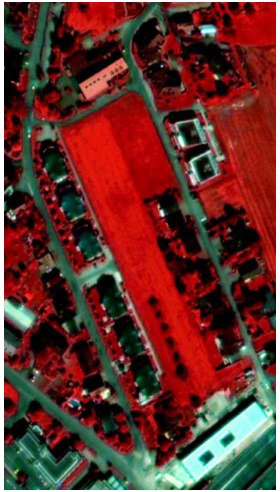
(each pixel is colored by the average color of the members of the cluster it belongs to)


2 clusters


3 clusters


4 clusters


8 clusters

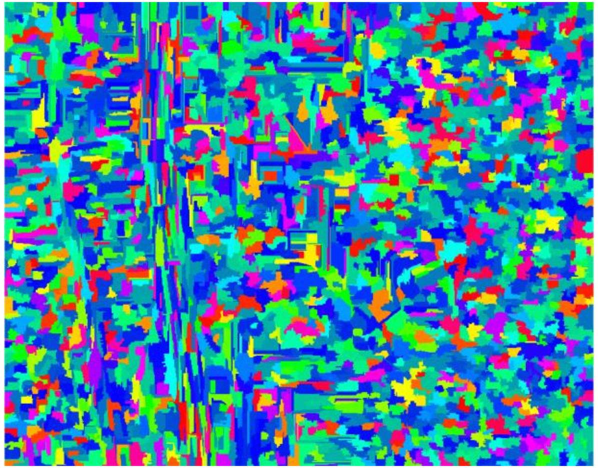

Original image

# Example on a very high resolution image

## Superpixelisation

(each pixel is grouped to neighbors according to their spectral similarity)

True image



Clustering result
500 clusters
(colors are cluster IDs)

VS

Semantic information

# Unsupervised learning – dimensionality reduction

- Compress information based on data redundancy
- Without any prediction objective, you aim at removing un-necessary, redundant information
- Typically one aims at finding correlated variables and remove them

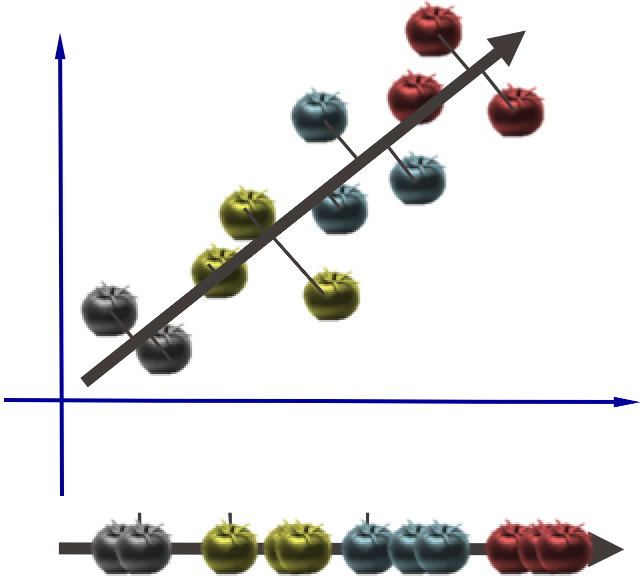
- A famous method is Principal Component Analysis (PCA)

# A dimensionality-reduction example



## Structure

From 2 correlated variables



To a single uncorrelated one
(with same information)

Dimensionality reduction

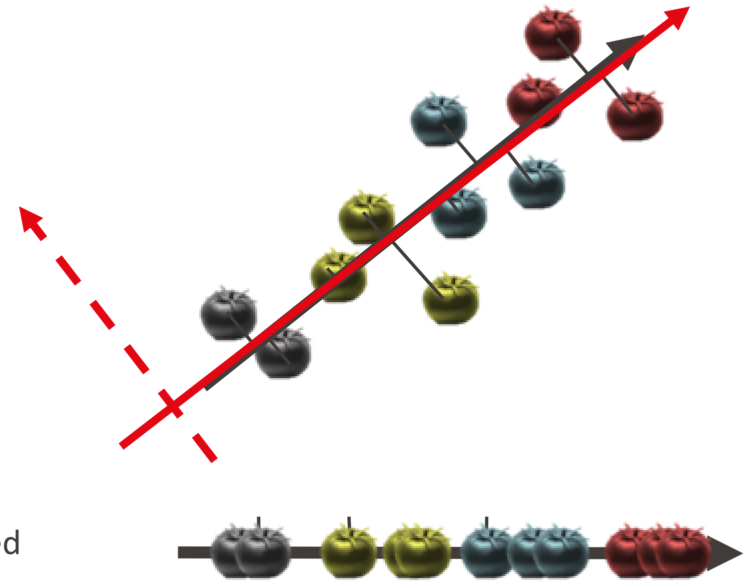
# A dimensionality-reduction example



## Structure

From 2 correlated variables

With a rotation of the axis



To a single uncorrelated
Principal component
(with same information)

Dimensionality reduction

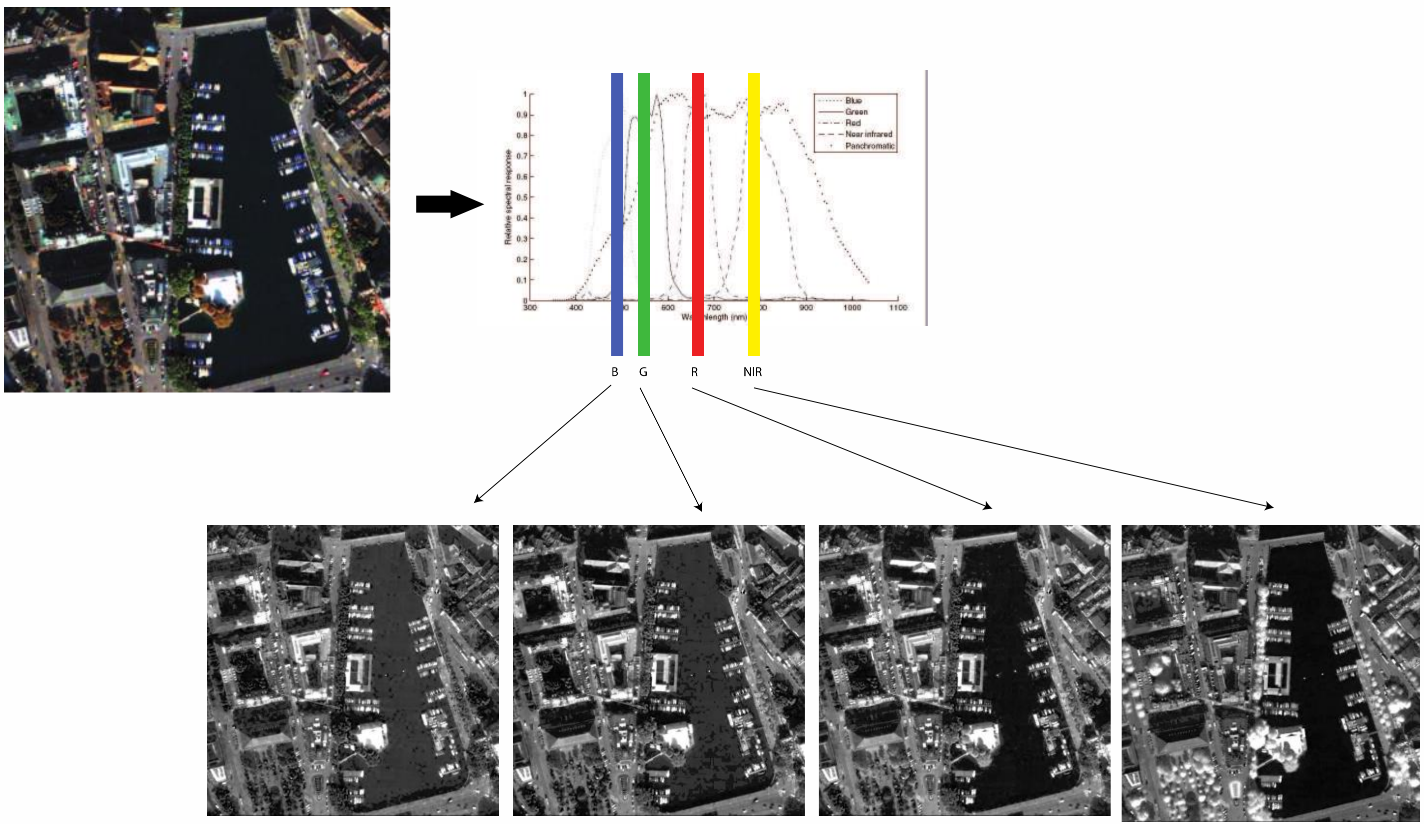
# Example of PCA on satellite data: multi-resolution image fusion



Multispectral image
Low res
color

Panchromatic image
High res
BW

# Example of PCA on satellite data






D. Tuia. ECEO

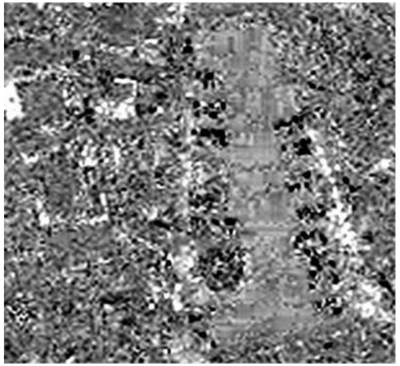
# Example of PCA on satellite data



PC1, the information correlated across all bands

PC2, information specific to single bands

PC3, very high frequencies



PC4, what’s left

# Projection matrix (from last course)

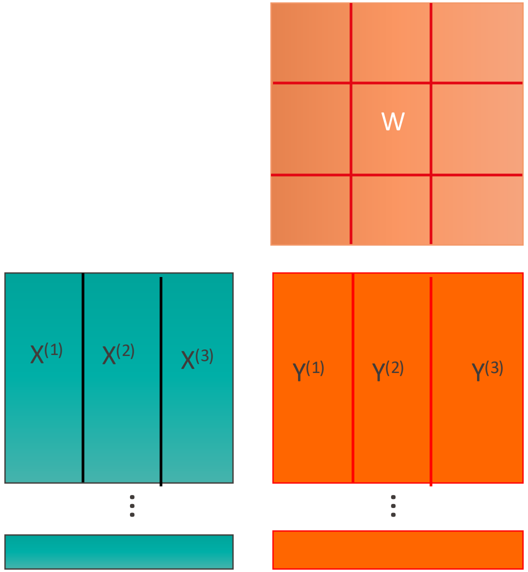
$$Y = XW$$

- W is a square matrix
- As many lines and columns as bands
- Ex: 3 bands

$$W = \begin{bmatrix} W(1,1) & W(1,2) & W(1,3) \\ W(2,1) & W(2,2) & W(2,3) \\ W(3,1) & W(3,2) & W(3,3) \end{bmatrix}$$

Multiplies band 2

Projects into component 1

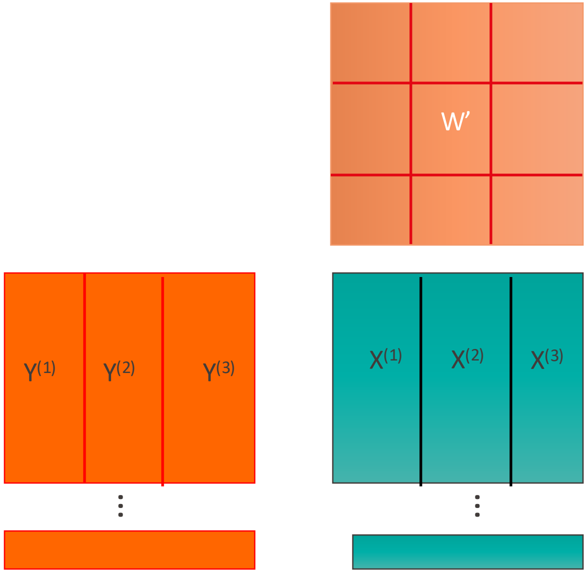


# Projection matrix (from last course)

- W is an invertible matrix!
- Since it is orthogonal, $W^{-1} = W'$
- Given the principal components, we can go back to the original bands exactly.
- Now… what if we change the first PC a little bit?

$$Y = XW$$

$$\boxed{X = YW^{-1} = YW'}$$



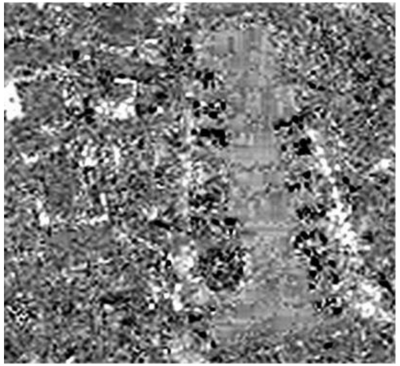
# Example of PCA on satellite data

D. Tuia. ECEO

PC1, the information correlated across all bands

PC2, information specific to single bands

PC3, very high frequencies



PC4, what's left

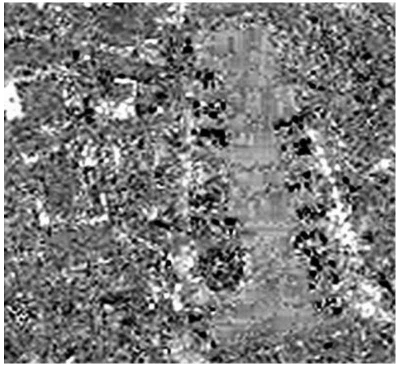
# Example of PCA on satellite data



Replace with
High resolution
Grayscale band

PC2

PC3



PC4

# Feature fusion

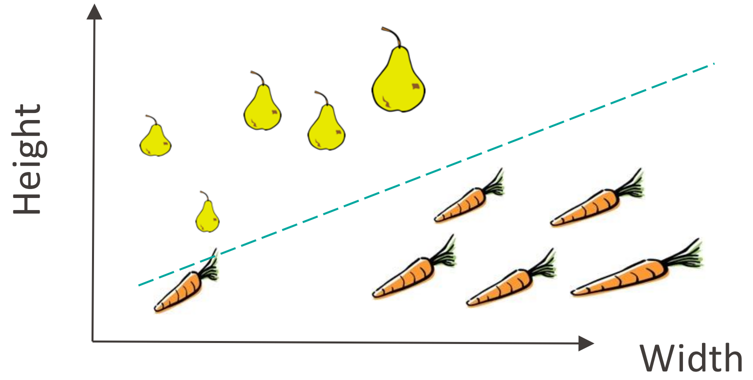
High res
BW

Low res
color

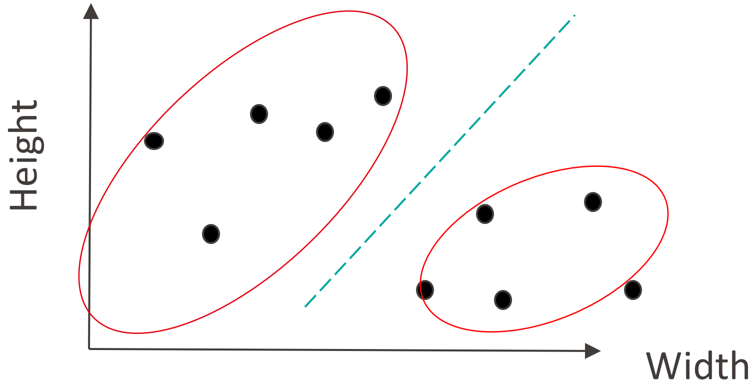
High res
color!

# A taxonomy of machine learning methods

- Supervised learning
  - Classification
  - Regression




- Unsupervised learning
  - Clustering
  - Structure



# Supervised learning

- Supervised learning
  - Learning the relationship $f$ between input $\mathrm{X}$ and output $\mathrm{Y}$ based on training data $\{(x_1, y_1), \ldots (x_n, y_n)\}$
  - $\mathrm{Y} = f(\mathrm{X})$

  - Input $\mathrm{X}$ is called predictor, independent variable, feature
  - Output $\mathrm{Y}$ is called response, dependent variable

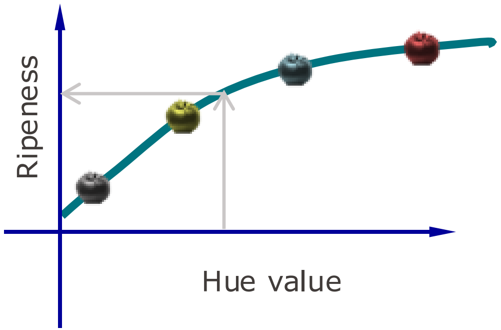
- Two types of outputs: <u>classification</u> and <u>regression</u>

# Regression: continuous variables

New image

Ripeness: 6.2

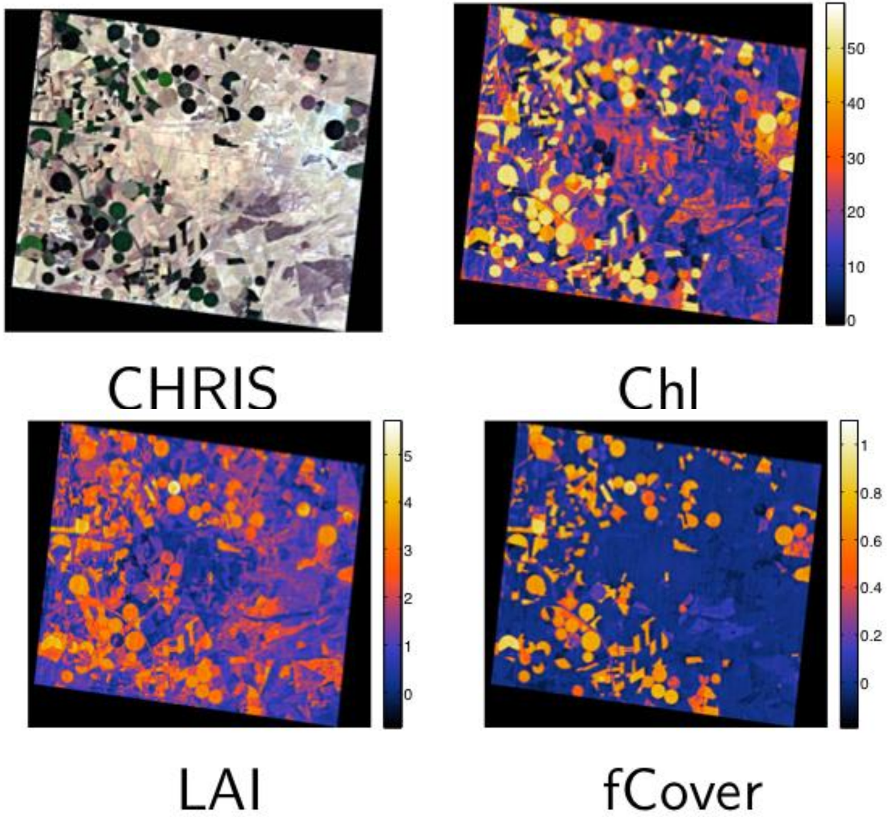
Regression model




Ripeness: 3.2 5.8 8.1 11.3

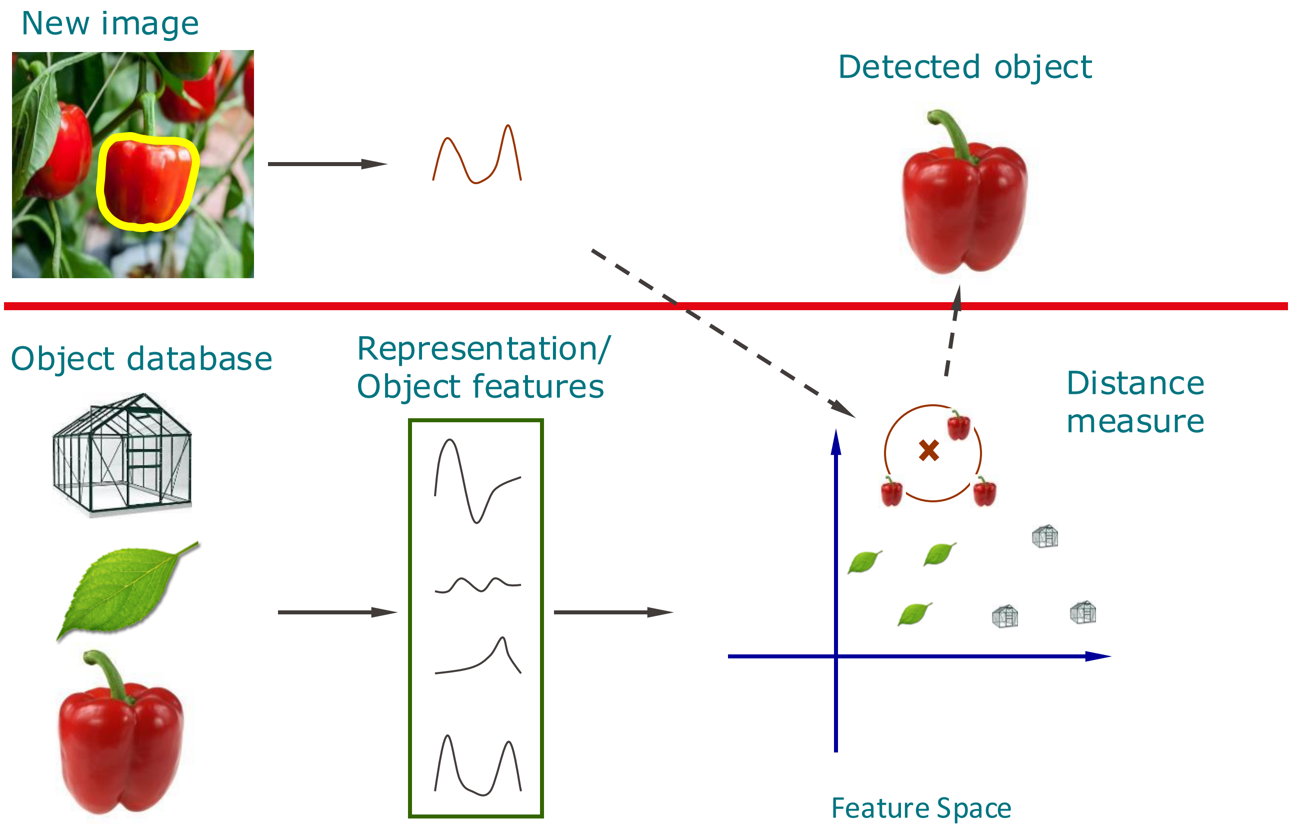
# Regression with remote sensing data

More about regression later in this course!

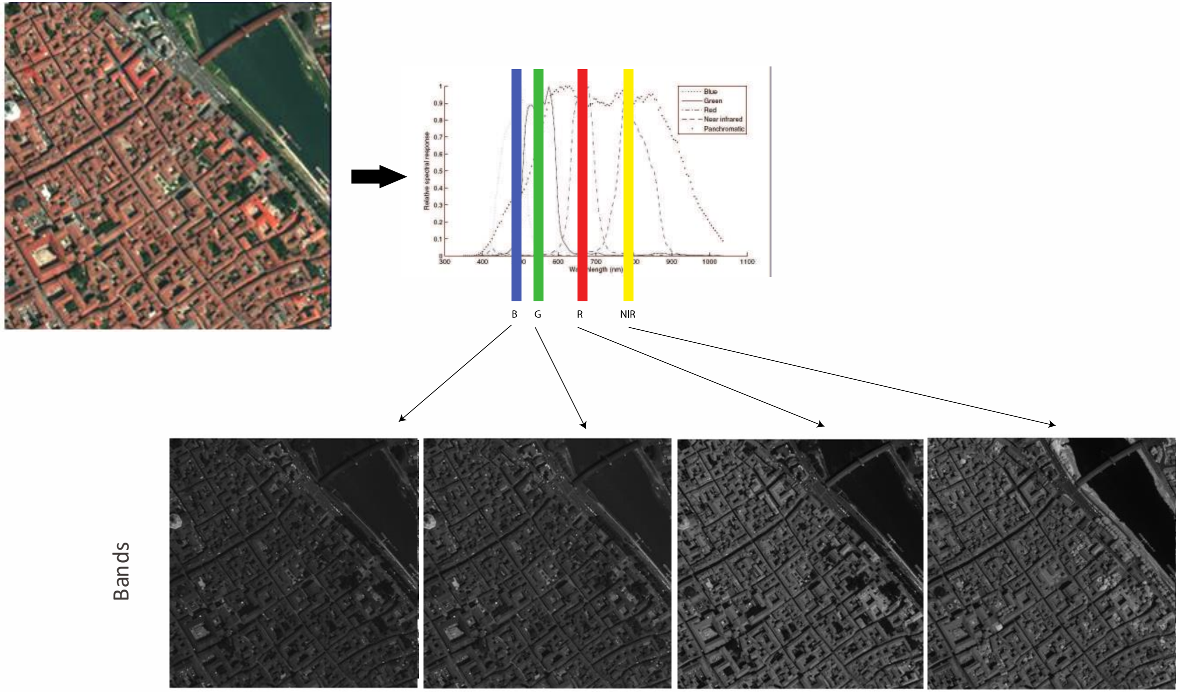



[Tuia et al., 2011, GRSL]

# Classification: discrete (classes)



# Classifying remote sensing data
# 1 – digital information



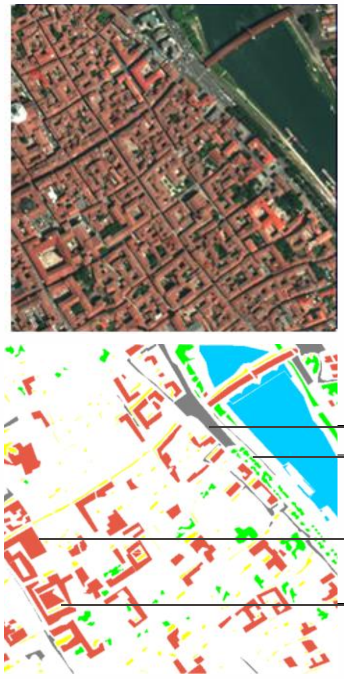
# Classifying remote sensing data
# 2 – labeled examples

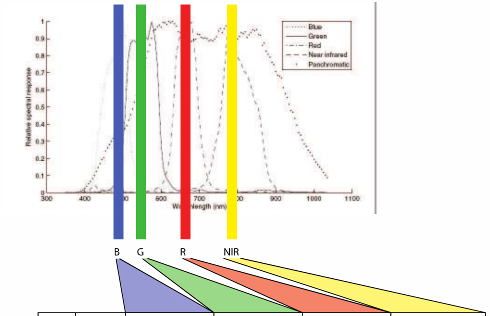


input-output pairs

Bands

class






| | Class | Blue | Green | Red | NIR |
|---|---|---|---|---|---|
| $x_1$ | R | $x_{1,B}$ | $x_{1,G}$ | $x_{1,R}$ | $x_{1,NIR}$ |
| $x_2$ | T | $x_{2,B}$ | $x_{2,G}$ | $x_{2,R}$ | $x_{2,NIR}$ |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| $x_3$ | B | $x_{3,B}$ | $x_{3,G}$ | $x_{3,R}$ | $x_{3,NIR}$ |
| ⋮ | · | ⋮ | ⋮ | ⋮ | ⋮ |
| $x_4$ | B | $x_{4,B}$ | $x_{4,G}$ | $x_{4,R}$ | $x_{4,NIR}$ |

Labeled information

EPFL

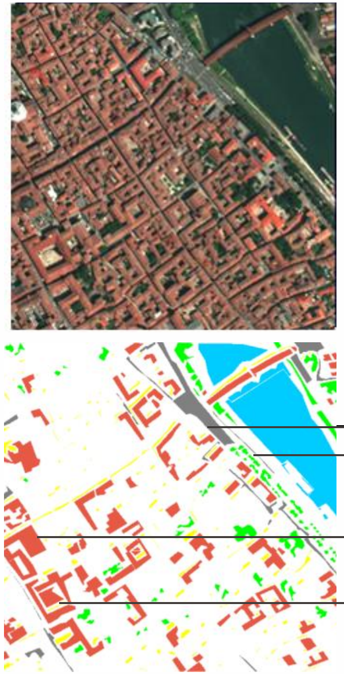
# Classifying remote sensing data
# 3 – build the model



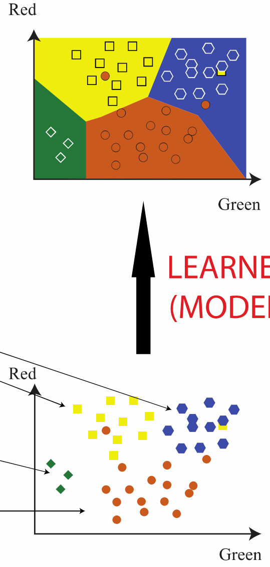


Bands

class



| | Class | Blue | Green | Red | NIR |
|---|---|---|---|---|---|
| $x_1$ | R | $x_{1,B}$ | $x_{1,G}$ | $x_{1,R}$ | $x_{1,NIR}$ |
| $x_2$ | T | $x_{2,B}$ | $x_{2,G}$ | $x_{2,R}$ | $x_{2,NIR}$ |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| $x_3$ | B | $x_{3,B}$ | $x_{3,G}$ | $x_{3,R}$ | $x_{3,NIR}$ |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| $x_4$ | B | $x_{4,B}$ | $x_{4,G}$ | $x_{4,R}$ | $x_{4,NIR}$ |

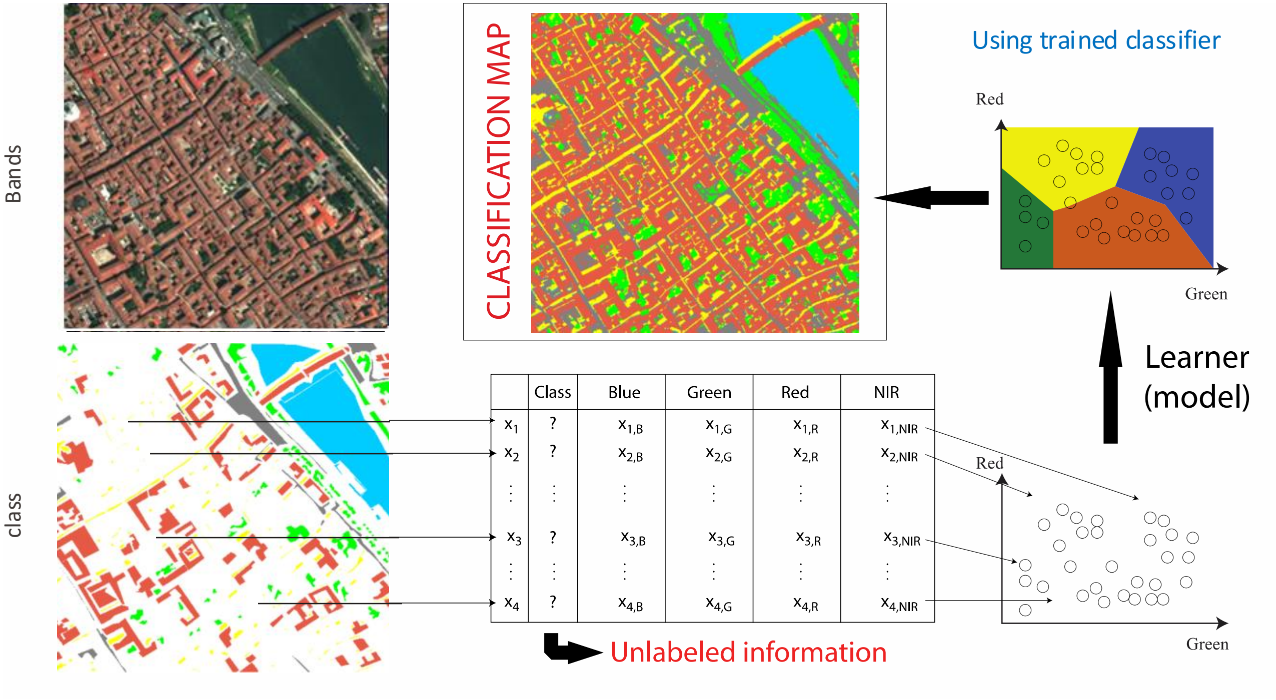



LEARNER (MODEL)

Training classifier

# Classifying remote sensing data
# 4 – predict unseen data






| | Class | Blue | Green | Red | NIR |
|---|---|---|---|---|---|
| $x_1$ | ? | $x_{1,B}$ | $x_{1,G}$ | $x_{1,R}$ | $x_{1,NIR}$ |
| $x_2$ | ? | $x_{2,B}$ | $x_{2,G}$ | $x_{2,R}$ | $x_{2,NIR}$ |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| $x_3$ | ? | $x_{3,B}$ | $x_{3,G}$ | $x_{3,R}$ | $x_{3,NIR}$ |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| $x_4$ | ? | $x_{4,B}$ | $x_{4,G}$ | $x_{4,R}$ | $x_{4,NIR}$ |

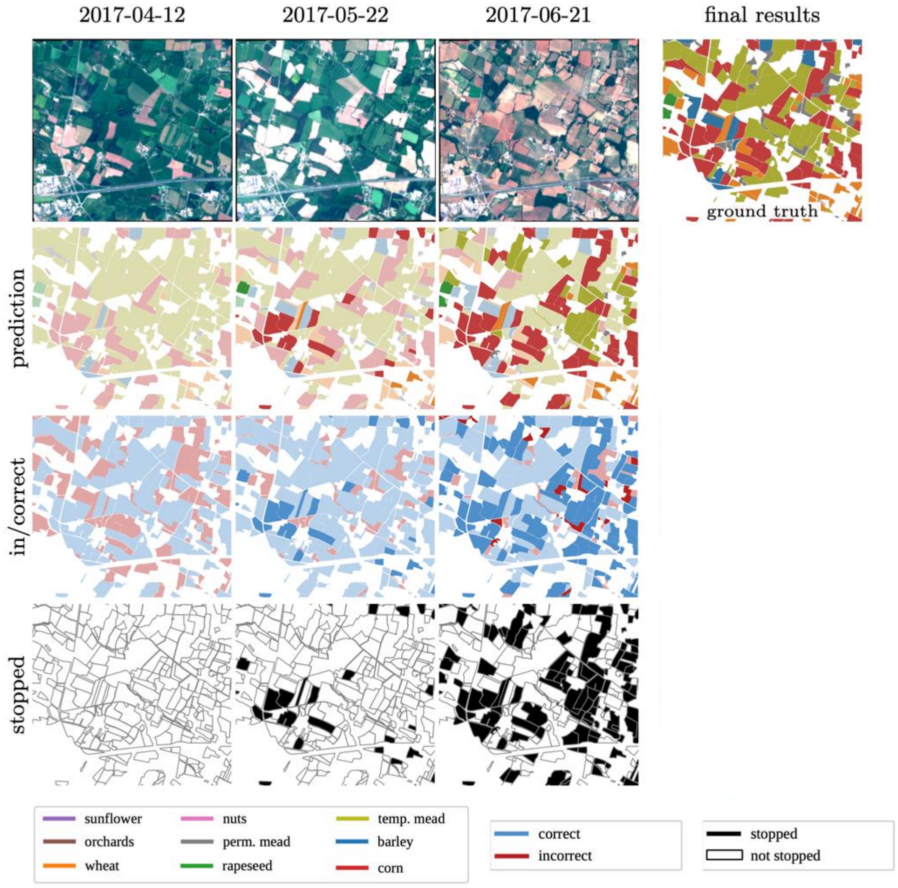
Classifying satellite images

- Here we predict the crop being cultivated using a time series of sat data.


M. Russwurm, Courty C., R. Emonet, S. Lefèvre, D. Tuia, and R. Tavenard. ELECTS: End- to-end learned early classification of time series. *ISPRS J. Int. Soc. Photo. Remote Sens.*, 196:445–456, 2023.

# Classifying satellite images

- Here we predict the crop being cultivated using a time series of sat data.
- The prediction get more accurate with more temporal information



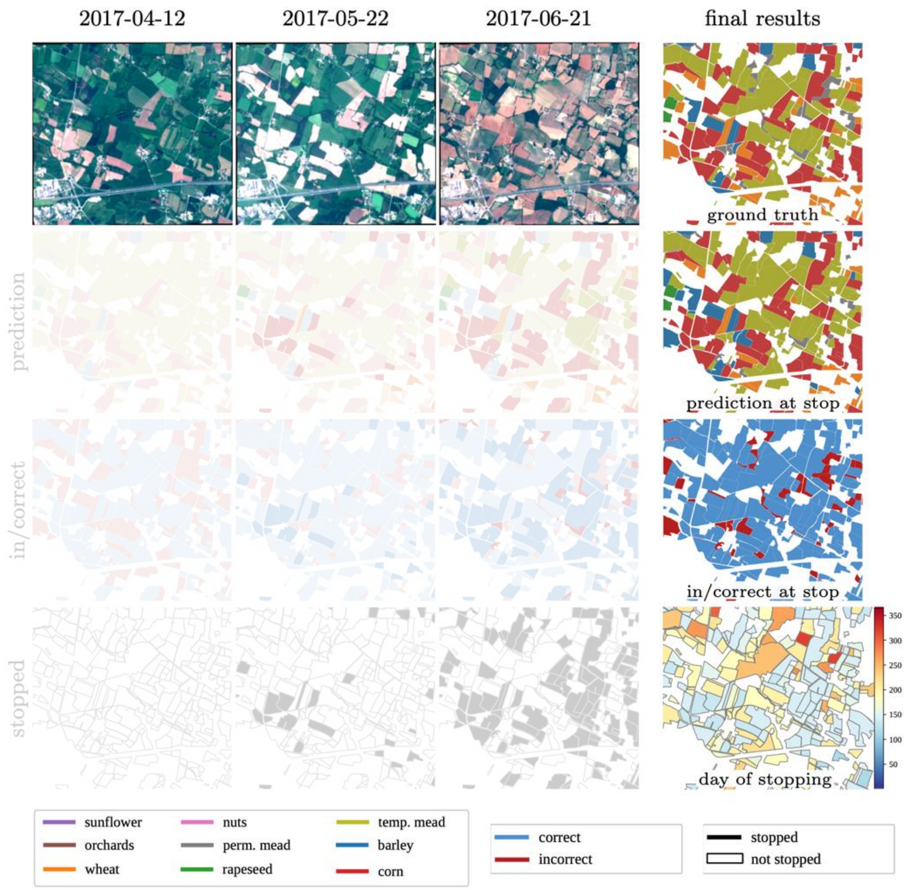

M. Russwurm, Courty C., R. Emonet, S. Lefèvre, D. Tuia, and R. Tavenard. ELECTS: End- to-end learned early classification of time series. *ISPRS J. Int. Soc. Photo. Remote Sens.*, 196:445–456, 2023.

# Classifying satellite images





- Here we predict the crop being cultivated using a time series of sat data.
- The prediction get more accurate with more temporal information




M. Russwurm, Courty C., R. Emonet, S. Lefèvre, D. Tuia, and R. Tavenard. ELECTS: End- to-end learned early classification of time series. *ISPRS J. Int. Soc. Photo. Remote Sens.*, 196:445–456, 2023.

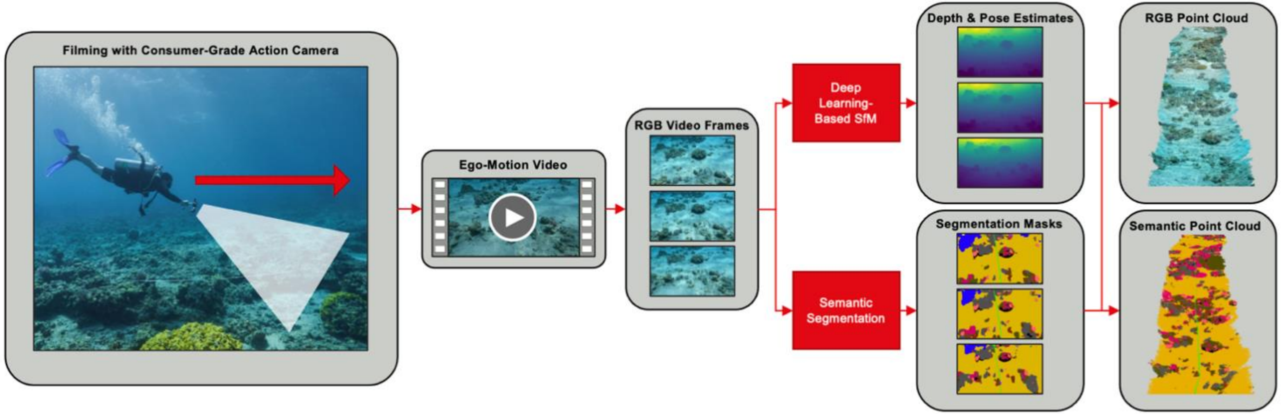
EPFL

Transnational Red Sea Center
Bridging Science & Diplomacy for the Future of Corals

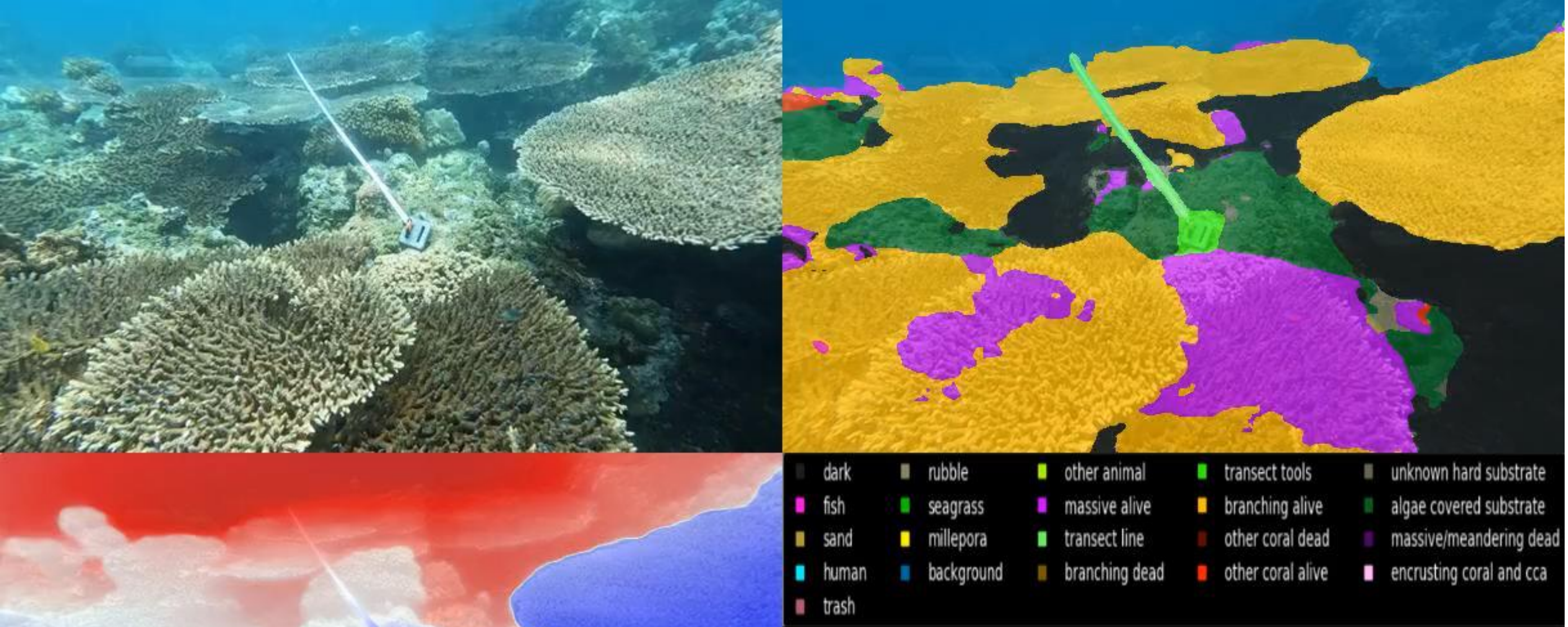
# Classify other types of images.

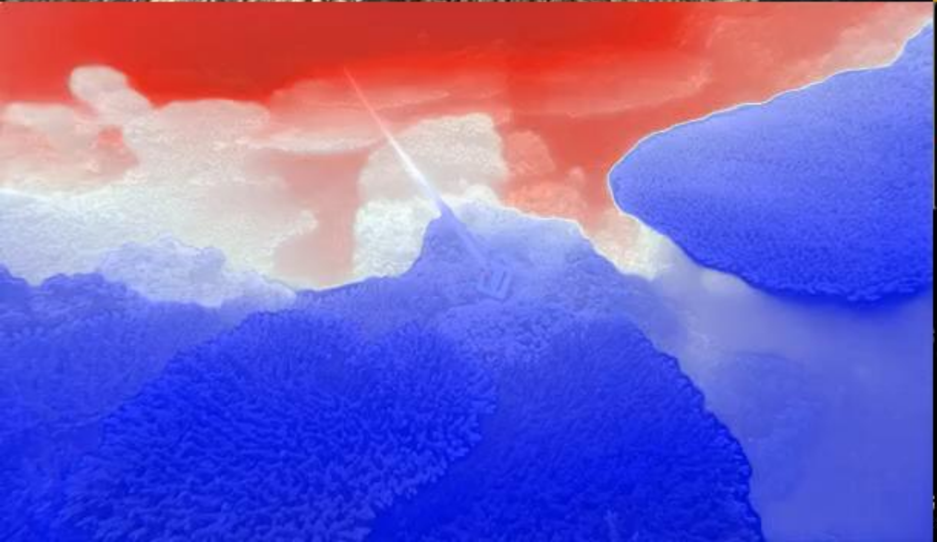
- With custom-built, affordable imaging setup
- A model that works on videos, leveraging 2 tasks

- Once trained, 3D reconstructs a 100m transect in playing video time
- Tested in Isreal, Jordan and Djibouti in 2022




J. Sauder, G. Banc-Praudi, A. Meibom, and D. Tuia. Scalable semantic 3d mapping of coral reefs with deep learning. *Methods in Ecology and Evolution,* 2024.


dark
rubble
other animal
transect tools
unknown hard substrate
fish
seagrass
massive alive
branching alive
algae covered substrate
sand
millepora
transect line
other coral dead
massive/meandering dead
human
background
branching dead
other coral alive
encrusting coral and cca
trash



# In the next course…

- We will consider regression problems

- In a regression problem, we estimate a process that is continuos, e.g.
  - Temperature (in °C)
  - Precipitations (in mm)
  - Chlorohpyll concentrations
  - …

- Discrete processes, or classification problems, will be covered in other courses (e.g. ENV-540, IPEO)