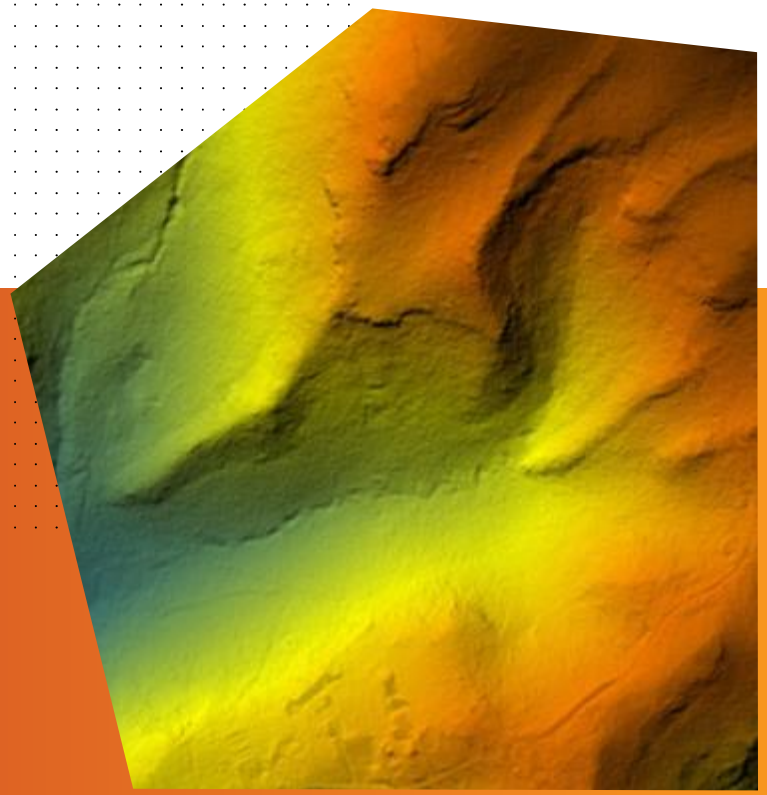
INSTITUT NATIONAL
DE L'INFORMATION
GÉOGRAPHIQUE
ET FORESTIÈRE



DEEP LEARNING EXPERIMENTATIONS

First steps in introducing deep learning into major IGN data production processes

Content

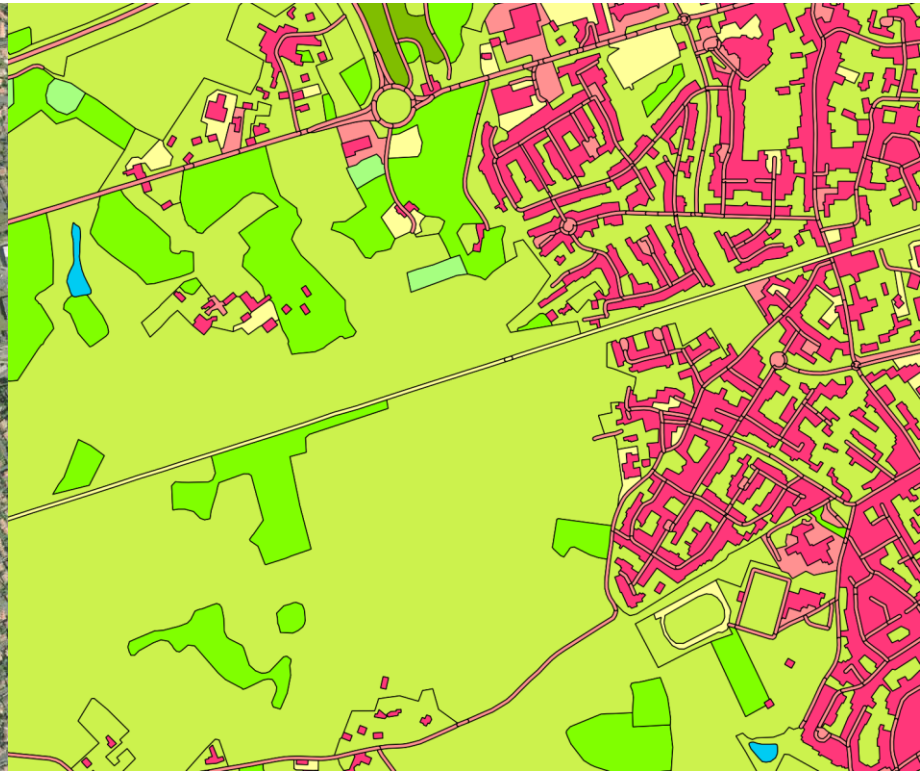
1. Illustrative Example
2. Deep Learning 101
3. Insights from a year of experiments
4. Zoom on timber : data generation
5. Summary and perspectives

IGN

Illustrative example

Land artificialisation monitoring based on OCS GE, a land cover/usage map

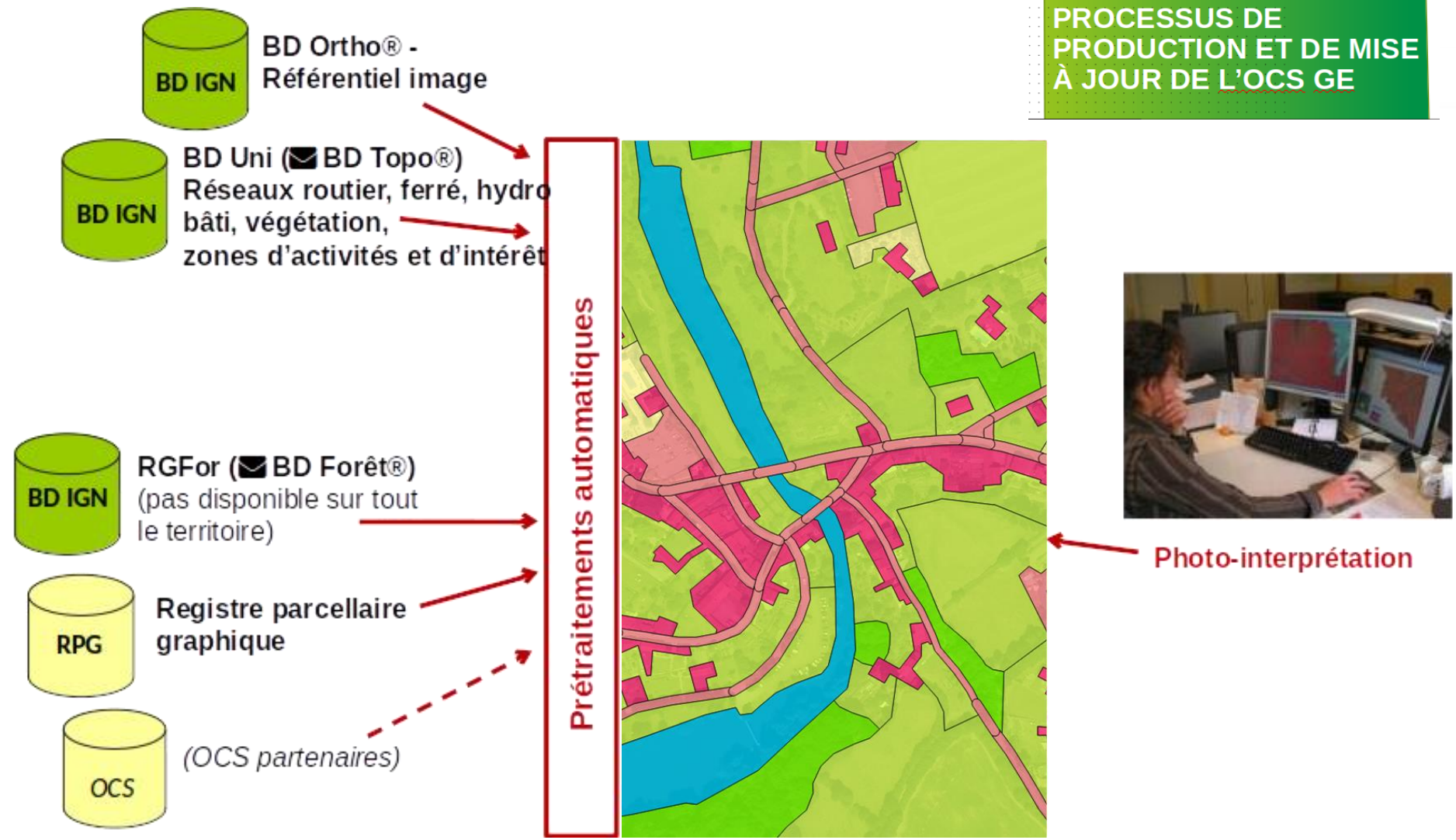
ORTHO HR®



OCS GE

Orthophotos (25cm/pix), Land Cover (14 fields)

ÉVOLUTION DU PROCESSUS DE PRODUCTION ET DE MISE À JOUR DE L'OCS GE



Current process

Current project using OCS GE

OCS GE Qualities / Defaults

- better definition than other land cover products
  - Corine Land Cover : 25 ha
  - OCS GE : 200 up to 2500 m²
- sufficiently detailed to build land artificialisation indicators
- updated every 3 years (French territory)

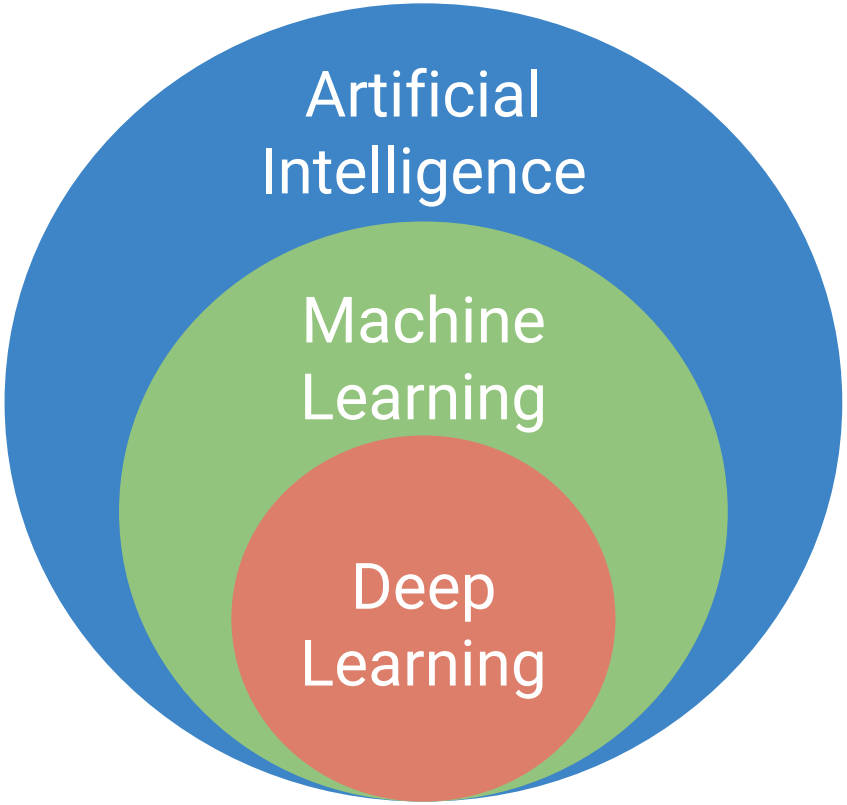
- complex to build
  - lots of external sources
  - lots of manual processes
  - change detection and updates processes are not yet fully defined

Expected improvements

- **minimizing** the number of external resources
- **limiting** manual work
- **shorter** chain of process
- computation launched **as soon as** the orthophotos are produced
- **no need to wait** for updated databases
- **stability** of the detection year after year

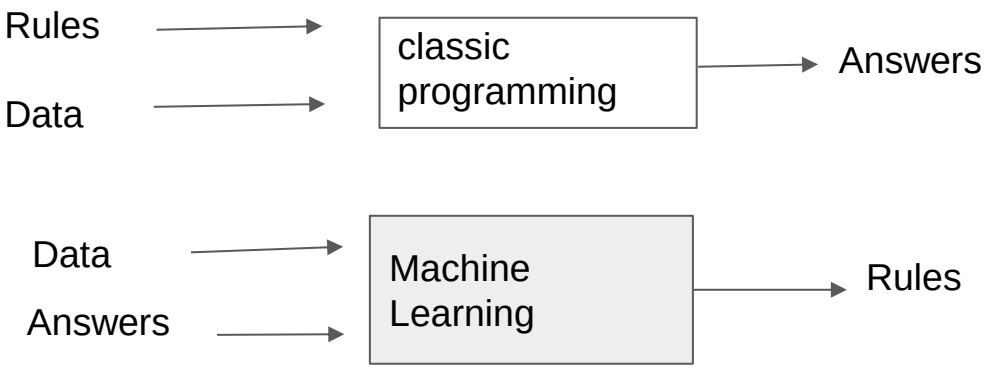
IGN

Deep Learning 101

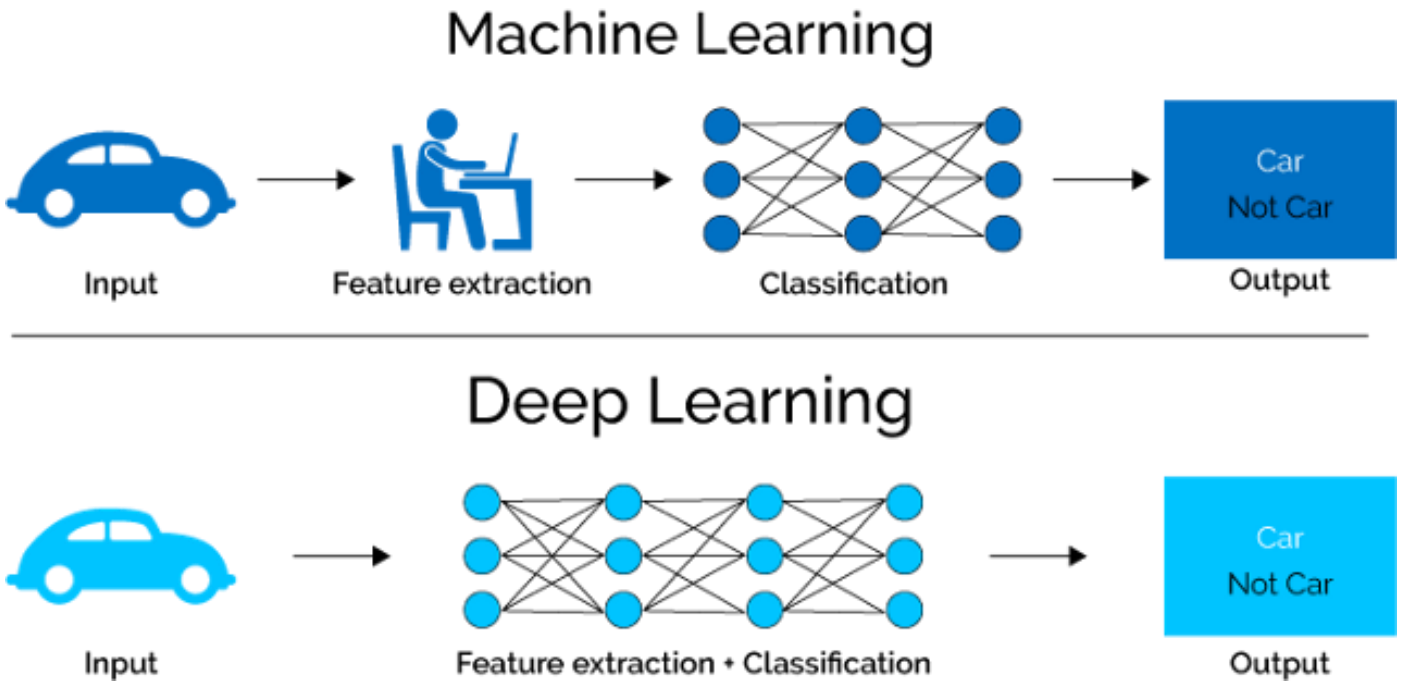


AI domains :

- planification
- natural language processing
- optimisation
- expert systems
- games
- ...

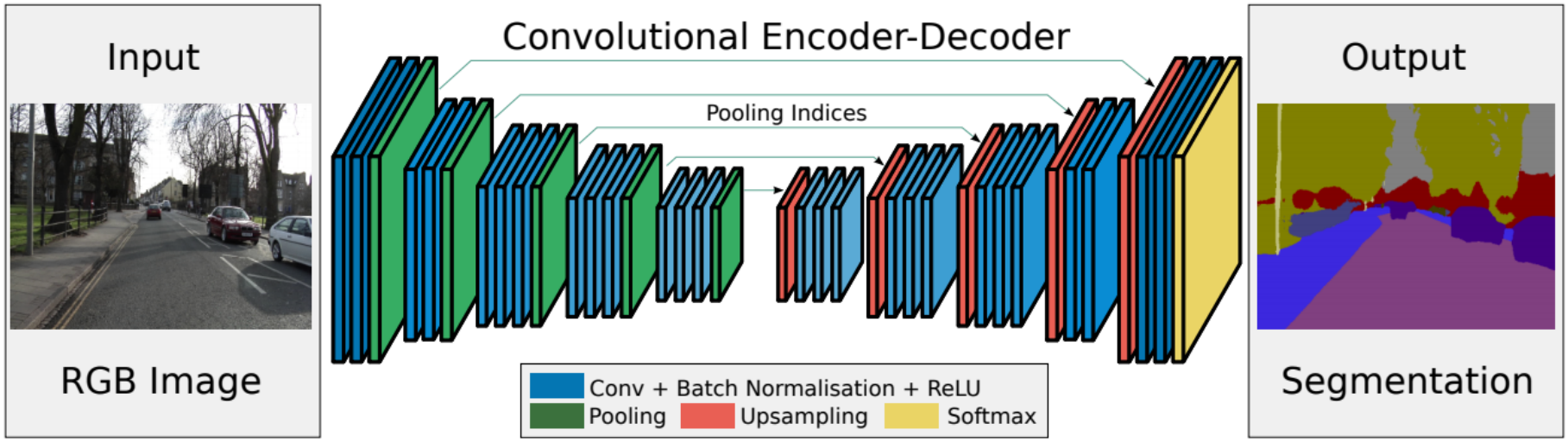


Machine Learning vs Deep Learning



- Less manual work
- Proven better than ML for image interpretation

At the core of deep learning : a model

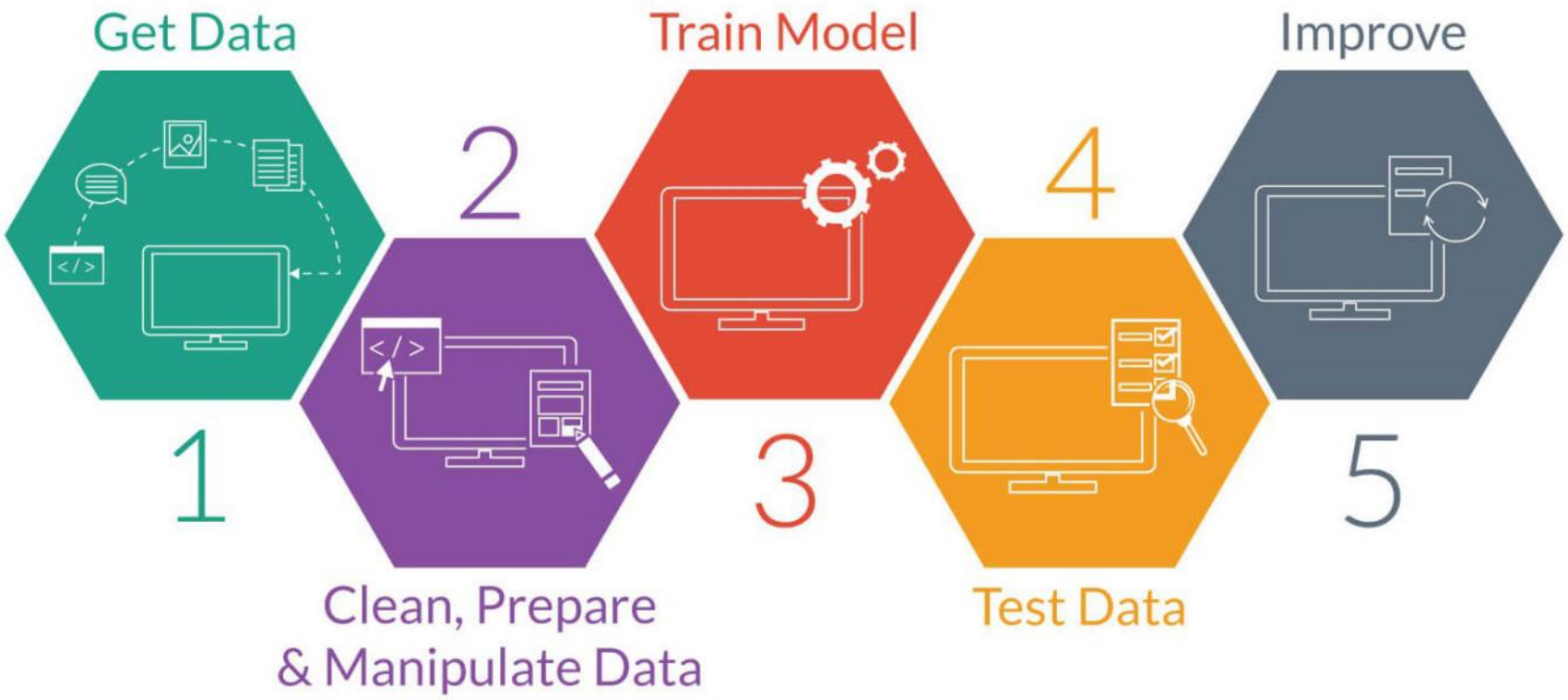


Model architectures :

- **Convolutional Neural Network** (image, video)
  - image classification
  - object detection
  - **semantic segmentation**
- Recurrent Neural Network (temporal data, text, sound)
- Generative Adversarial Network (image, text, sound)
- ...

IGN

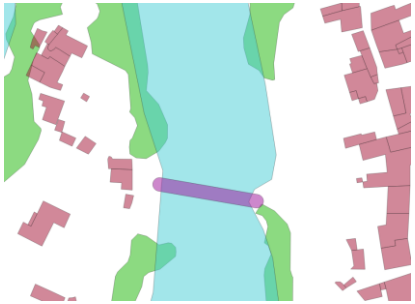
Deep Learning workflow (before deployment in production)

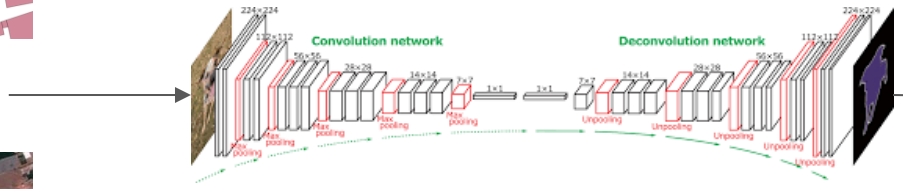


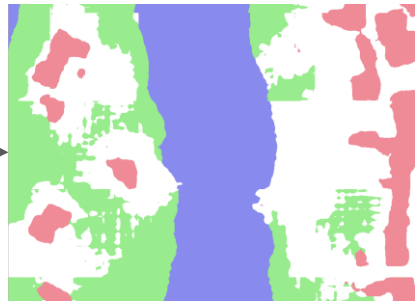
Once the models are stabilized and adequate
→ Integration into a production pipeline

IGN

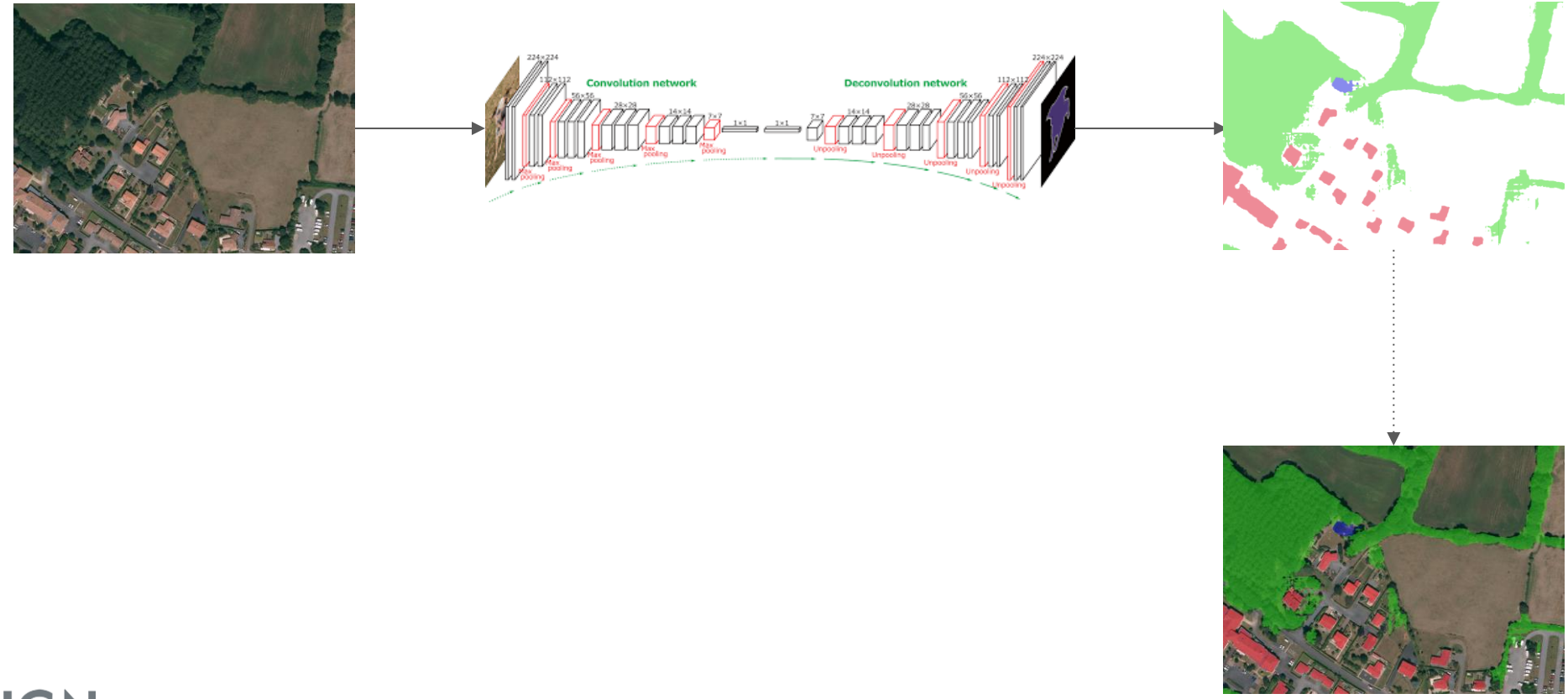
Deep Learning : learning phase







Deep Learning : detection phase



Insights from a year of experiments

Organization - timeline

Summer 2018 (1 month)

- Gathering of r**aw data** (aerial & satellite, vectorized data)
- **Infrastructure** setup
- **Data set** generation
- **Basic working codes** for several objectives (forest id, detection of change in buildings, building segmentation)

Hackaton (2 days)

- **Presentation**
- **Modification** of the proposed codes
- New and better datasets
- **Evaluation** (metrics, visualisation)
- **Improvement** proposals

Termos v1 (4 month)

- More **diverse datasets**
- New **models**
- Generation of **metrics** to rank the results
- **Bigger** territories investigated
- **Multiclass** detection
- ...

Termos v2 and beyond

- **Training** of additional personnel
- New **objects** of interest
- **Target : OCS GE**

Organization - workforce and infrastructure

Developers and data specialists

- Deep Learning
- Satellite data
- BDUni
- RGFor
- OCS GE
- ...

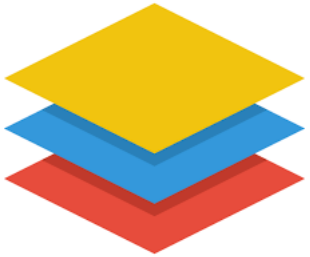
Computation power

- 6 upgraded computers with GPU (8 and 11 GB) and 32 GB of RAM
- 6 workstations for code development, data treatment, data visualisation, etc.

Lots of meetings with :

- researchers
- potential end users
- decision makers
- software developers
- ...

Mode of operation

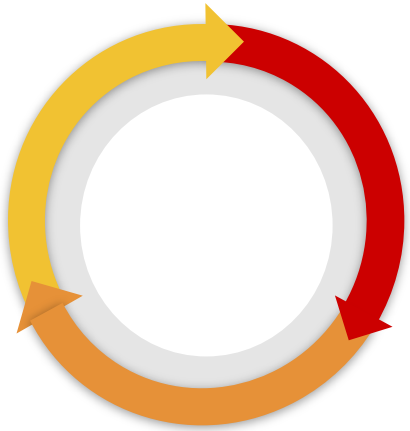


Training datasets :

- several zones in France
- several objet type

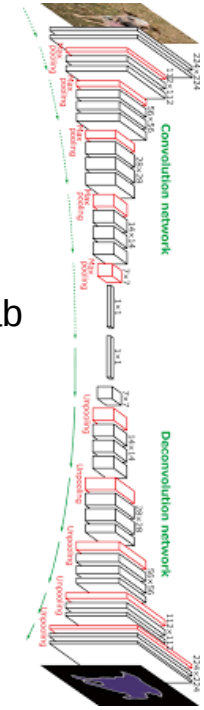
Raw data sources :

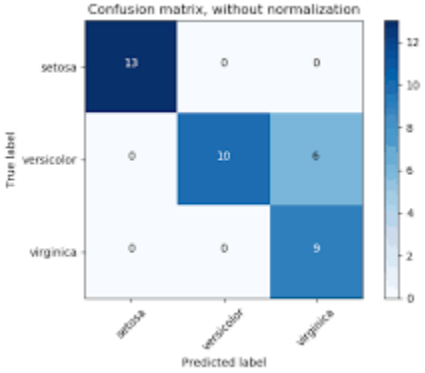
- BD ORTHO
- BD ALTI
- BD UNI
- RG For
- SPOT
- …



Models :

- UNet
- DeepLab





Evaluation :

- qualitative
- quantitative

Model **Comparaisons**
Analysis of detection failures

Ground truth from BDUNI/RGFor

Termos v1 Deep Learning output

IGN

Zoom on timber

or how to relatively quickly build a ground truth data

What is in the database (RGFor)

What we need : exhaustivity

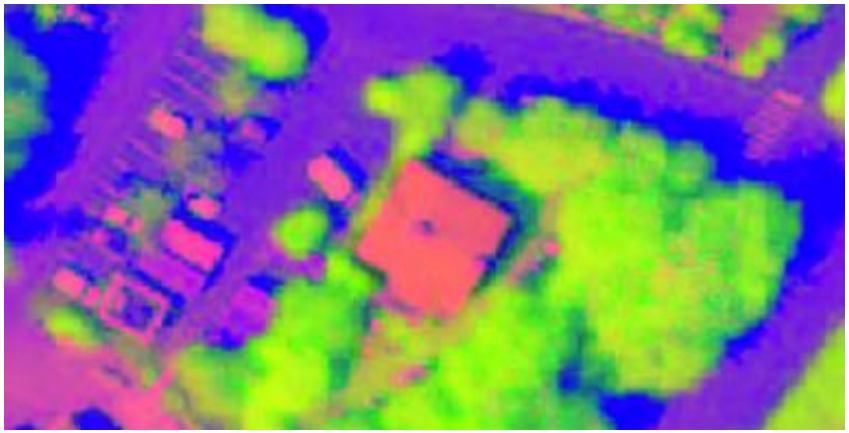
Methodology using Orfeo toolbox

Apply a geometric segmentation at the right level of detail

→ no shadows, separation between trees and grass

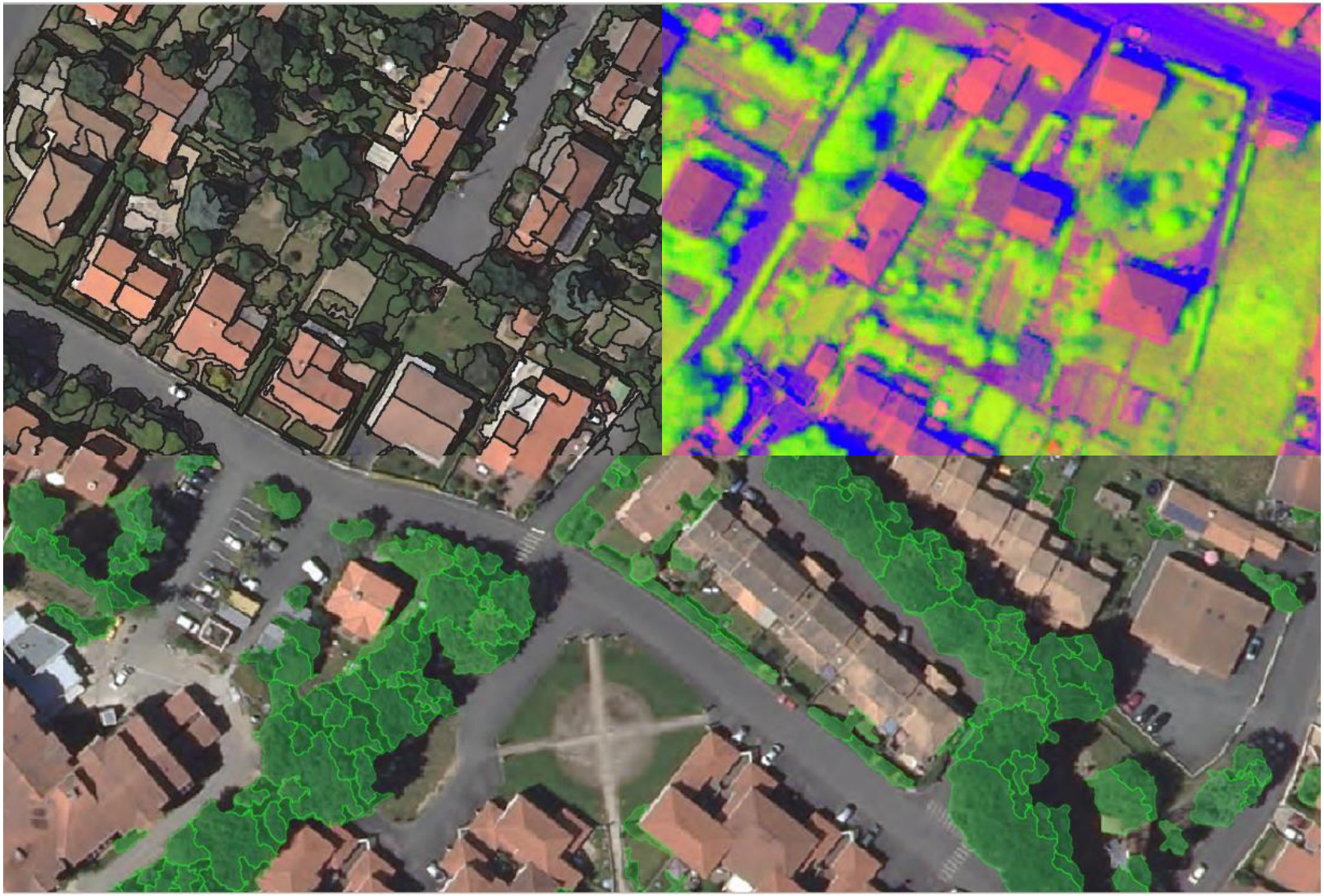
Use **Machine Learning**

- Manually define features (ndvi, bi2, ...)
- Apply Random Forest algorithm to distinguish timber from the rest



Orfeo ToolBox : Open Source processing of remote sensing images (https://www.orfeo-toolbox.org/)

IGN

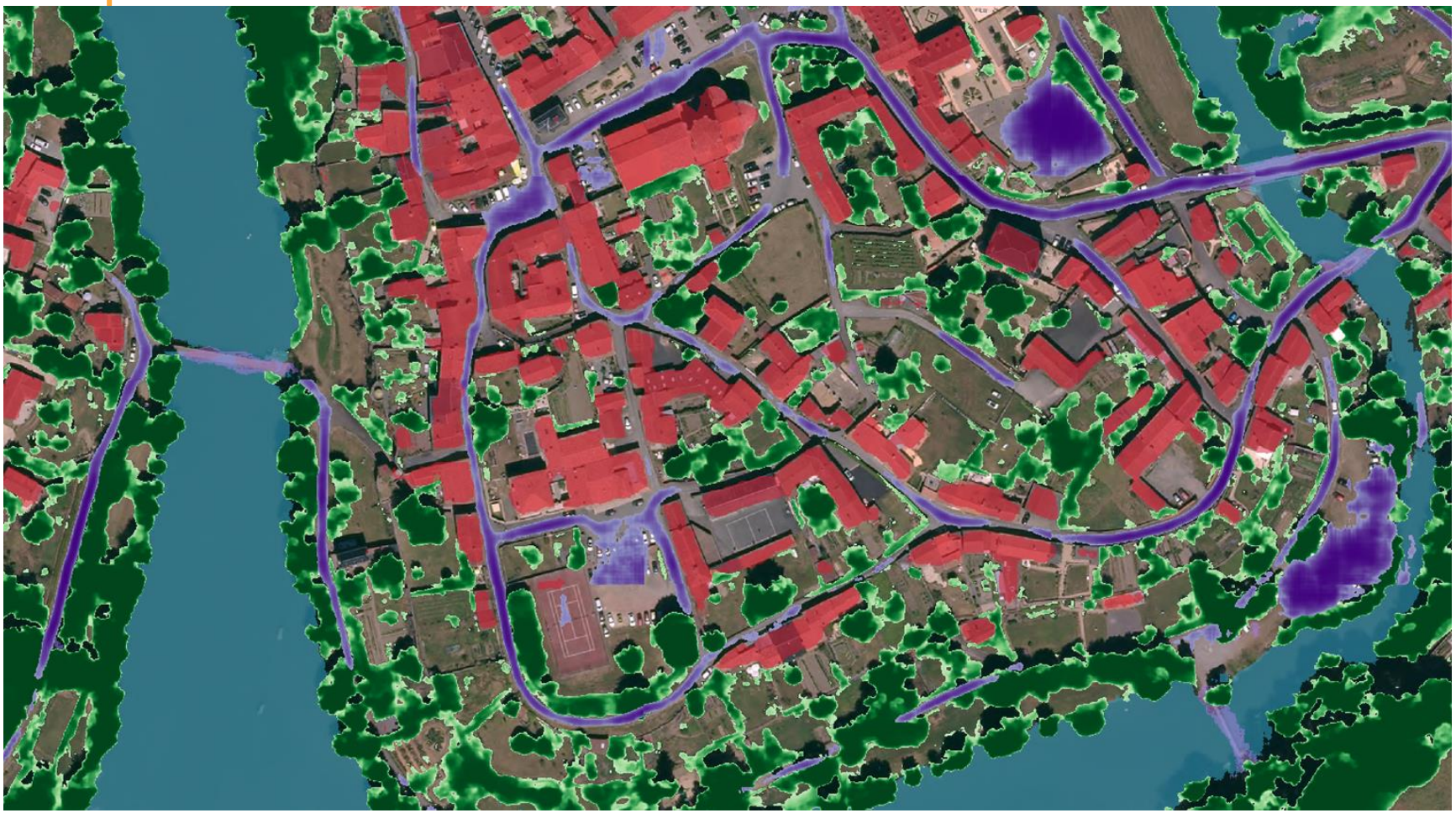


Geometric segmentation and feature representation (top)
Machine learning output (bottom)

IGN

Summary and perspectives

Up to date results : better timber, work in progress on asphalt



What is a “good data” ?

Nomenclature

- IGN: complex (minimum surface, mixed component)
- DL: what you see is what you learn / get

Visual rendering

- IGN: vector
- DL: raster

Semantic

- IGN: usage or cover
- DL: mostly cover

IGN strength

We are currently in the process of producing a very accurate and clean dataset for **urban areas** (with 7 general classes and several subclasses) in France (10 "departements" from 2015 up to 2018)

→ datasets are the most valuable asset in the Deep Learning community (famous dataset : ImageNet, Mapillary, ...)

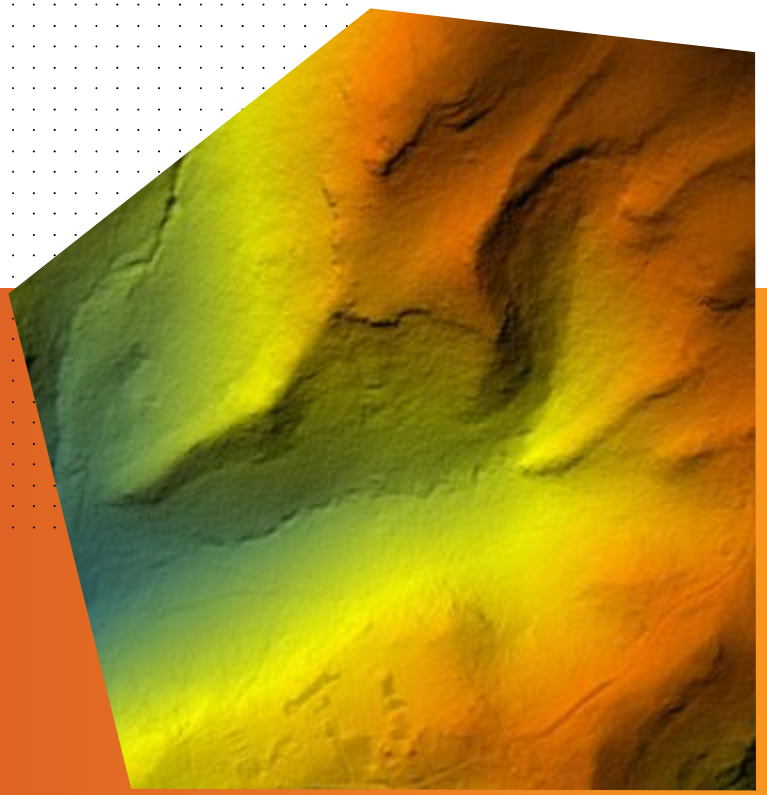
We have access to more data, particularly height information, close infrared images

→ allows us to limit the size of dataset (<< 1M samples) but we can't profit from pre-trained models

Perspectives

- Change detection and alert system
- Accelerate the update of main IGN databases (BDUni, RGFor, ...)
- Fault detections during the transformation from PVA to orthos photos
- Real time Mobile Mapping
- ...

IGN
INSTITUT NATIONAL
DE L'INFORMATION
GÉOGRAPHIQUE
ET FORESTIÈRE



THANK YOU FOR YOUR ATTENTION

ign.fr