RÉPUBLIQUE
FRANÇAISE

Liberté
Égalité
Fraternité

IGN
INSTITUT NATIONAL
DE L'INFORMATION
GÉOGRAPHIQUE
ET FORESTIÈRE

Institut national de l'information géographique et forestière

RÉPUBLIQUE
FRANÇAISE
Liberté
Égalité
Fraternité

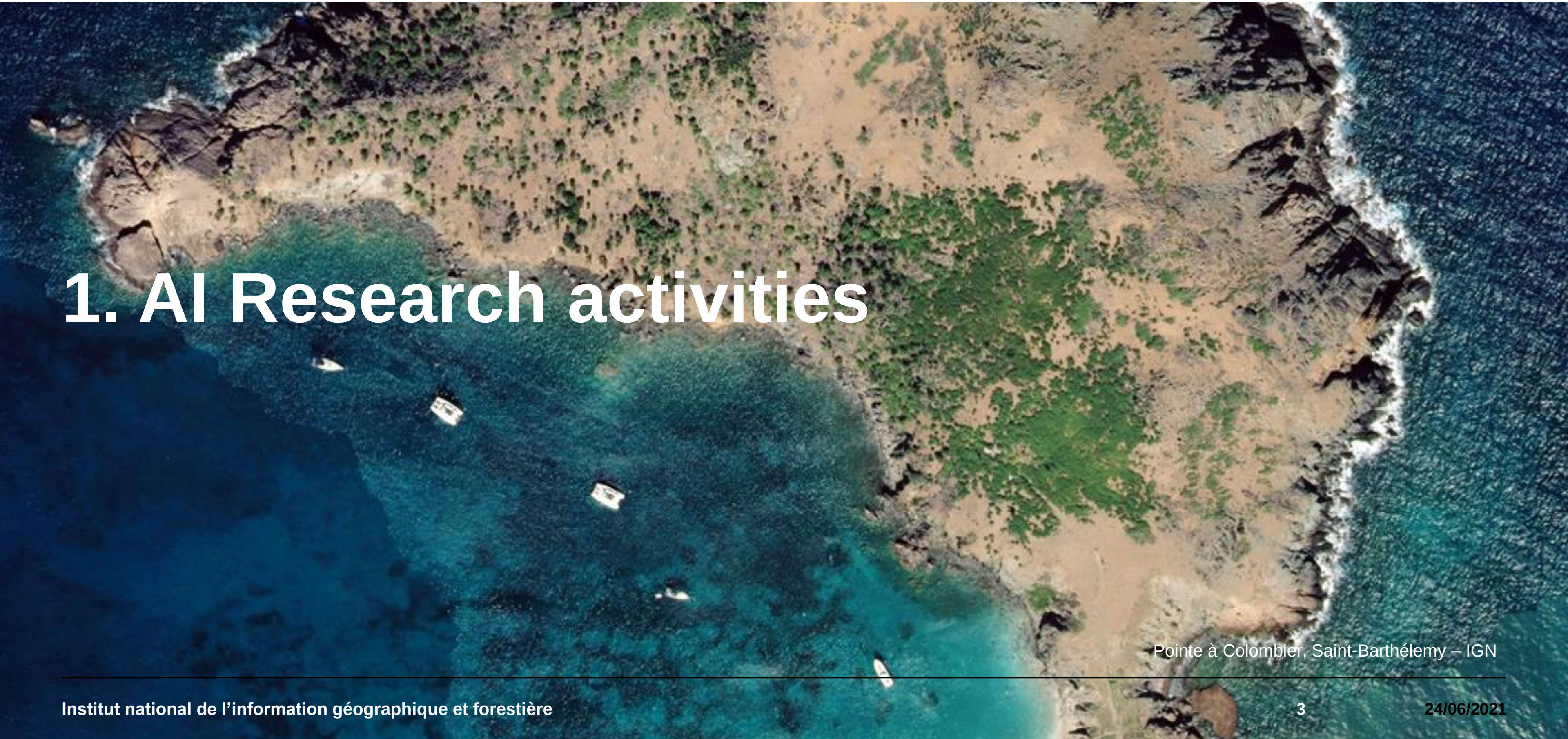
INSTITUT NATIONAL
DE L'INFORMATION
GÉOGRAPHIQUE
ET FORESTIÈRE

ARTIFICIAL INTELLIGENCE AT IGN FRANCE

24/06/2021

Matthieu Porte

1. AI Research activities

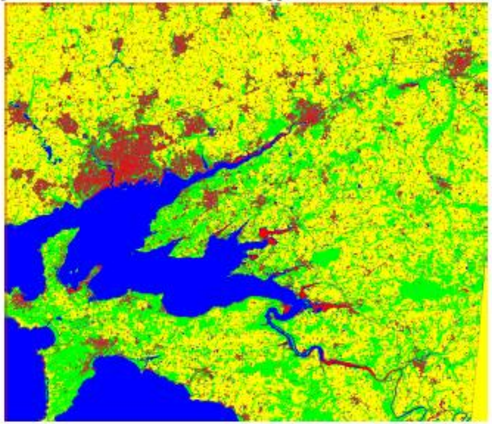


Pointe à Colombier, Saint-Barthélemy – IGN

Land cover mapping

- DL methods studied since 2014-2018 (T. Postadjian doctoral thesis)
- Ongoing research : other sensors, historical data,…
- Neural architectures with relevant inductive biases

(A) Région A, classifiée avec un modèle appris sur la zone délimitée en noire.



(B) Occupation des sols sur la zone du Finistère avec le réseau entraîné sur ROI-1 (la zone de Brest étudiée en section 3.2).
Les dimensions de cette zone sont de 39×45 km.
● *Bâti*, ● *Route*, ● *Culture*, ● *Végétation*, ● *Eau*.

Vers une occupation du sol France entière par imagerie satellite à très haute résolution, T. Postadjian (2018)

Land cover mapping

Satellite Image Time Series Classification with Pixel-Set Encoders and Temporal Self-Attention, Sainte-Fare-Garnot et al.

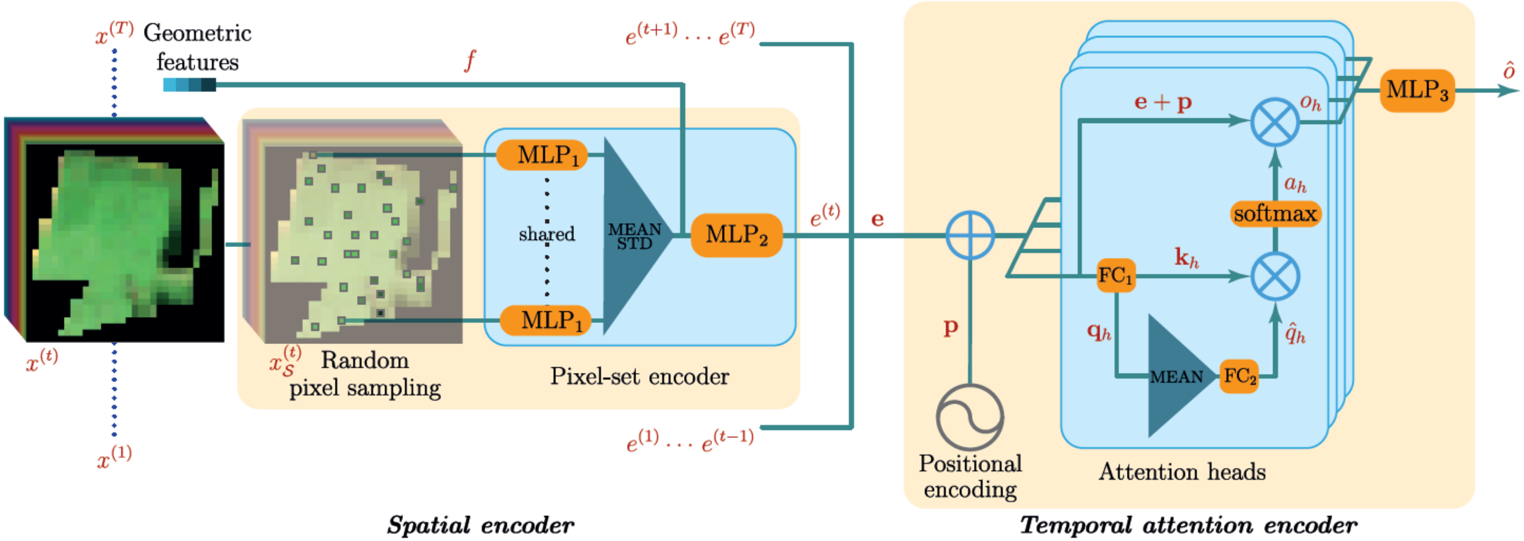
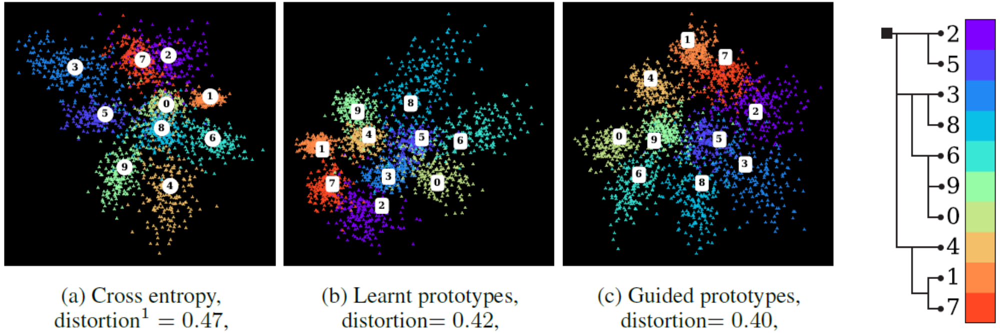


Figure 2: Schematic view of our spatio-temporal encoder. Variables in bold are tensors concatenated along the temporal dimension, *e.g.* $\mathbf{e} = [e^{(0)}, \cdots, e^{(T)}]$.

Land cover mapping

Leveraging class hierarchies with Metric-guided prototype learning, Sainte-Fare-Garnot and Landrieu.



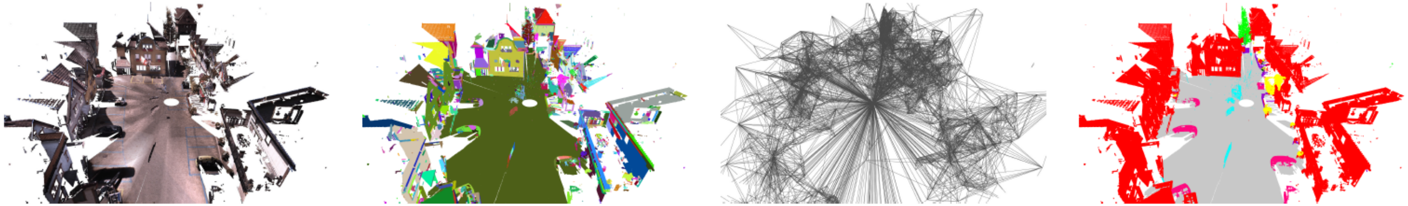
(a) Cross entropy, distortion[1] = 0.47, ER= 15.2%, AC = 0.81

(b) Learnt prototypes, distortion= 0.42, ER= 14.2%, AC = 0.75

(c) Guided prototypes, distortion= 0.40, ER= 11.9%, AC = 0.52

Figure 1: Mean class representation ○, prototypes □, and 2-dimensional embeddings learnt on perturbed MNIST by a 3-layer convolutional net with three different classification modules: (a) cross-entropy, (b) learnt prototypes, and (c) learnt prototypes guided by a tree-shaped taxonomy (constructed according to the authors' perceived visual similarity between digits). The guided prototypes (c) have lower distortion[1] with respect to the tree-induced metric: classes with low error cost are closer. This is associated with a decrease in the *Average Cost* (AC), as well as *Error Rate* (ER), indicating that our taxonomy may contain useful information for learning better visual features.

3D data



(a) RGB point cloud (b) Geometric partition (c) Superpoint graph (d) Semantic segmentation

Figure 1: Visualization of individual steps in our pipeline. An input point cloud (a) is partitioned into geometrically simple shapes, called superpoints (b). Based on this preprocessing, a superpoints graph (SPG) is constructed by linking nearby superpoints by superedges with rich attributes (c). Finally, superpoints are transformed into compact embeddings, processed with graph convolutions to make use of contextual information, and classified into semantic labels.

PyTorch Points 3D

Large-scale Point Cloud Semantic Segmentation with Superpoint Graphs, Landrieu and Simonowsky (CVPR 2018).

3D data

- Classification of 3D triangulations (inside/outside tetrahedrons) with graph-based neural networks ;
- Correlation and depth estimation

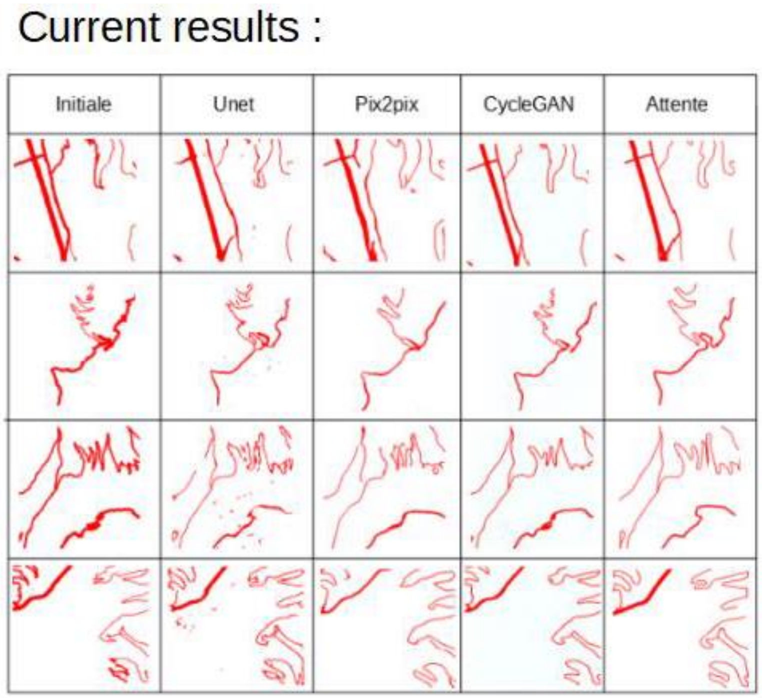
« Image-to-image » translation

- Map generalization
- Style transfer
- Super-resolution

- Heavy use of generative models

« Image-to-image » translation

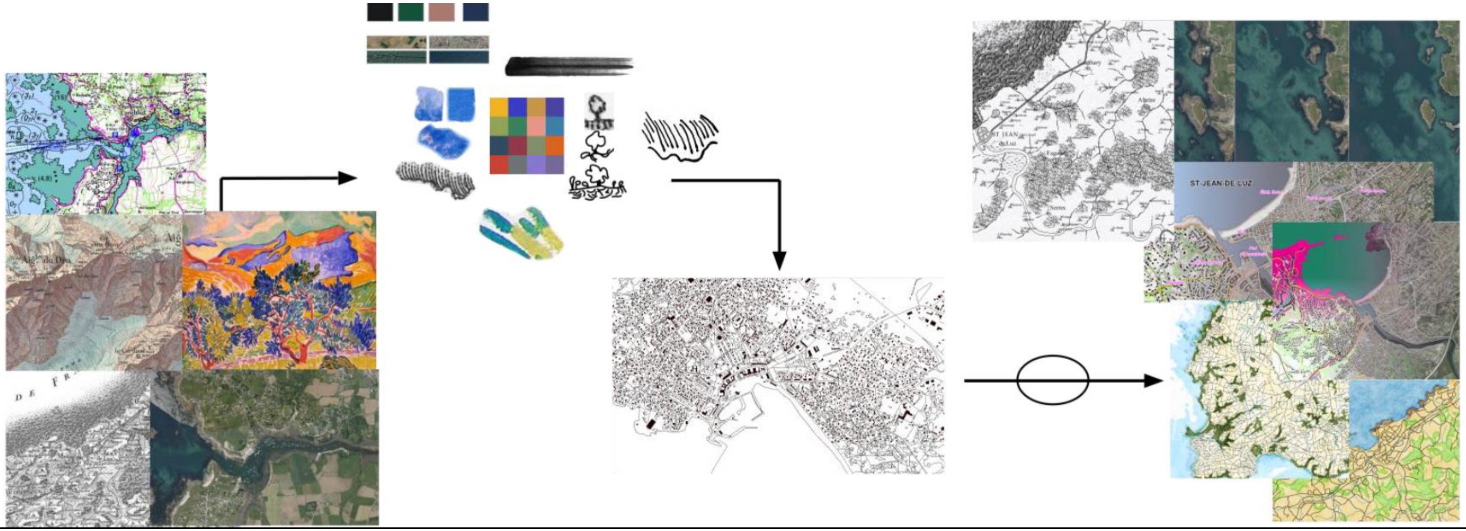
Map generalization



- *Exploring the Potential of Deep Learning Segmentation for Mountain Roads Generalisation*, Courtial et al. (2020)

« Image-to-image » translation

Style transfer

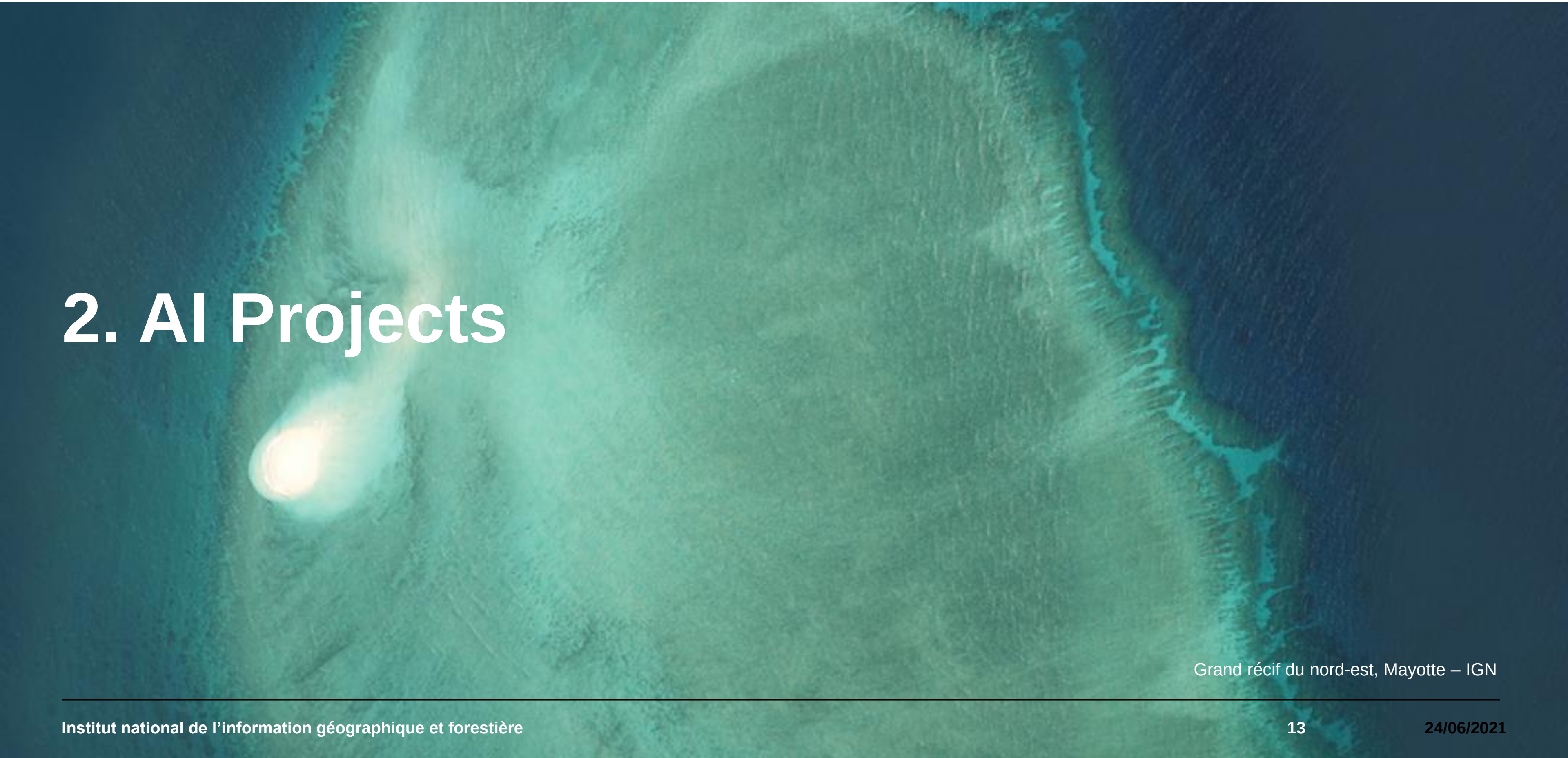


- Ongoing research with AI techniques

Other research areas

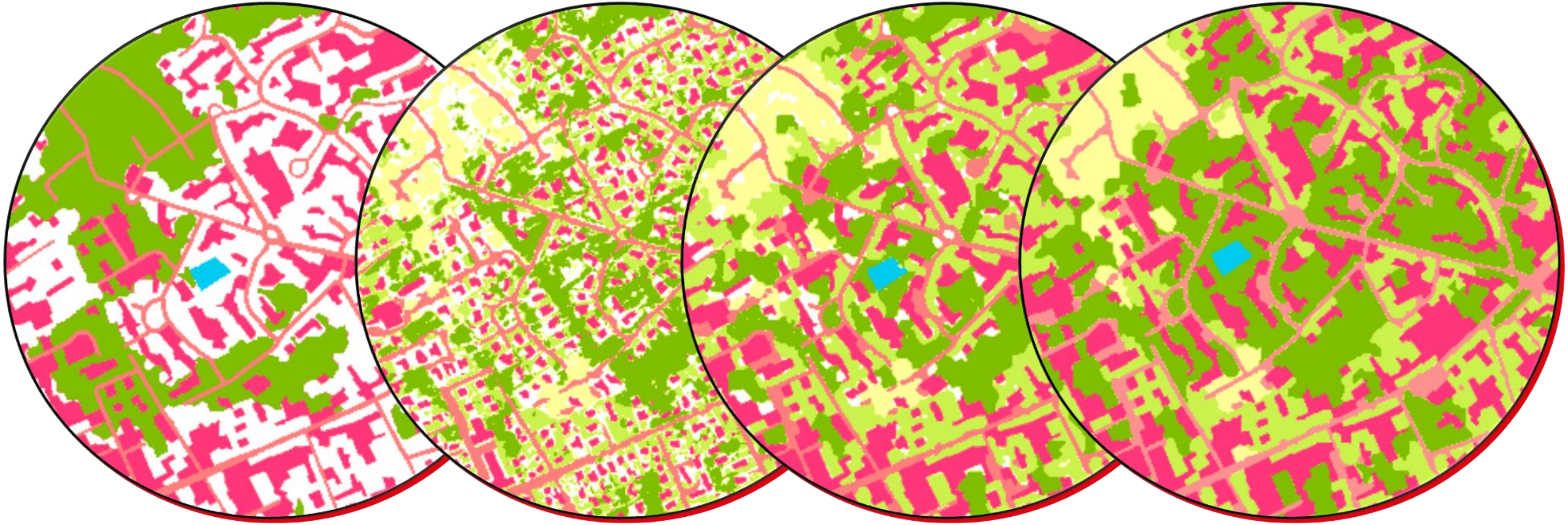
- Representation learning :
  - Multi-task learning
  - Multi-sensor representation fusion
  - Handling domain shifts
  - Pre-trained representations, unsupervised / weakly-supervised learning…
- Human-in-the-loop : active learning, uncertainty estimation…

2. AI Projects



Grand récif du nord-est, Mayotte – IGN

OCS GE upgrading



TRAITEMENT AUTOMATIQUE SANS DEEP LEARNING

DÉTECTION DEEP LEARNING

OCSGE AVEC DEEP LEARNING

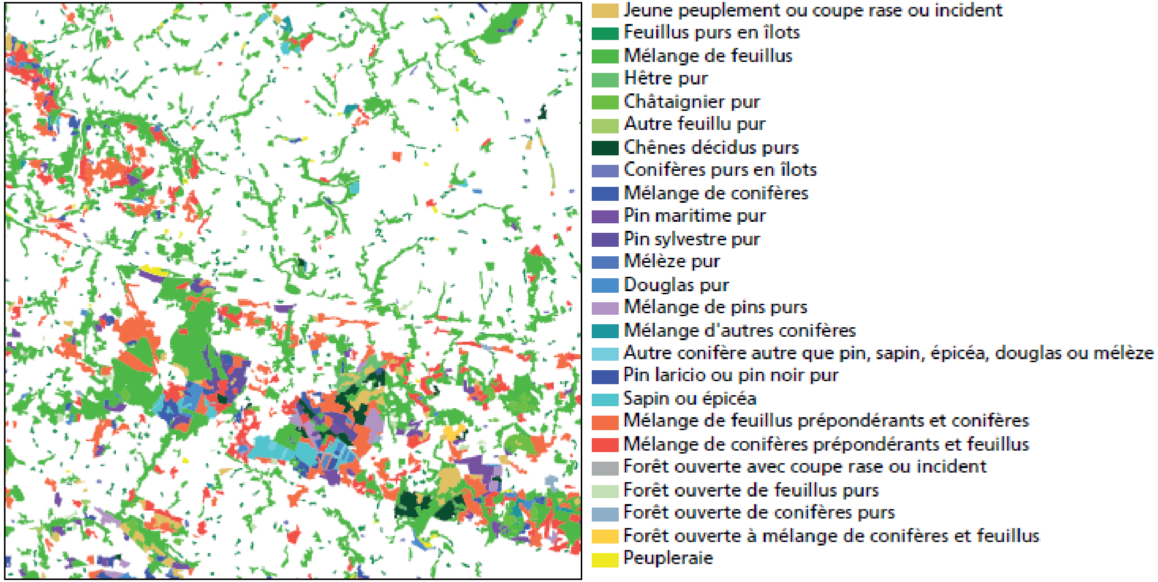
OCSGE (COUVERTURE) FINALE

OCS GE upgrading

- Deep learning is only one of many inputs
- Unable to use the historical OCS GE database as ground truth
- Production of new semantic segmentation training data
- Post-processing steps required
- Operational questions on best model stratification strategies

BD Forêt upgrading

Fig.1 - Extraits de la BD Forêt® V2 et de la BD ORTHO® infrarouge couleur, département du Morbihan (Nord-Ouest des Landes de Lanvaux, secteur ville de Baud)



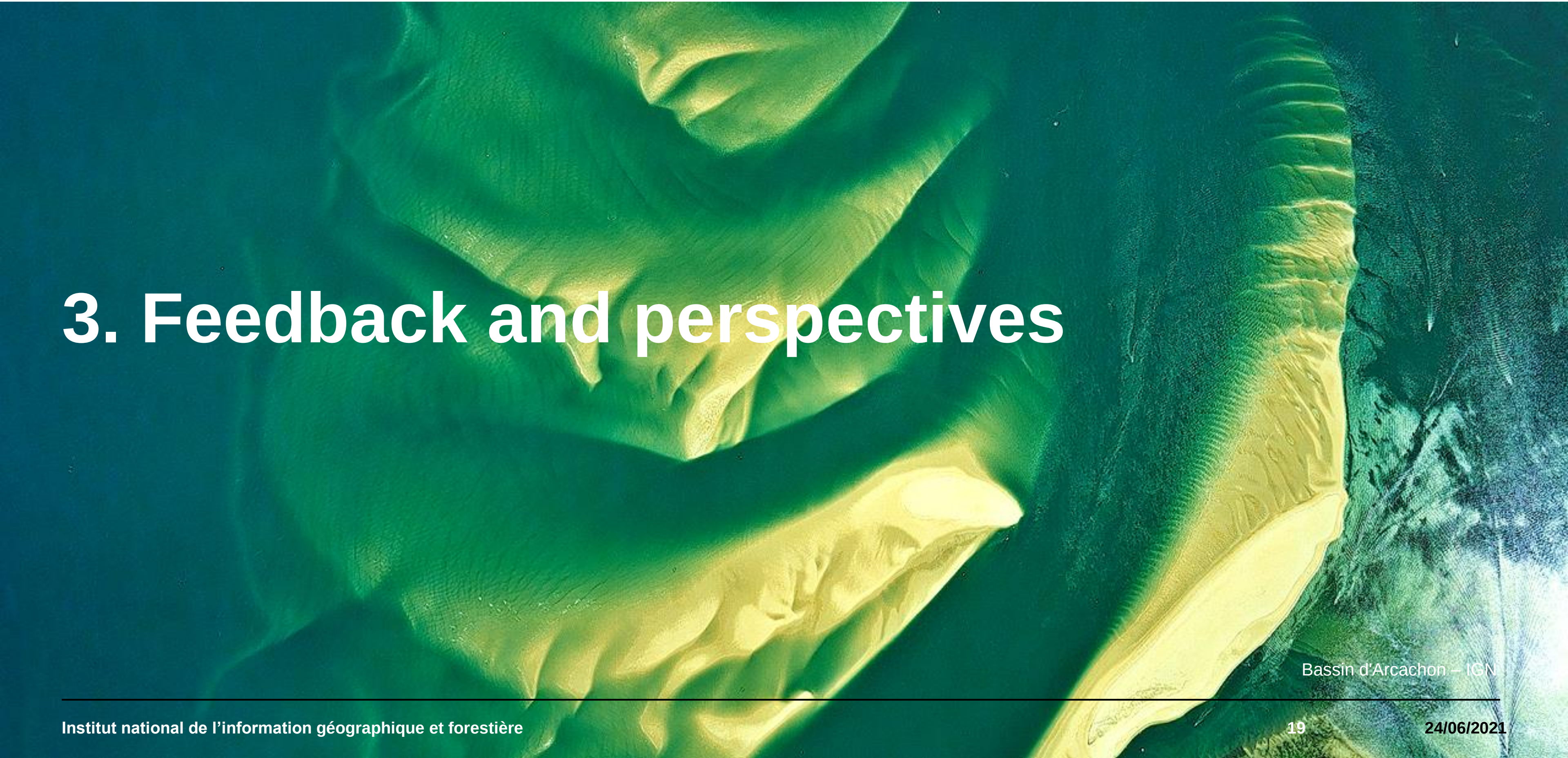
- Technical challenges :
  - Some common with OCS GE (semantic segmentation, data annotation, geographic/temporal generalization)
  - Noise in historical data / outdated data
  - Mixtures and hierarchies
  - Leveraging image time series with proper encoders

National Lidar acquisition (10pts/m²)

- 2021-2025 national acquisition
- AI experiments to improve our point cloud classification components

Other projects

- AI4GEO consortium (Airbus, CLS, CNES, CS Group, Geosat, IGN, ONERA, Quantcube, Qwant) : AI tools and applications for aerial/satellite images. Open source libraries should be released in near future.
- GNSS time series anomaly detection
- …



3. Feedback and perspectives

Bassin d'Arcachon – IGN

Handling ML projects : feedback

- Several critical and cumulative conditions :
  - Proper AI framing and delineation of the task to be automated
  - Availability, or capacity to produce, large amouts of labelled data
  - Setting up a proper AI project team, with all relevant expertise : ML/DL engineering, data engineering, thematic expertise, software engineering,… Provide management familiarity with AI techniques, limits and potentialities
  - Frequent thematic feedback to assess model performance and adjust metrics
- Historical databases are sometimes inadequate as training data, labelling effort is required
- AI models are parts of larger processing pipelines.

Thoughts and perspectives

- Need for proper model architectures, with relevant inductive biases
- Large amounts of unlabelled data should be leveraged with unsupervised/SSL and geographically sound domain adaptation techniques
- Our operational challenges are technically and scientifically relevant and should gain attention : we need more open datasets and benchmarks for those.

RÉPUBLIQUE
FRANÇAISE
Liberté
Égalité
Fraternité

INSTITUT NATIONAL
DE L'INFORMATION
GÉOGRAPHIQUE
ET FORESTIÈRE

THANK YOU FOR YOUR ATTENTION!