Artificial Intelligence for Earth Observation - ESA Φ-lab

Artificial Intelligence (AI) Research

AI for Earth Observation (EO) - AI4EO

19 July 2024

Presenter: **Nikolaos Dionelis**
Research Fellow at the European Space Agency (ESA), Φ-lab, ESRIN, Italy
Nikolaos.Dionelis@esa.int

ESA UNCLASSIFIED – For ESA Official Use Only

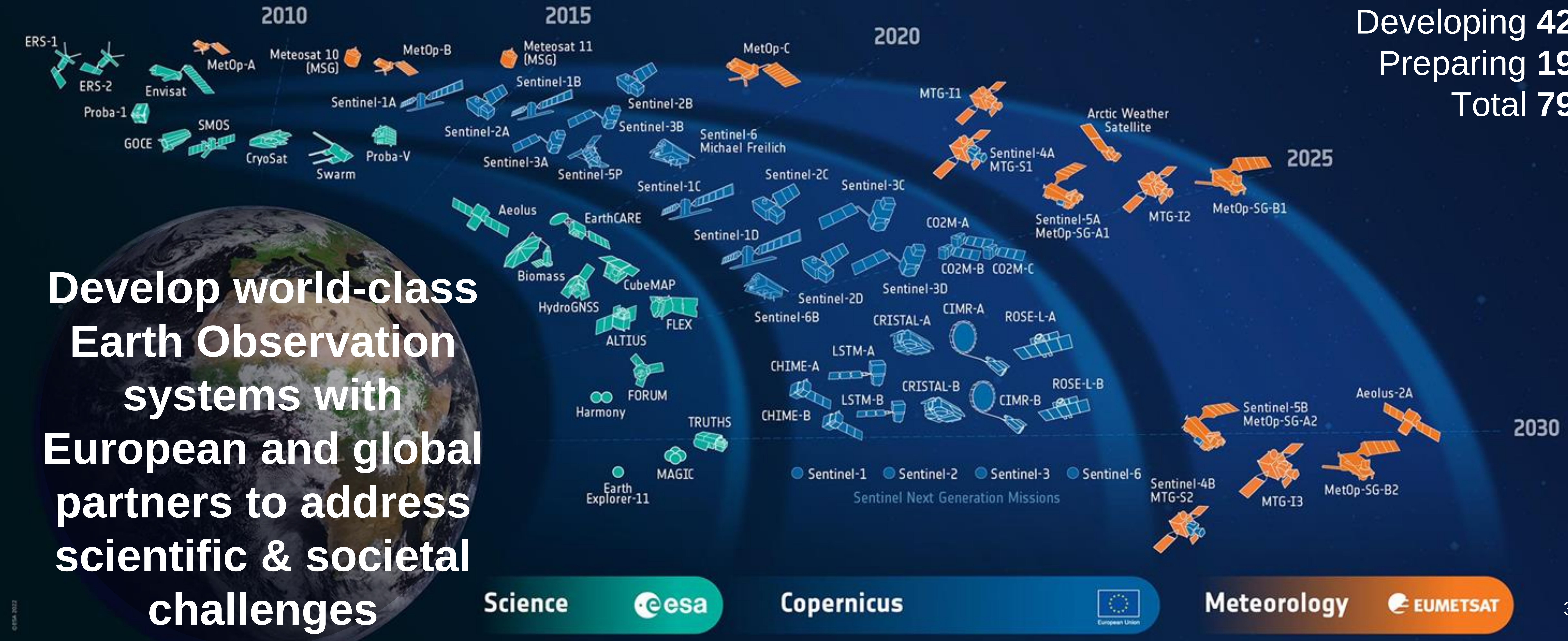
Contents

- **Artificial Intelligence** for Earth Observation (**AI4EO**)
- The Φ-lab **Explore and Invest Offices**
- **Φ-lab collaborations**
- **ESA Φ-lab Current Projects:**
    - PhilEO: Earth Observation Foundation Model and Evaluation Framework
        - **Satellite** data
        - EO Foundation Models
    - Major TOM: Expandable **Datasets** for EO and Remote Sensing
    - Learning from unlabelled data: Domain adaptation
        - Application/ **use case**: Ground-to-aerial image matching
    - **PINNs:** Physics Informed Neural Networks

ESA's Earth Observation Mission

Satellites

Heritage **04**
Operational **14**
Developing **42**
Preparing **19**
Total **79**



Develop world-class Earth Observation systems with European and global partners to address scientific & societal challenges

→ THE EUROPEAN SPACE AGENCY

from Earth Observation to Earth Action

from data to actionable information

AI opening a new dimension for EO

On Board Autonomy

Process Automatio
n

Data Science

AI

Detection/ Classification

Big Data Analytics

Super Resolution

esa

Φ-lab

ESA Φ-lab Offices: **Explore** and **Invest**

We strongly believe in
truly transformative ideas and in
the power of compelling partnerships
to accelerate the Earth Observation future

Giuseppe.Borghi@esa.int

The ESA Φ-lab Offices

Φ-lab Explore Office

Explores the innovation universe and connects together EO and digital revolution

A team of Researchers and innovation seed funding (FutureEO)

Φ-lab Invest Office

Stimulates competitiveness by fostering the growth of entrepreneurial initiatives through investment actions from ESA Member States and private investors

A team of business innovators and a commercial co-funding programme (InCubed)

The ESA Φ-lab Explore Office

Our goal is to enable a connected network of talents to brings together expertise and ideas from researchers, industry, ICT players, innovators and investors to foster scalable learning from EO data, Earth Science and Computer Science

We operates as a ESA 'hub' within the network of EO academic and industrial researchers across Europe and globally

Research Lab

Flag Ships

Φ-lab Community

Challenge Program

[Link to ESA Φ-lab Explore Office - Research & Innovation strategy, 2020]

ESA Φ-lab Invest Office

Our goal is to stimulate and develop all transformative actions that ESA can leverage to strengthen European EO industrial competitiveness and leadership

In this role ESA behaves not as customer but as the de-risking partner and facilitator stimulating competitive growth of the economic operator

in³

Φ-lab co-invest program

Invest Action

Φ-lab Community

The ESA Φ-lab Invest Office

Φ-lab co-invest program

Offers investment opportunities to support and develop innovative and commercially viable products and services. Encourages high-risk/high-potential developments mitigating the technical and financial risks. Implemented via the ESA InCubed+ Program

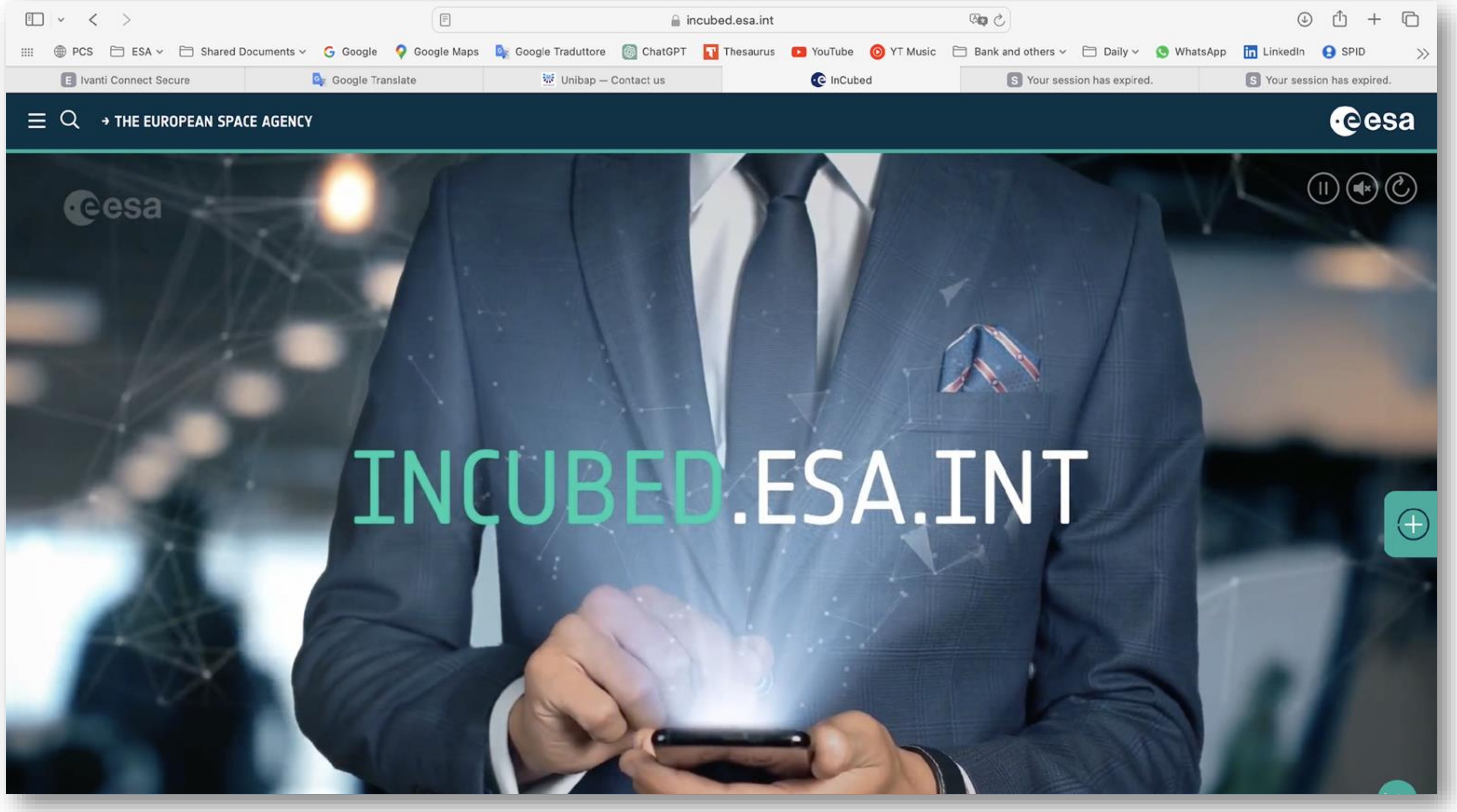
Invest Action

Accelerates access to risk capital tools for innovation funding to our ecosystem, in particular start-ups and SMEs

Φ-lab Community

Fosters industry-to-industry and industry-to-academia synergies and cooperation to accelerate adoption of innovative business solutions

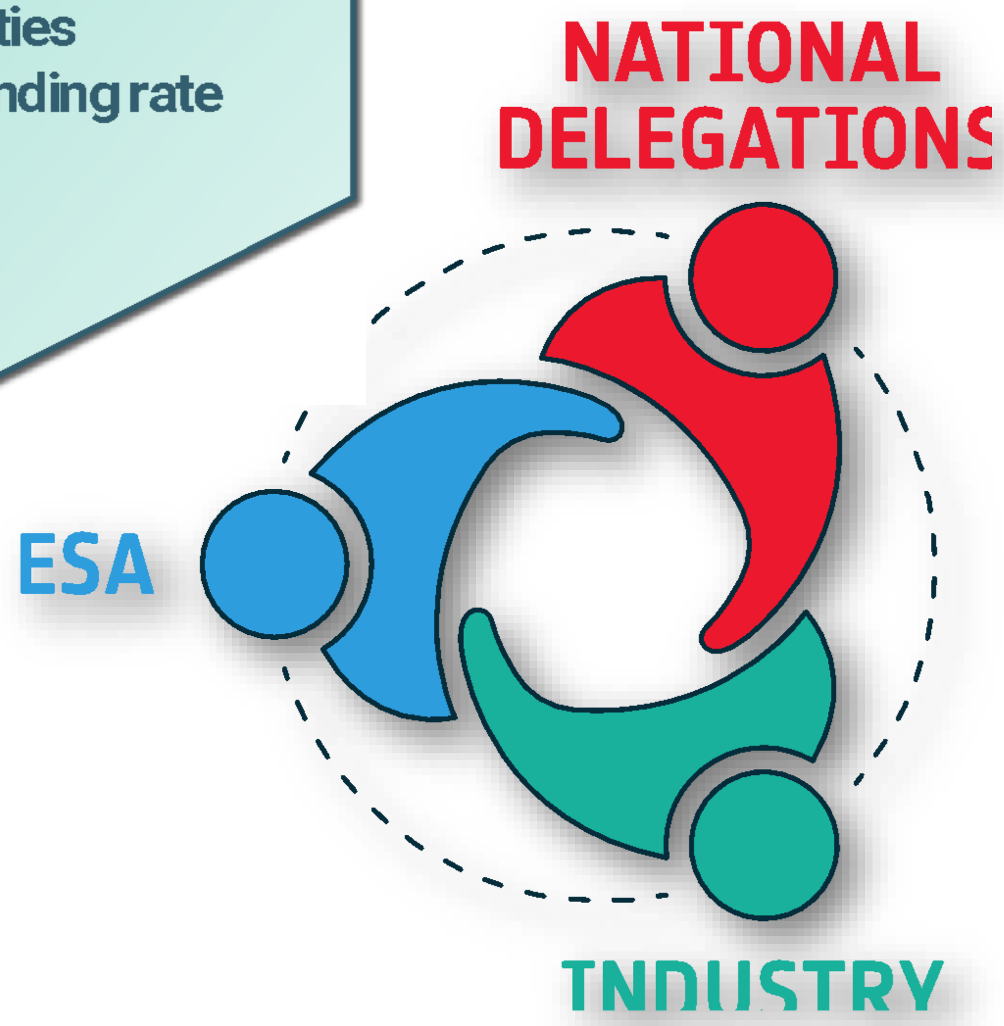
Invest in Industrial Innovation (InCubed)



€205M
InCubed fund size

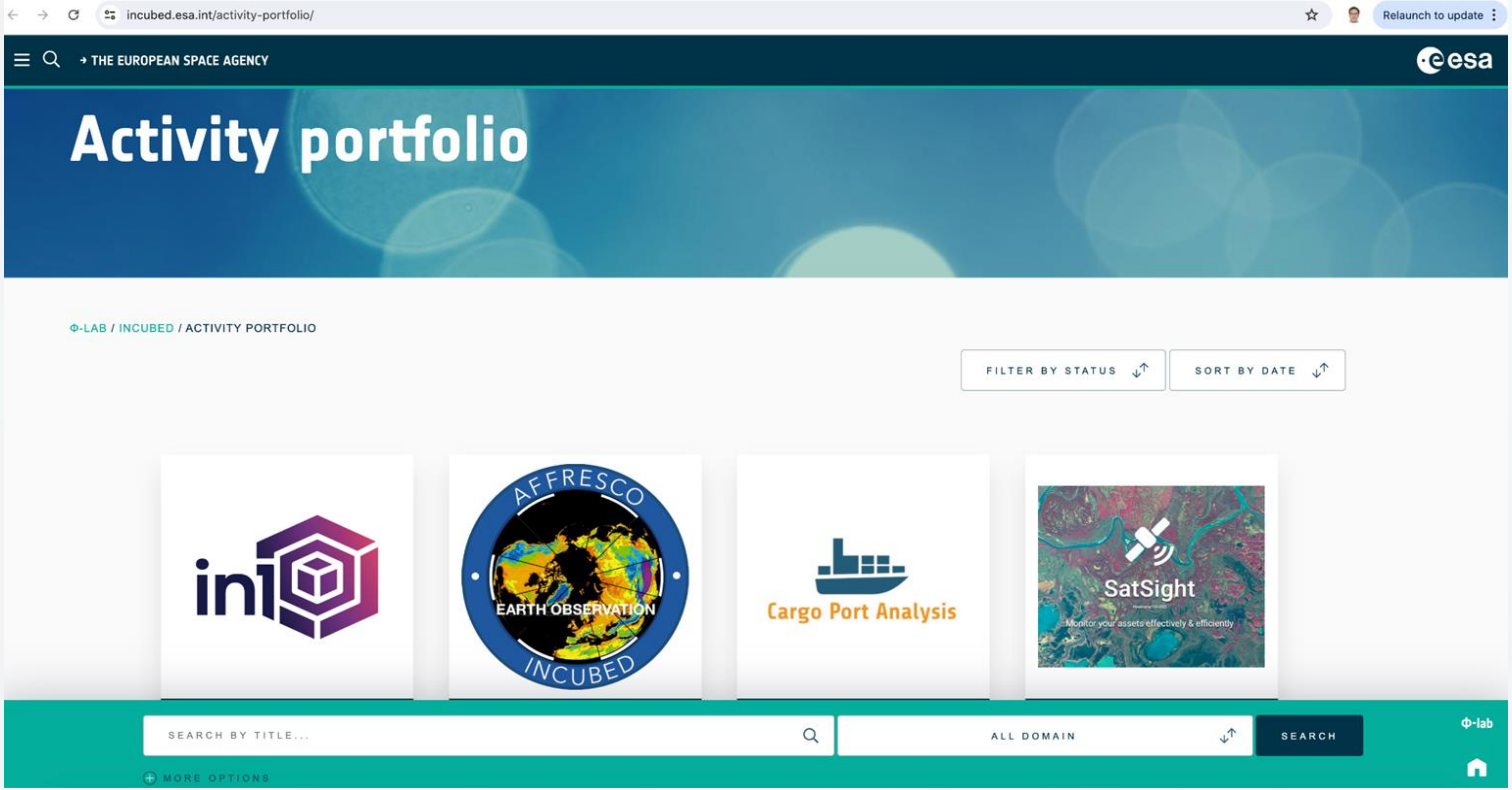
150
Activities
@62% co-funding rate

100k€ to >15m€
Project size



- Personalised technical and commercial guidance
- Zero-equity and zero-IPR
- ESA stamp of credibility
- Privileged access to commercial services enabling your development
- Access to ESA EO facilities and Φ-lab community

http://incubed.esa.int/activity-portfolio



esa

Φ-lab

ESA Φ-lab collaborations

We strongly believe in
truly transformative ideas and in
the power of compelling partnerships
to accelerate the Earth Observation future

Giuseppe.Borghi@esa.int

(some) Collaborations and parternships

esa

Massachusetts Institute of Technology

ETH

TUM

Technische Universität Berlin

POLITECNICO MILANO 1863

EPFL

CERN

DLR

ASI Agenzia Spaziale Italiana

SatCen European Union Satellite Centre

ECMWF

Die Österreichische Hagelversicherung

VNIVERSITAT DE VALÈNCIA

unicef

WFP

Università di Roma Tor Vergata

Jožef Stefan Institute

ISTITUTO NAZIONALE DI GEOFISICA E VULCANOLOGIA

KP LABS

IIASA

Università degli Studi del Sannio

PI SCHOOL
MACHINE INTELLIGENCE MEETS HUMAN CREATIVITY

PRIMO SPACE

ICEYE

cosine

SATLANTIS

ThalesAlenia Space
a Thales / Leonardo company

OHB

planet.

OPEN COSMOS

CGI

LEONARDO

OVHcloud

TELESPAZIO
a LEONARDO and THALES company

e-geos
AN ASI/TELESPAZIO COMPANY

Maker Faire

ellis
European Laboratory for Learning and Intelligent Systems

FDL

CLAIRE

Collaboration opportunities at Φ-lab

Join the open Φ-lab as an Industrial or University Visiting Researcher,
Visiting Professor, Research Fellow, PhD, YGT, Intership, etc.
to explore together transformational ideas

With funding

1. Φ-lab's Invitation To Tender on ESA-STARS
   - Foundation Models, Generative AI, QC4EO, Edge computing, Web 3.0, etc..
2. InCubed : partnership development of commercial products or services
3. Open Space Innovation Platform : co-funded research or researchers
4. EO Science4Society : no SOW, 100/200K, 6/18 months
5. ESA Technology Programmes like GSTP and TDE

esa

Φ-lab

ESA Φ-lab Current Projects

We strongly believe in
truly transformative ideas and in
the power of compelling partnerships
to accelerate the Earth Observation future

Giuseppe.Borghi@esa.int

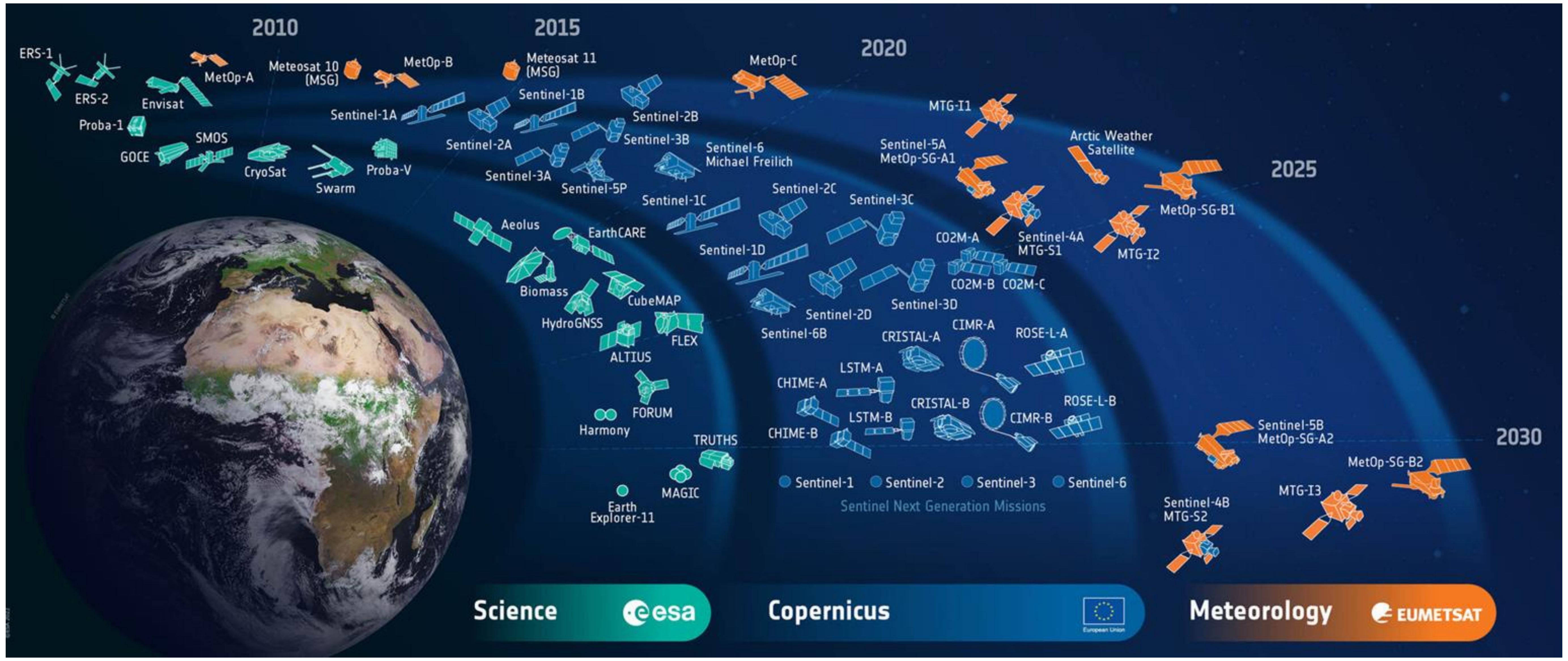
Ongoing Research

- **ESA Φ-lab:**
  - PhilEO: Earth Observation Foundation Model and Evaluation Framework
    - **Satellite** data
    - EO Foundation Models
  - Major TOM: Expandable **Datasets** for EO and Remote Sensing
    - Dataset in HugginFace, **Sentinel-2** & Sentinel-1
  - Learning from unlabelled data: Domain adaptation
    - Application/ **use case**: Ground-to-aerial image matching
  - Weather forecasting for **solar energy**
  - **PINNs:** Physics Informed Neural Networks

EO Foundation Model and Evaluation Framework

Φ-lab: Nikolaos Dionelis, Jente Bosmans
Joint work with: Casper Fibaek, Luke Camilleri, Andreas Luyts, Bertrand Le Saux

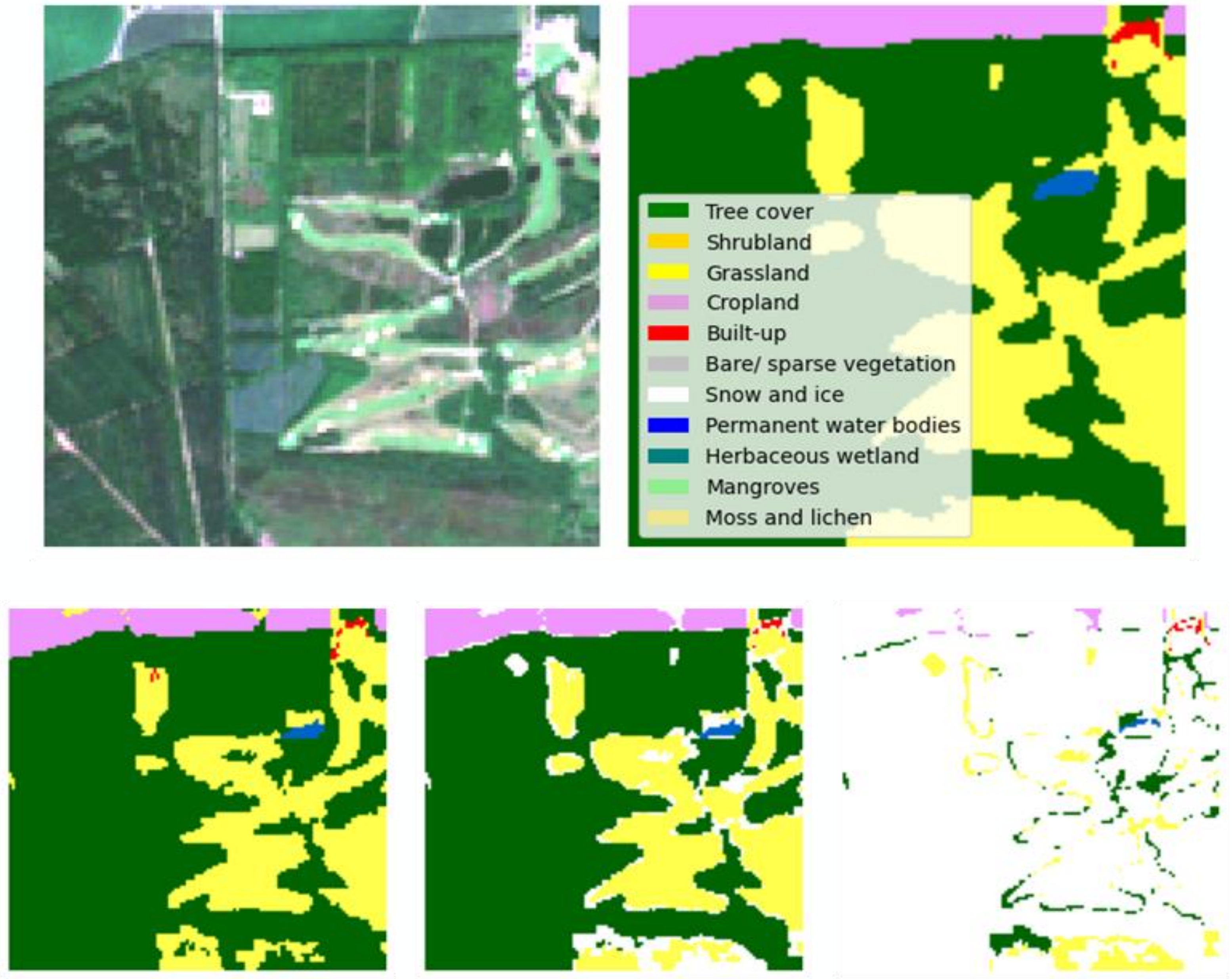
Introduction



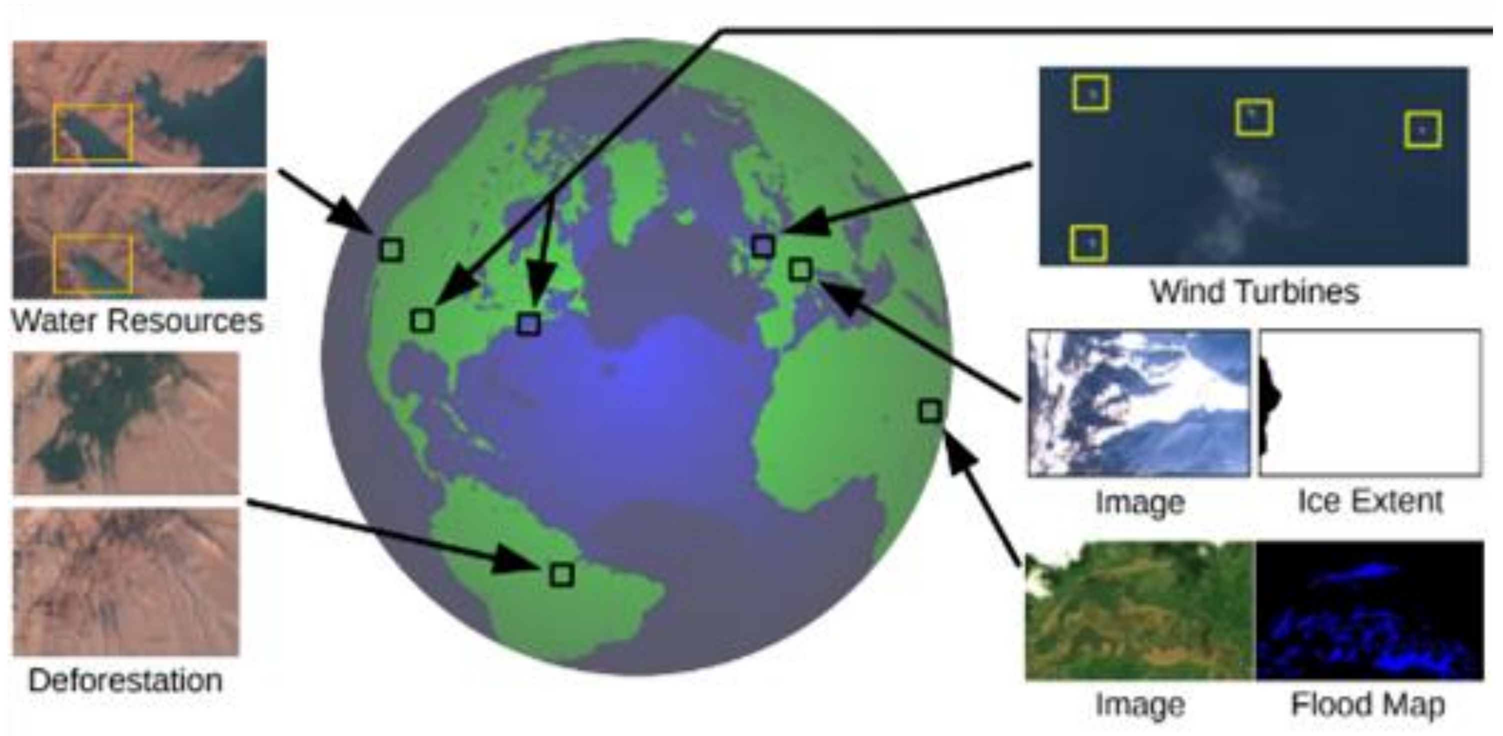
Importance & Motivation

- Large amounts of **unlabelled data** are captured by satellites
  - Copernicus Sentinel-2 constellation: Generates 1.6TB of data *daily*
- EO and remote sensing: Data-rich domain
  - Well-suited to AI and deep learning
- **Lack of annotated data**
  - Labels
    - Need *time*, are expensive, & can be labour-intensive & imperfect
- The focus is on **Foundation Model** approaches
  - **Self-supervised learning** methods using unlabelled data
  - Satellite data information:
    - **Geo-location longitude & latitude**

Advantages and Role of Foundation Models in EO



From top left: Sentinel-2, Prediction by our model, Ground truth, *Correct* classifications of the model, Incorrect classifications



"SatlasPretrain," http://arxiv.org/pdf/2211.15660

Target applications of EO Foundation Models

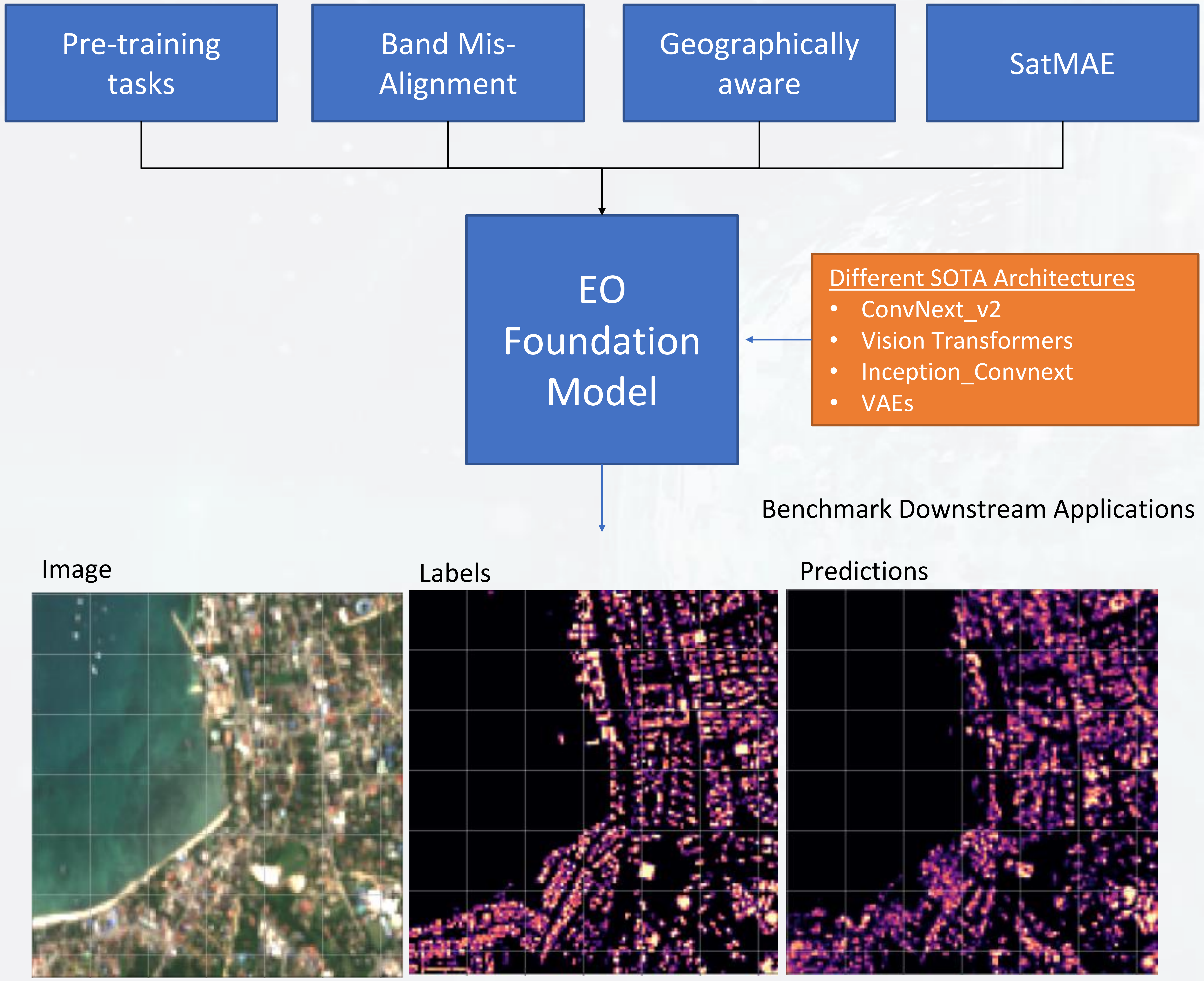
- Start with: *End goal*
- Several **different target applications:** Downstream tasks of Foundation Model

- We focus on solving **groups** of downstream tasks that are of interest to ESA
- The problems we are trying to *solve* at ESA:
  - e.g.,- **Land cover classification**
  - Building density estimation
  - Road density regression

- **Performance**
  - Deal with several tasks **jointly**
    - For each task: A prescribed *high* performance (e.g., accuracy 95%)

- ***Sharing* across tasks:** A common module
- Label efficiency

PhilEO Bench Evaluation Framework



The problem we want to solve

- EO satellites: Massive amounts of non-annotated data
  - Sentinel-2: Every day: 1.6TB of data
- **Aim:** Semantic segmentation using few labelled data, i.e. being label efficient
  - Limited number of labels
- Start from an EO Foundation Model and perform semantic segmentation as a downstream task
  - Model retraining
  - Also: Regression, Patch classification
- ESA Evaluation Framework for any Foundation Model for EO and geospatial AI

The more general problem

- General methodology: Labelled & Unlabelled data
- Unlabelled data: Pre-training/ pre-text task:
  - Geo-location prediction: Longitude & Latitude

Related Work

- *Many* **different EO Foundation Models**

Table 1. Current Foundation Models, trained by self-supervised learning and evaluated on disparate downstream tasks.

Name	Dataset	Dataset features	Architecture	Downstream tasks	Benchmark
Prithvi	HLS V2 L30	Coverage: USA, 30m res, 6 bands, Size: 1TB, Time	ViT, MAE, 3D, 100M params, Geolocation	Flood detection, Burn scars, Multi-temp crop	Sen1Floods11: mIoU 88.7%. HLS Burn Scars: IoU 73% burn class
SatMAE	fMoW	Size: 3.5 TB, Global, 10m res, 8 bands	ViT, MAE, With geolocation, With time	LULC, Multi-label, Building segmentation	NAIP: Acc: 72%, BEN: mAP: 82%, SpaceNet: mIoU: 78%
SeCo	SeCo	Size: 36 GB, 1M samples, Global, 10m res, RGB	ResNet, MoCo-v2, 23M params, Geolocation	LULC, Change detection, Fine-tuning and LP	BEN: mAP 87%. EuroSAT: Acc: 93%. OSCD: F1: 46.9%
Satlas	SatlasPre-train, 137 classes	Size: 30 TB, Global, 10m res & 1m, 3 bands (RGB) and 9, With time	Swin Transf, Shifted window self-attention, Labeled data, Geolocation	LULC, Segmentation of roads, buildings, ships. Multi-scale features	UCM: Acc: 99%, RESISC: 98%, AID: 88%, FMoW: F1-score: 44%, Roads: 87%, Buildings: 88%
				PhilEO (Ours): Building density estimation, Road extraction, LULC	Evaluation framework using annotated Sentinel-2 data that are global, 10m res, 10 bands

HLS = Harmonised Landsat Sentinel-2; LULC = Land Use Land Cover; LP = Linear Probing; BEN = BigEarthNet dataset, OSCD = Onera Satellite Change Detection

- Many additional EO FMs including, for example, **USat** and SkySense

Multi-modal EO Foundation Models

- Data efficiency: **Quicker time to value**
  - ***IBM*: Foundation models: Opportunities, risks and mitigations:**
  - http://www.ibm.com/downloads/cas/E5KE5KRZ

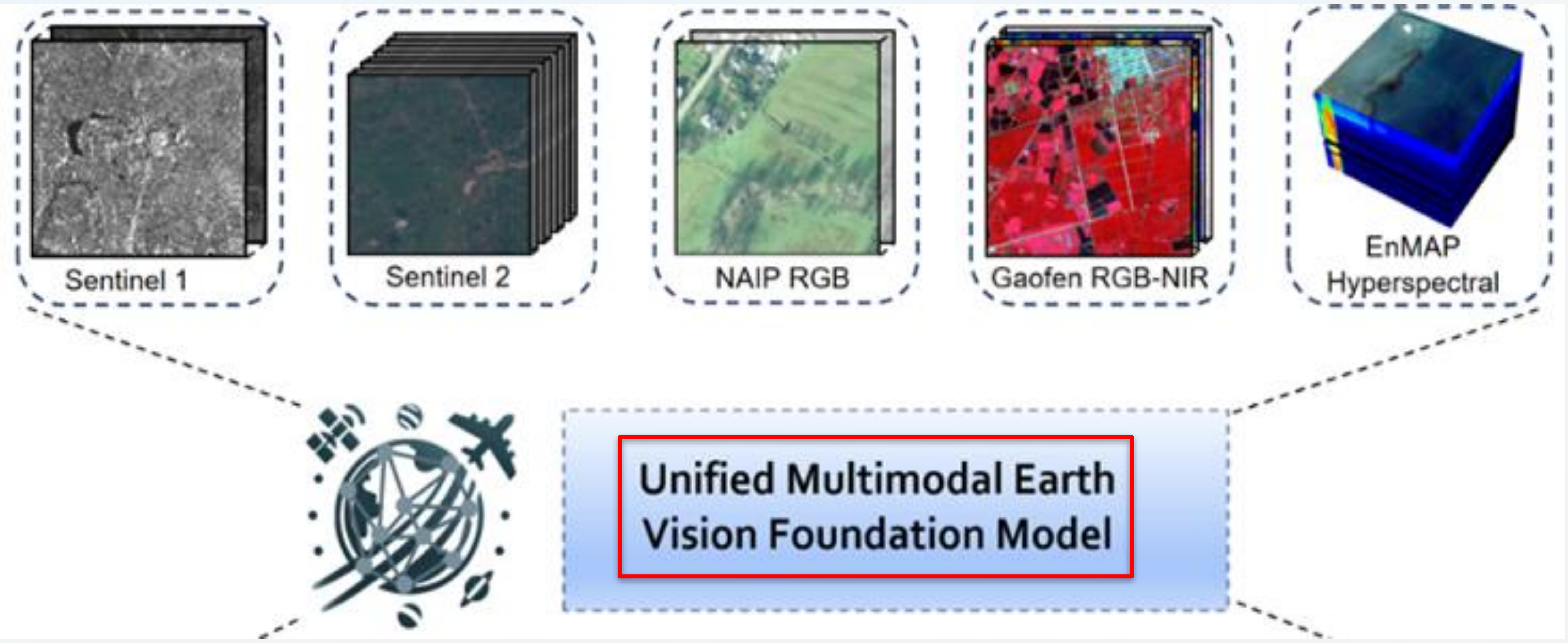
- FMs: *Paradigm shift*: http://polarview.org/news-press/foundation-models-for-earth-observation

\- **Foundation Model** for **Climate and Society** (FM4CS)
http://eo4society.esa.int/projects/fm4cs
Climate (glaciers, sea ice, icebergs)

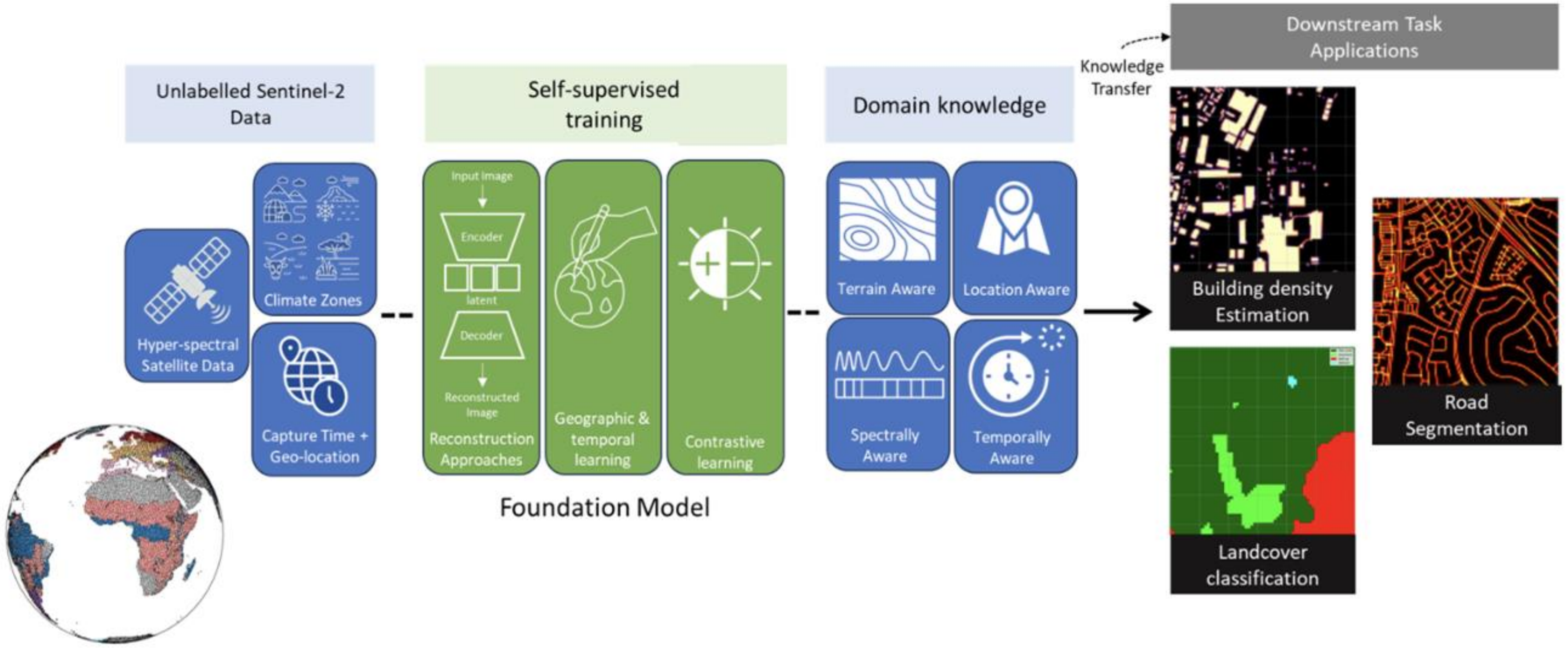
- FAST-EO
http://eo4society.esa.int/projects/fast-eo

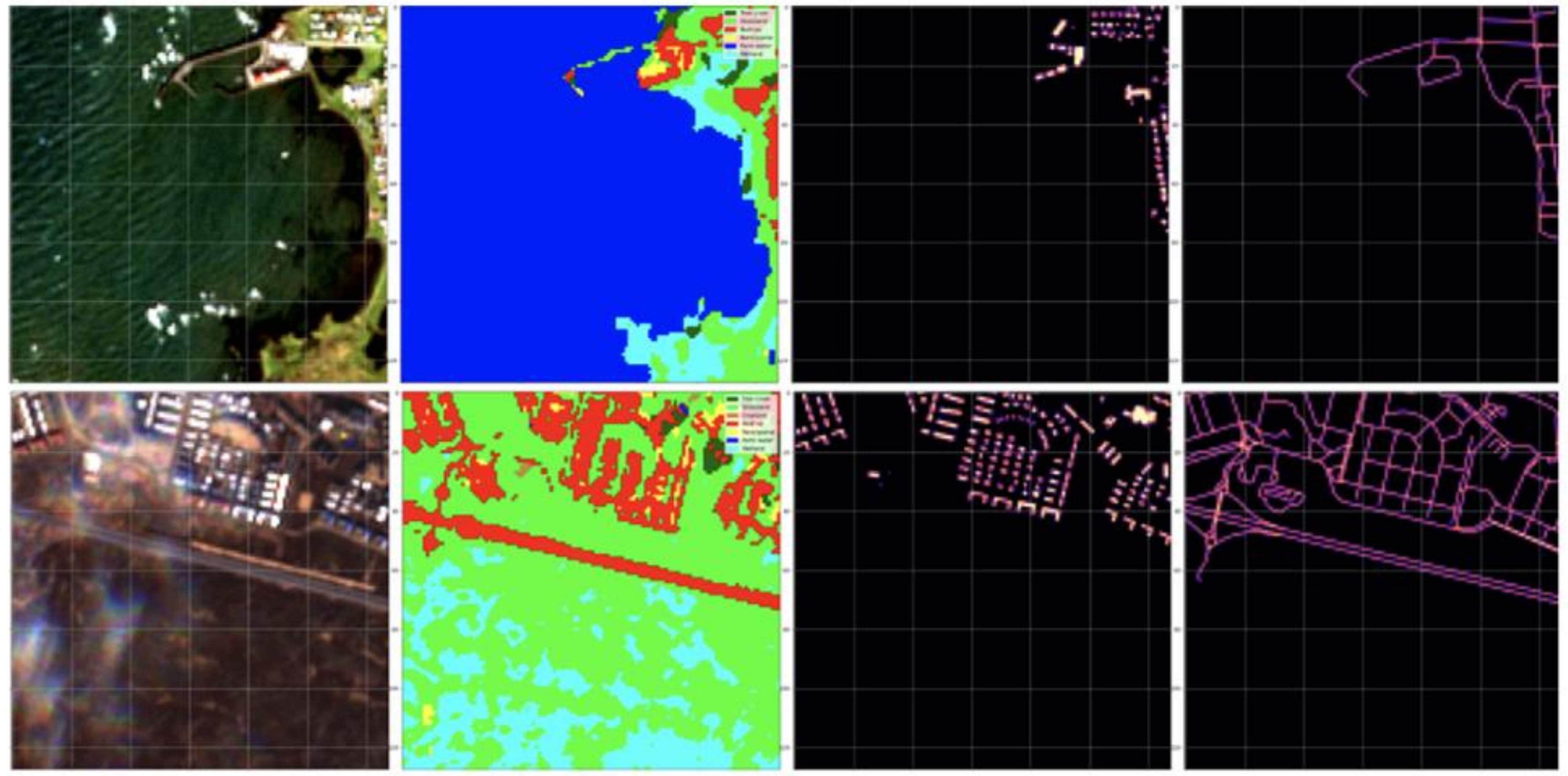
- FDL, S-1 SAR S-2 FM
http://arxiv.org/pdf/2310.00826v3

http://arxiv.org/pdf/2401.07527



Evaluation of EO Foundation Models



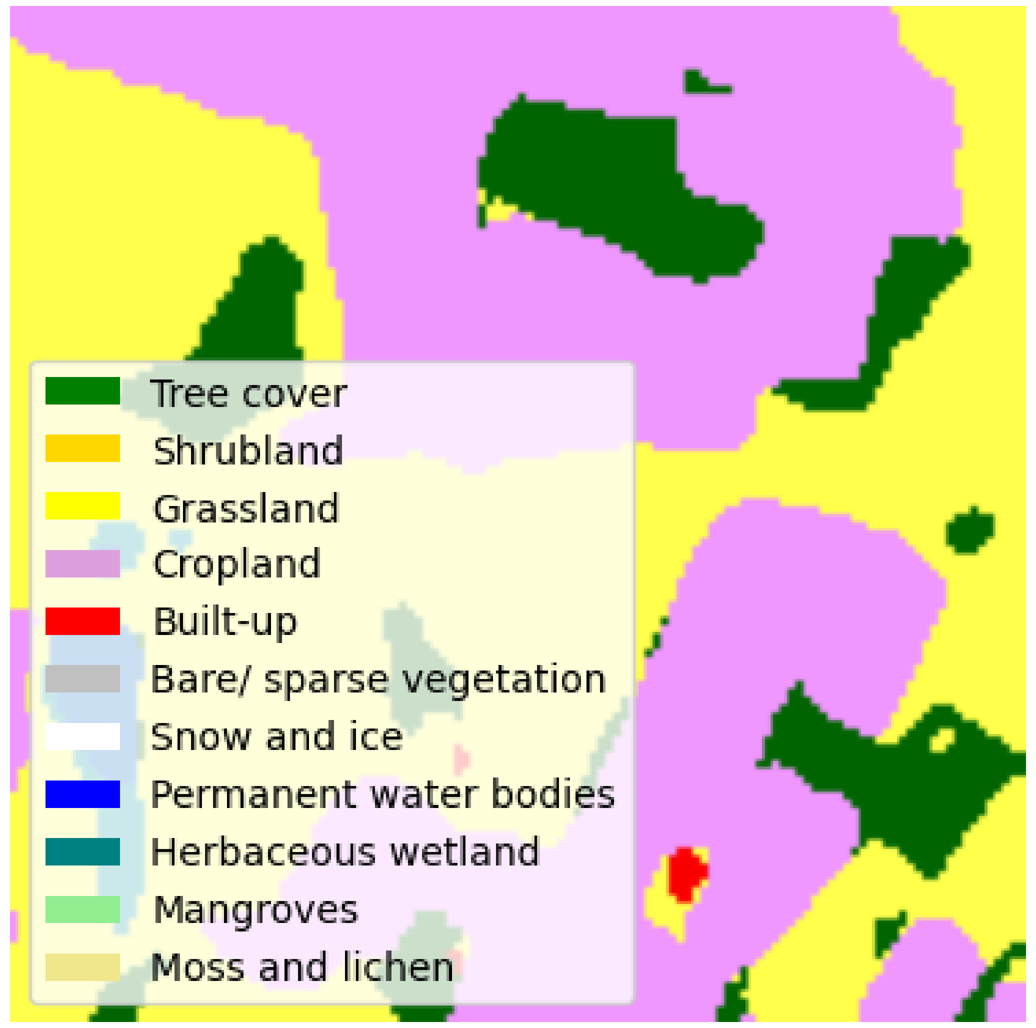


Sentinel-2, **Land cover labels, Building density, Road density**

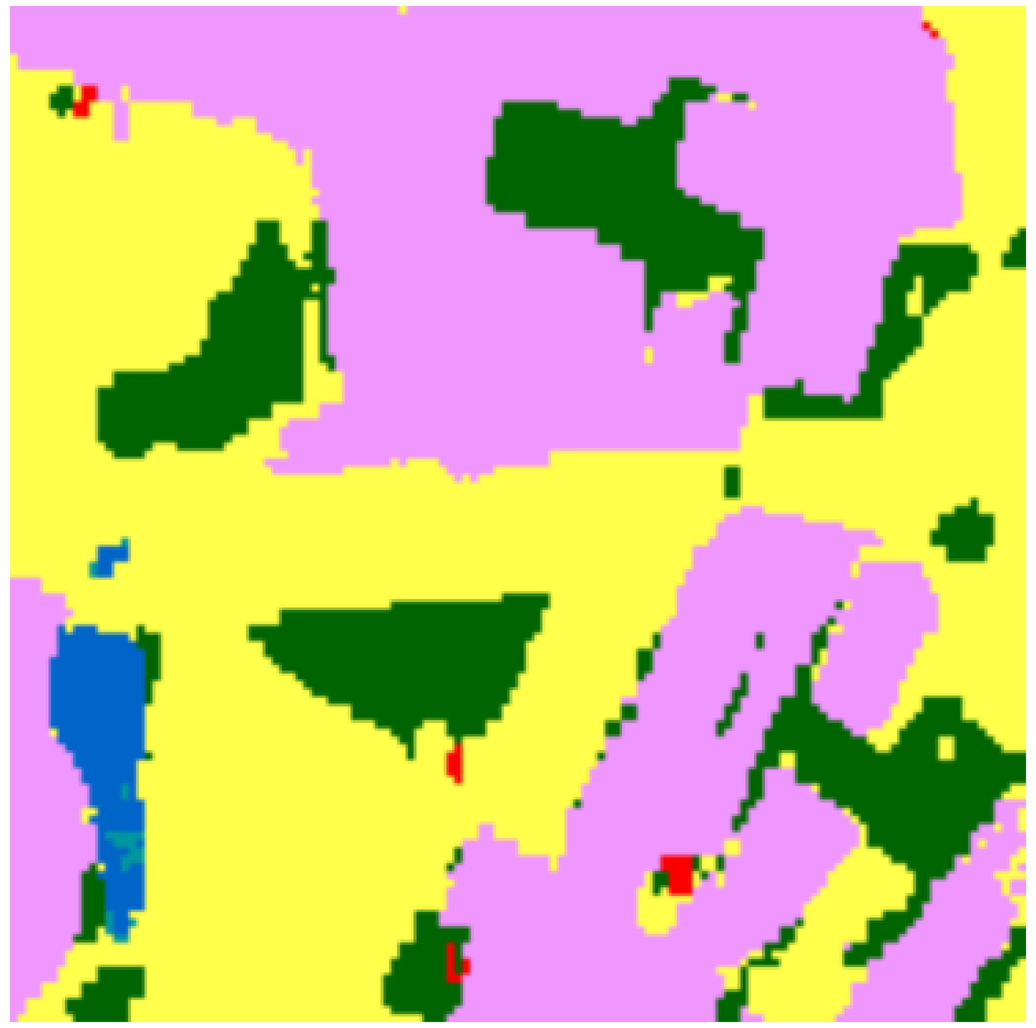
The proposed Evaluation Framework PhilEO

- **Self-supervised** learning and Foundation Model approaches
  - Reduce label requirements
    - ~20% of the labels otherwise needed
  - Research in EO and remote sensing has *focused* on these approaches
- In recent years, *many* new models have been introduced
- However, despite all the advancements:
  - It has become increasingly difficult to standardize a *fair* comparison across the many different Foundation Models
- **The PhilEO Suite**
  - Complex deep neural network models are trained using unlabelled data
  - **PhilEO Bench: *Evaluation***
    - On diverse downstream tasks

Evaluating and benchmarking Foundation Models



Left: Sentinel-2 data: Visualization in RGB (3 bands). Right: *Prediction* by the model



Ground truth **labels**: Classes like *Cropland*

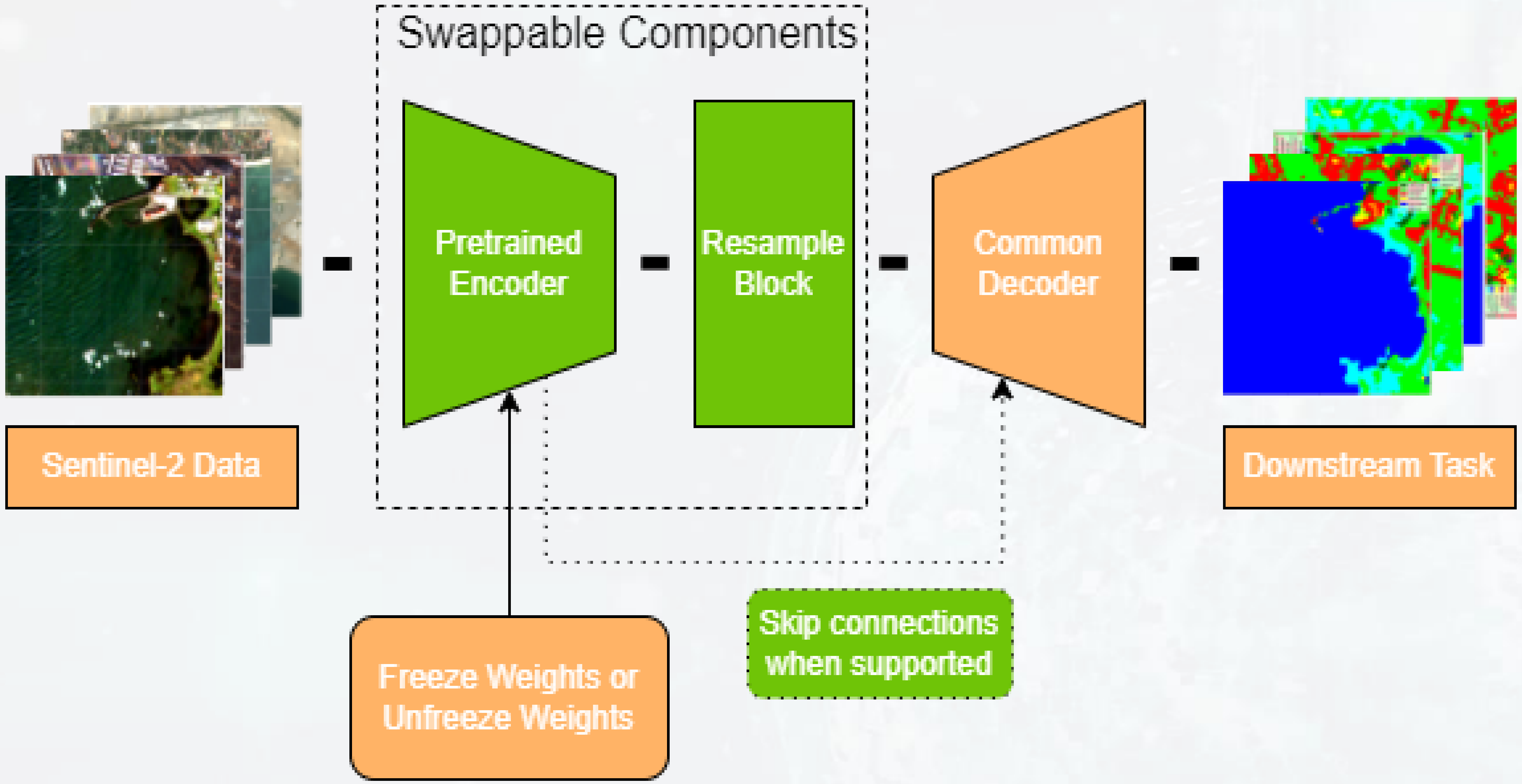
- RGB colours defined by ESA WorldCover

New dataset and framework

- PhilEO Bench: There is a need to evaluate different EO Foundation Models on a fair and uniform benchmark
  - **Testbed**
  - Novel 400GB global dataset of Sentinel-2 data with *labels* for 3 downstream tasks:
    - Land cover classification
    - Building density estimation
    - Road segmentation

- For the same **Sentinel-2 image** (L2A, 10 spectral bands):
  - ***Semantic* segmentation** land cover classification downstream task
    - Based on the **labelled dataset ESA WorldCover: 11 classes**
  - Estimation of how dense and *close* to each other buildings are: Regression task
  - **Road *regression* segmentation**

The proposed PhilEO Bench Evaluation Framework



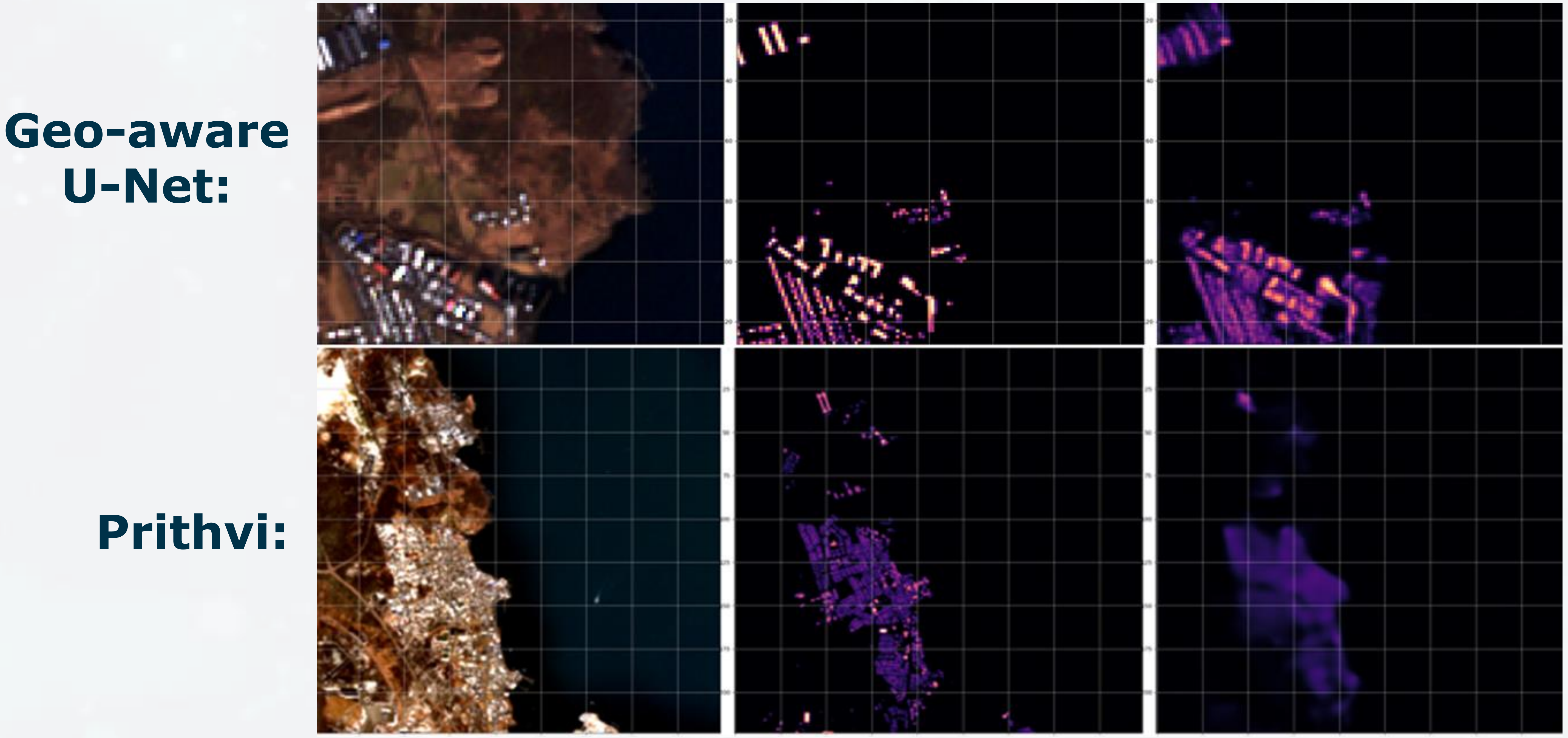
Minimizing the impact of confounding variables

- *One head* to rule them all
  - A common decoder head is used for all downstream task models
  - The performance of a model is a consequence of the effectiveness of the pre-training task and the representational strength of its latent space

- **Training configurations**
  - Fine-tuning
    - *All* model weights are updated during training including the FM encoder
  - Linear-probing
    - Only the decoder head weights are updated during training, *freezing* the FM encoder parameters

- **Evaluation:** Different regions: Stratification
- *n-shot*: Different dataset sizes

Qualitative evaluation



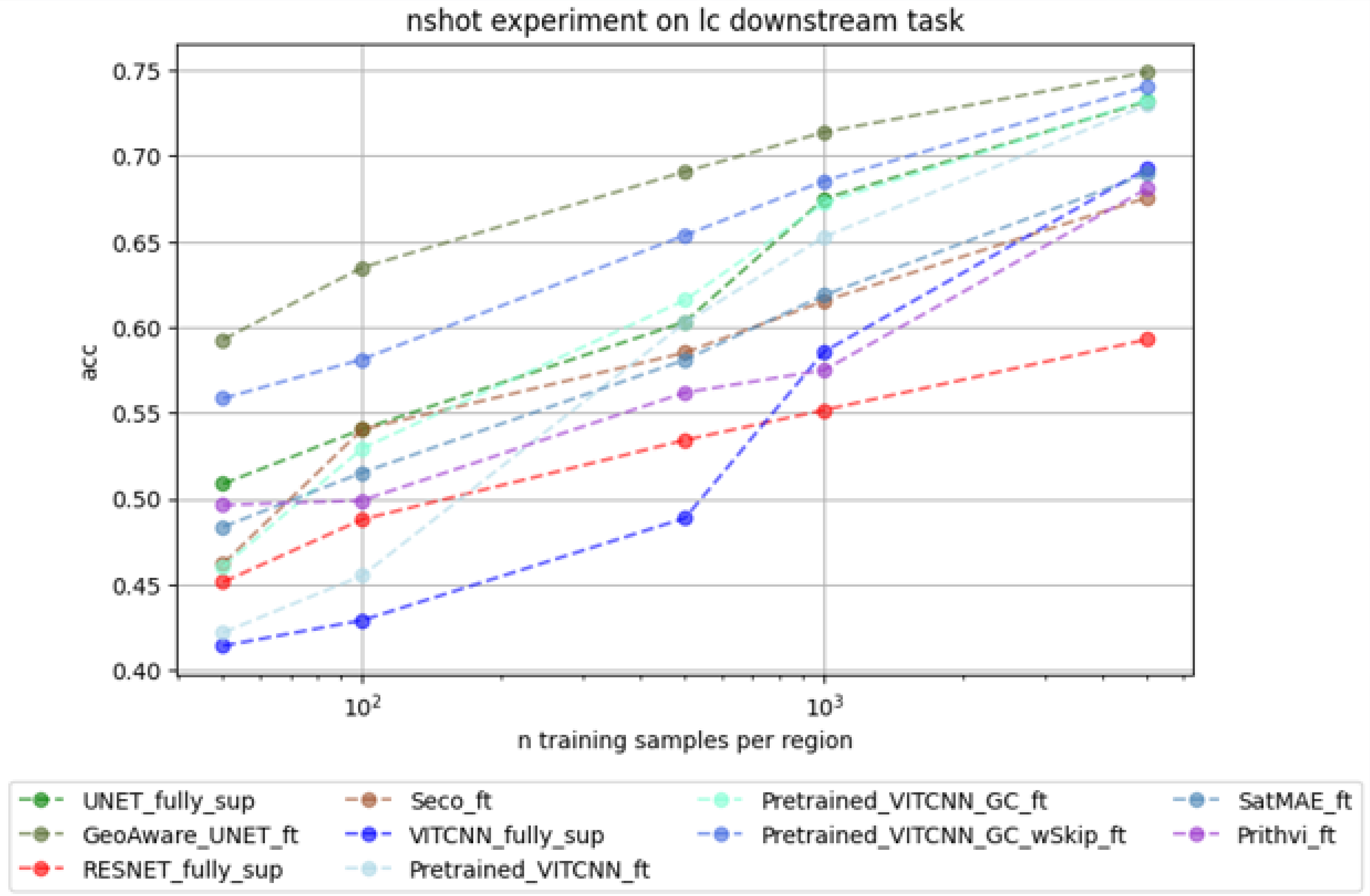
Building density estimation

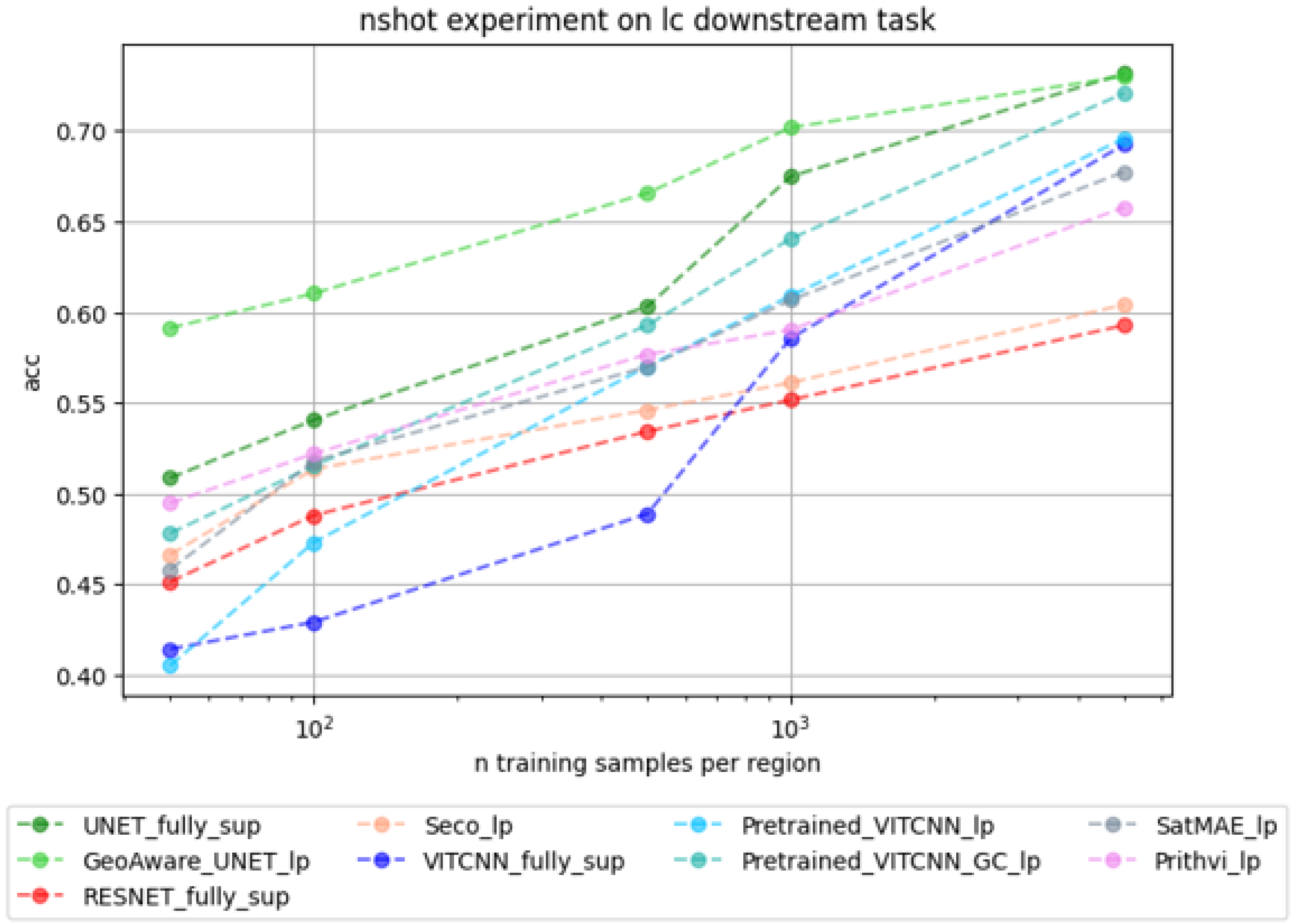
Output images

- Building density *regression task*
- Geo-aware U-Net
- Left: Sentinel-2 input. Middle: Ground truth. Right: *Prediction*
- Comparison with Prithvi

→ THE EUROPEAN SPACE AGENCY

Numerical evaluation: Semantic segmentation

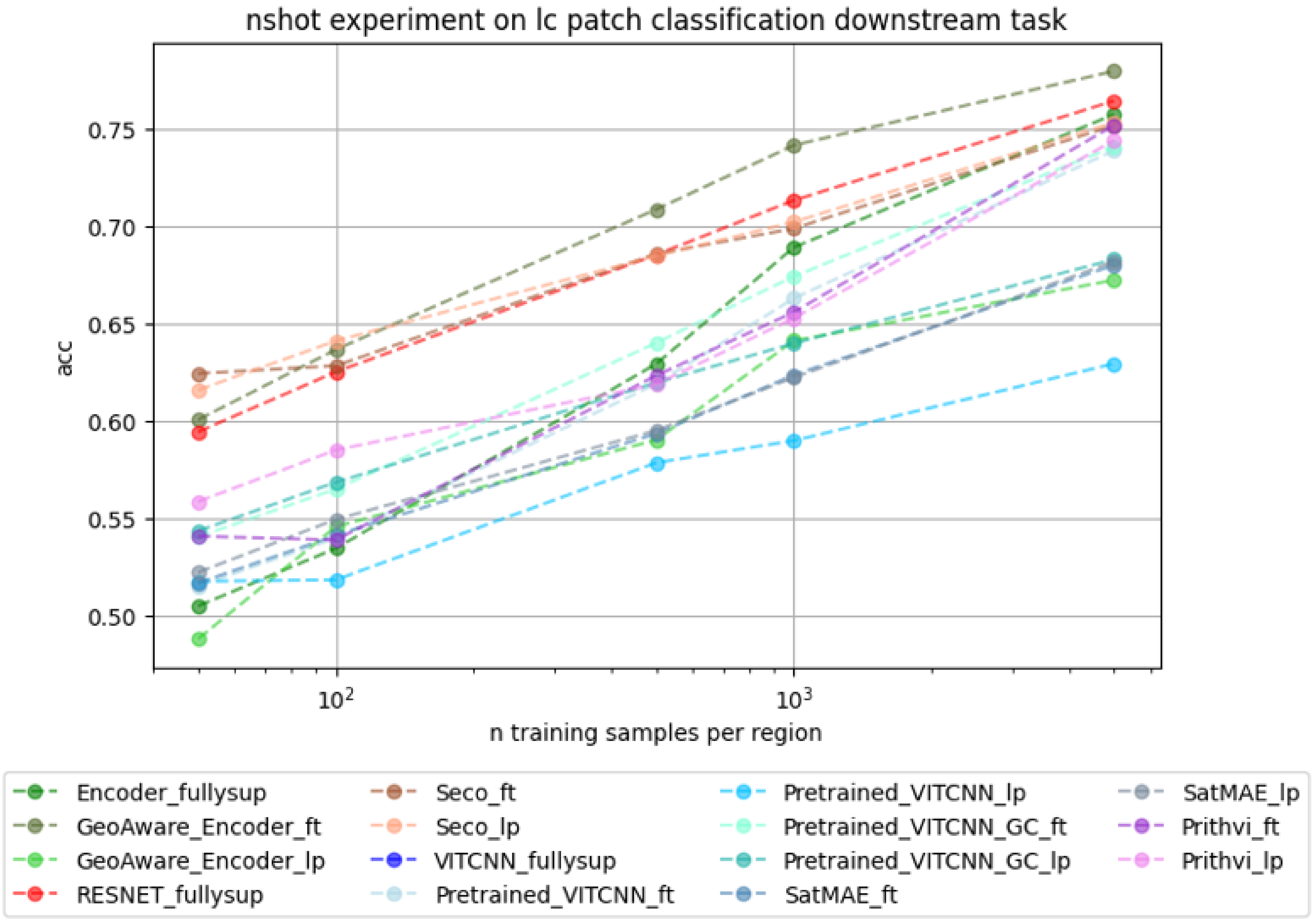




Land cover classification (lc):

- 11 classes, WorldCover
- Left: *Fine-tuning* (ft). Right: Linear probing (lp)
- Accuracy evaluation metric

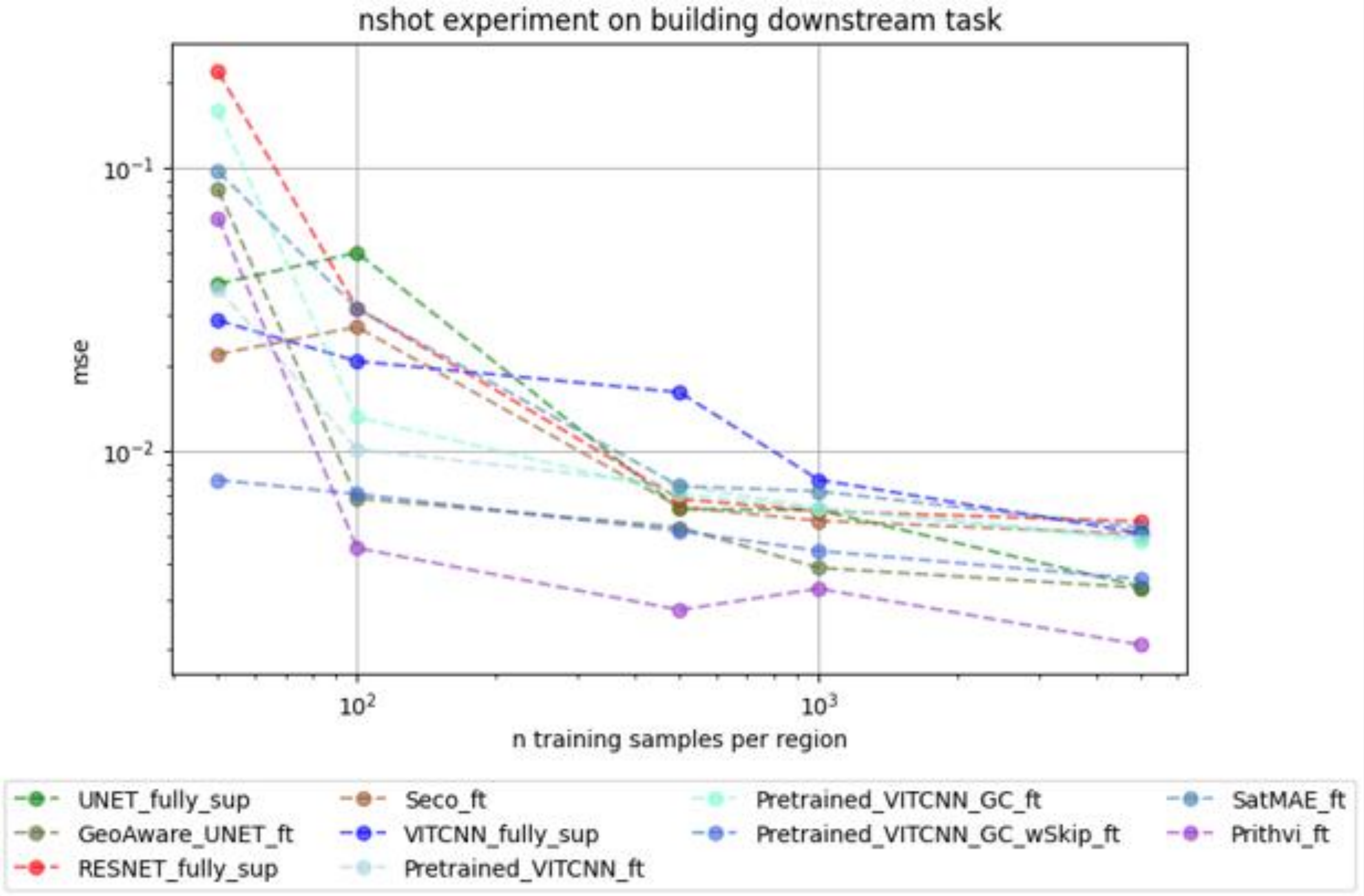
Patch classification land cover

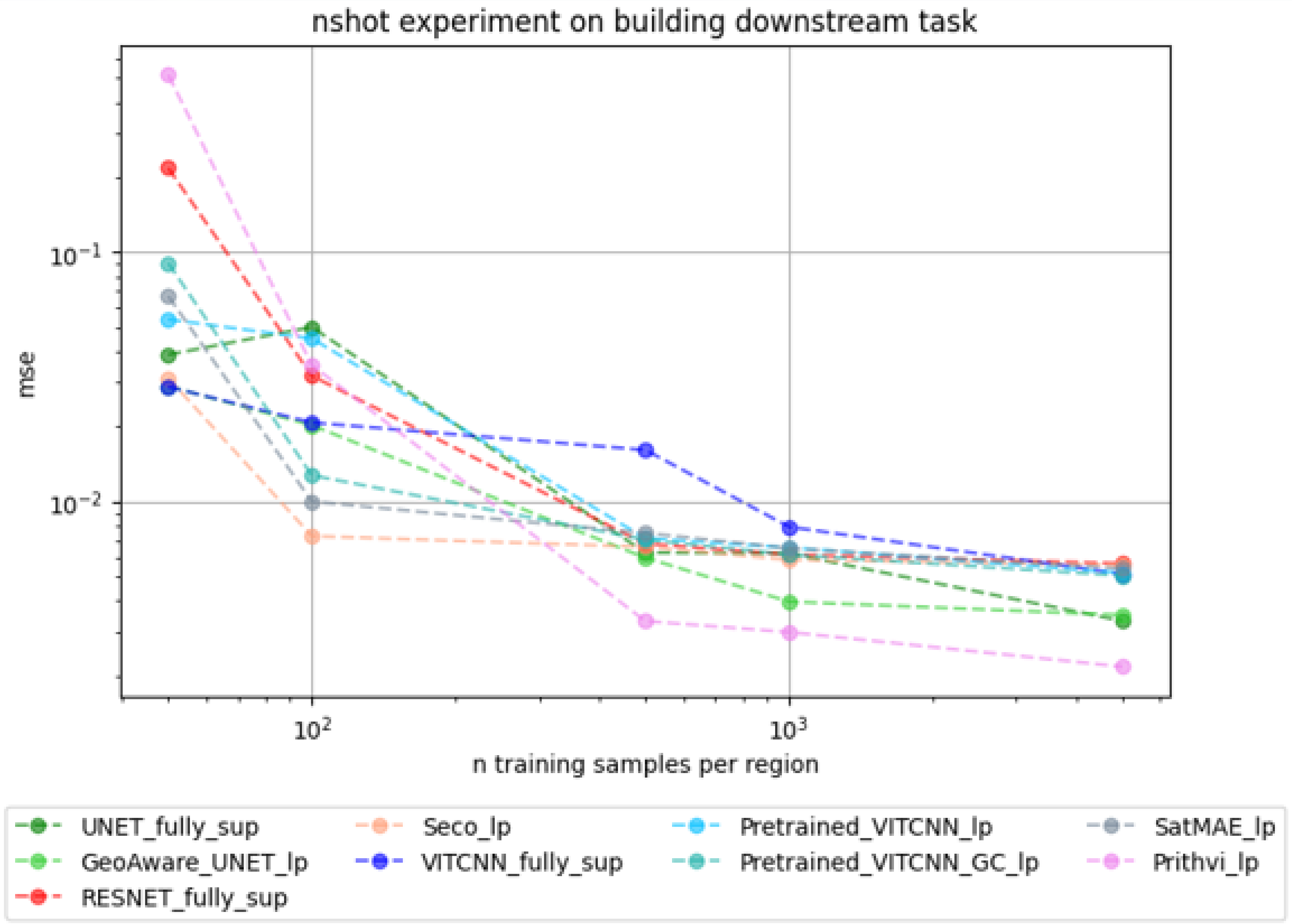


Labels at the image level

- The *majority class* in the image
  - Rather than semantic segmentation

Building density estimation downstream task

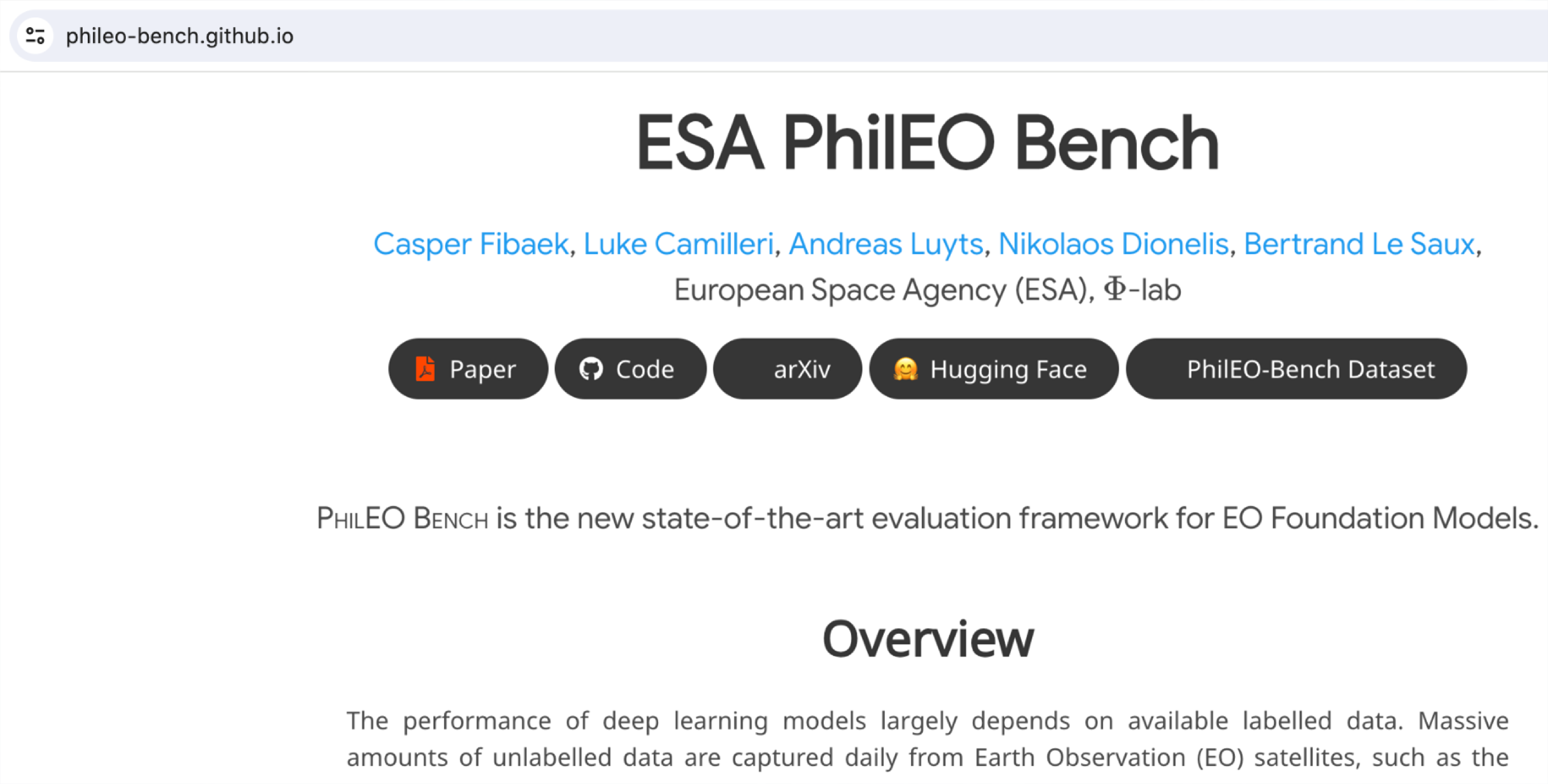




Building density *regression* task

- Evaluation metric: Mean Squared Error (MSE)
- **Left:** Fine-tuning (ft). **Right:** Linear probing (lp)
- Geo-aware U-Net

Conclusion


phileo-bench.github.io

ESA PhilEO Bench

Casper Fibaek, Luke Camilleri, Andreas Luyts, Nikolaos Dionelis, Bertrand Le Saux,
European Space Agency (ESA), Φ-lab

Paper | Code | arXiv | Hugging Face | PhilEO-Bench Dataset

PHILEO BENCH is the new state-of-the-art evaluation framework for EO Foundation Models.

Overview

The performance of deep learning models largely depends on available labelled data. Massive amounts of unlabelled data are captured daily from Earth Observation (EO) satellites, such as the

- Casper Fibaek, Luke Camilleri, Andreas Luyts, **Nikolaos Dionelis**, and Bertrand Le Saux, **“*PhilEO Bench: Evaluating Geo-Spatial Foundation Models*,” IEEE IGARSS, 2024**
- Bertrand Le Saux, Casper Fibaek, Luke Camilleri, Andreas Luyts, **Nikolaos Dionelis**, Giacomo Cascarano, Leonardo Bagaglini, and Giorgio Pasquali, **“*The PhilEO Geospatial Foundation Model Suite*,” European Geosciences Union (EGU), 2024**

PhilEO Bench

- Evaluation framework
- Testbed and new *global* labelled 400GB dataset
- **3 downstream tasks**
  - Land cover classification semantic segmentation
  - Building density estimation
  - Road regression segmentation
- GitHub: http://github.com/ESA-PhiLab/PhilEO-Bench
- Landing page: http://phileo-bench.github.io/
- Geo-aware U-Net
- ViT
- Comparison with Prithvi, SatMAE, SeCo

Expandable Datasets for Earth Observation

Φ-lab: Alistair Francis, Mikolaj Czerkawski

MajorTOM

- **Big models need big data…**
- AI companies compete by filtering the best and biggest datasets from internet data

Model	Training Data
GPT-3	500 billion tokens
GPT-4	13 trillion tokens
DALL-E	12 million image/text pairs
DALL-E 2	650 million image/text pairs
StableDiffusion1.1	170 million image/text pairs

- Like the internet, EO has vast amounts of public data too

What **SHOULD** we train big EO models on?

- **Huge data** (terabytes)
- Globally distributed – low bias
- Openly accessible data

What **DO** we train big EO models on?

- **Limited data** (gigabytes)
- **Geographically biased**
- **Mix of open and closed data**

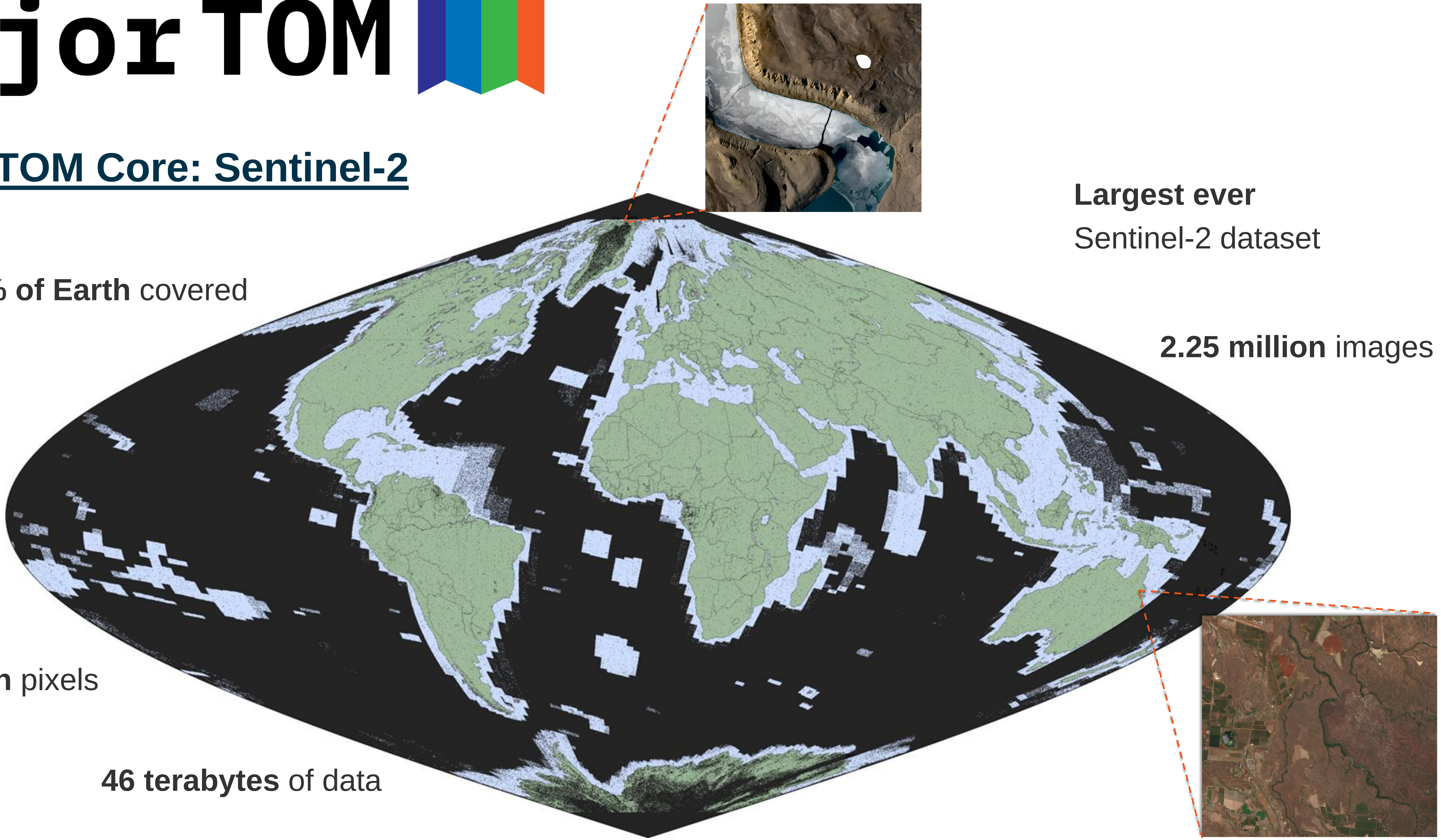
Major TOM

- **Major TOM: Terrestrial Observation Metaset**
- **Framework** to build **largest ever EO datasets** for AI.
- **Simple, repeatable format:** combine Major TOM datasets together easily
- **Distributed freely:** partnership with **Hugging Face** to deliver data to anyone, anywhere
- **Collaborative project:** expandable and managed by **open-source community**

Major TOM's grid system. Each grid point gets a sample of data. 200km grid visualised, real data in 10km grid.

MajorTOM

Major TOM Core: Sentinel-2



Largest ever
Sentinel-2 dataset

50% of Earth covered

2.25 million images

2.5 trillion pixels

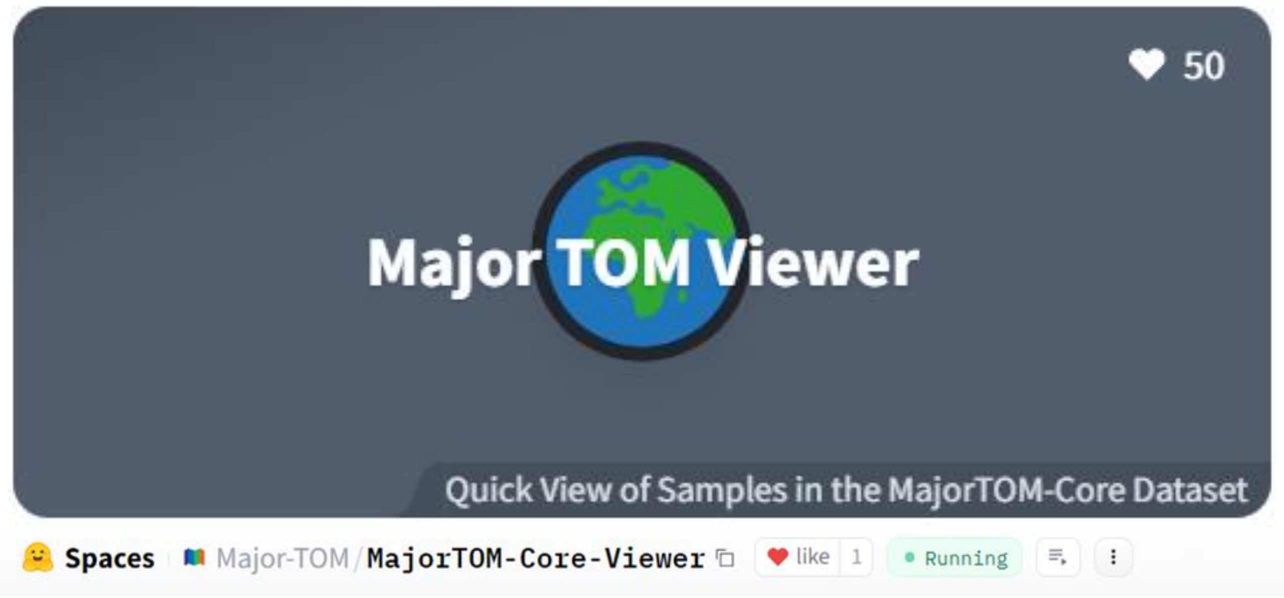
46 terabytes of data

Major TOM

Since recent release:

- Major TOM is now a **trending dataset** on Hugging Face
- The online viewer app is currently featured as a **HF space of the week**
- The **community** organisation on HF **is growing** rapidly with an influx of new members
- **Setting foundations for truly open EO data...**

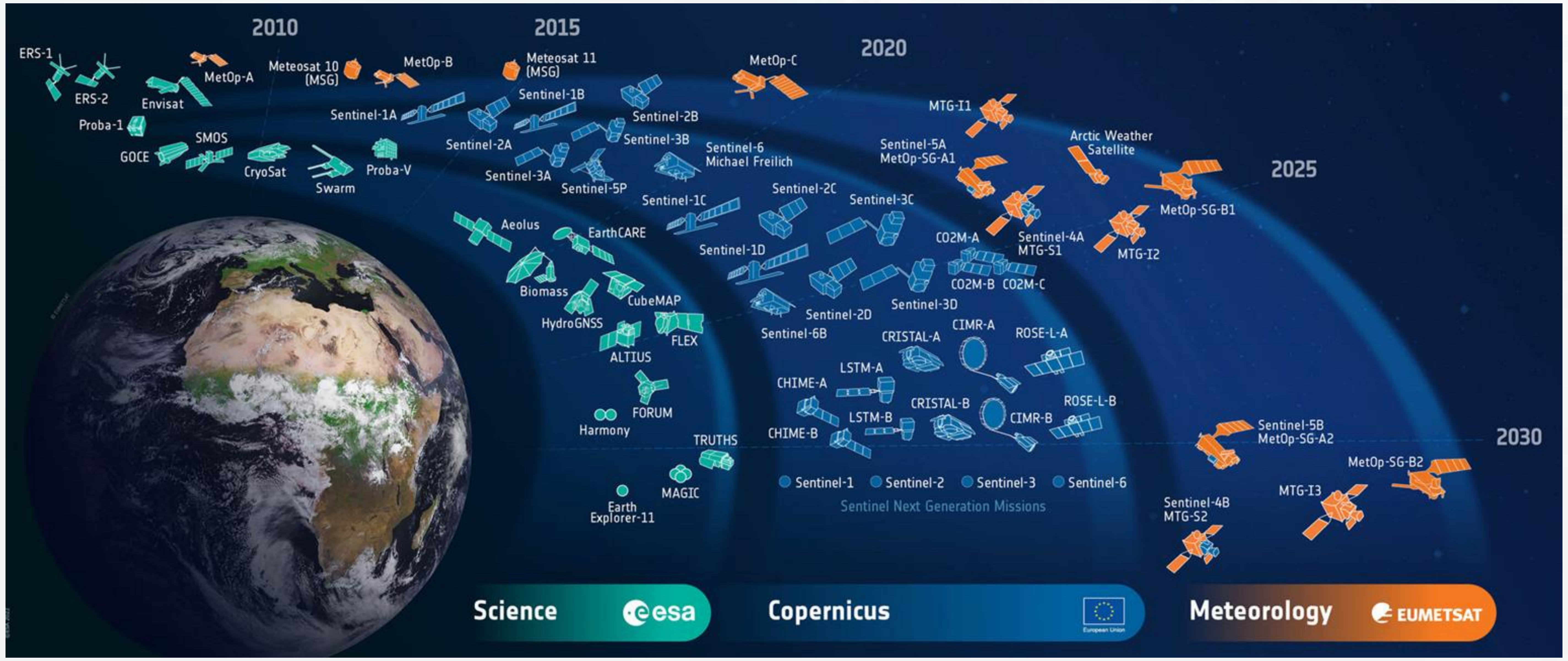
Explore data in our web app:



Learning from Unlabelled Data with Transformers: Domain Adaptation for Semantic Segmentation of High Resolution Aerial Images

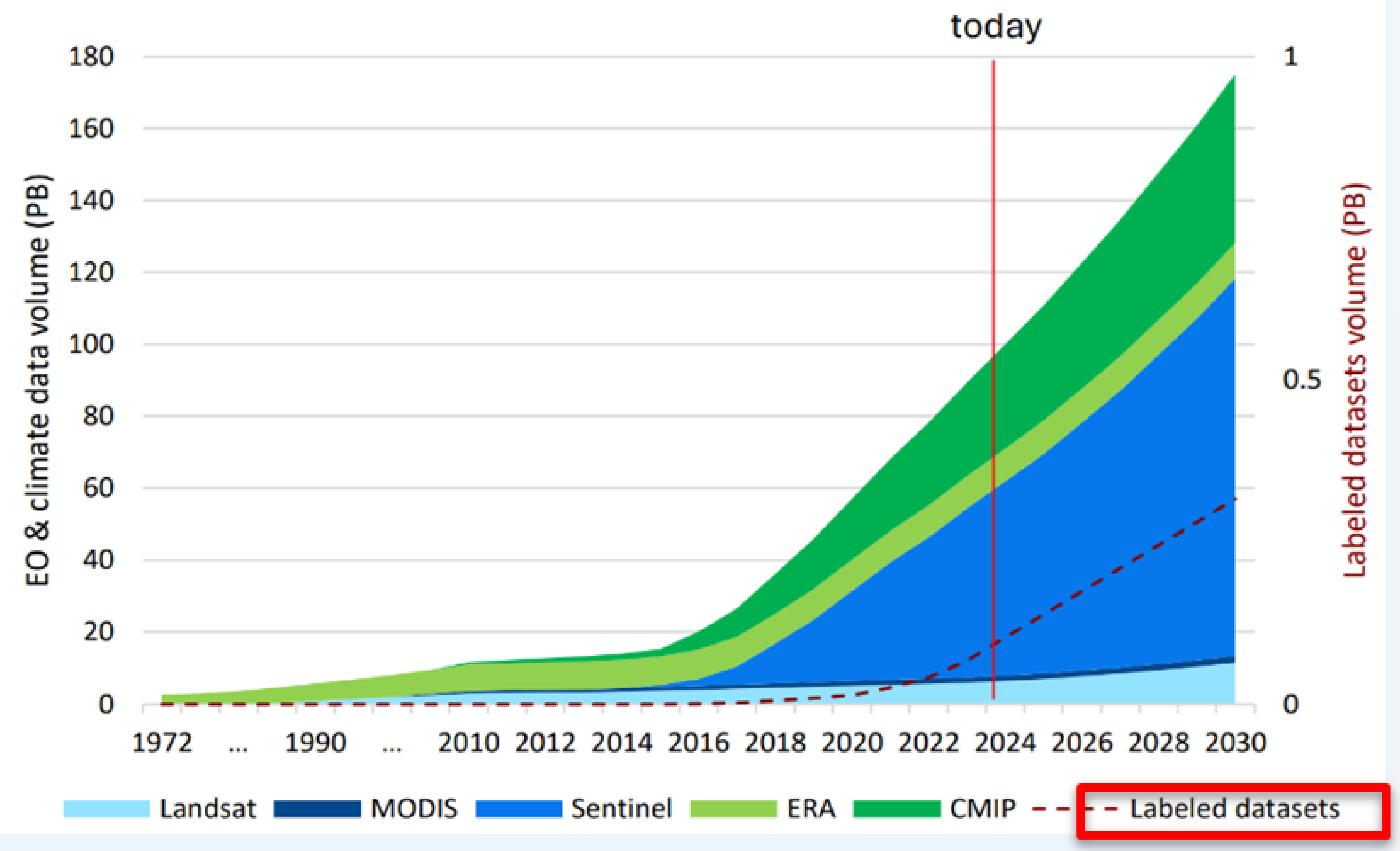
Nikolaos Dionelis[1], Francesco Pro[2], Luca Maiano[2], Irene Amerini[2], Bertrand Le Saux[1]
1 European Space Agency (ESA), Φ-lab
2 La Sapienza University of Rome, Italy

Introduction



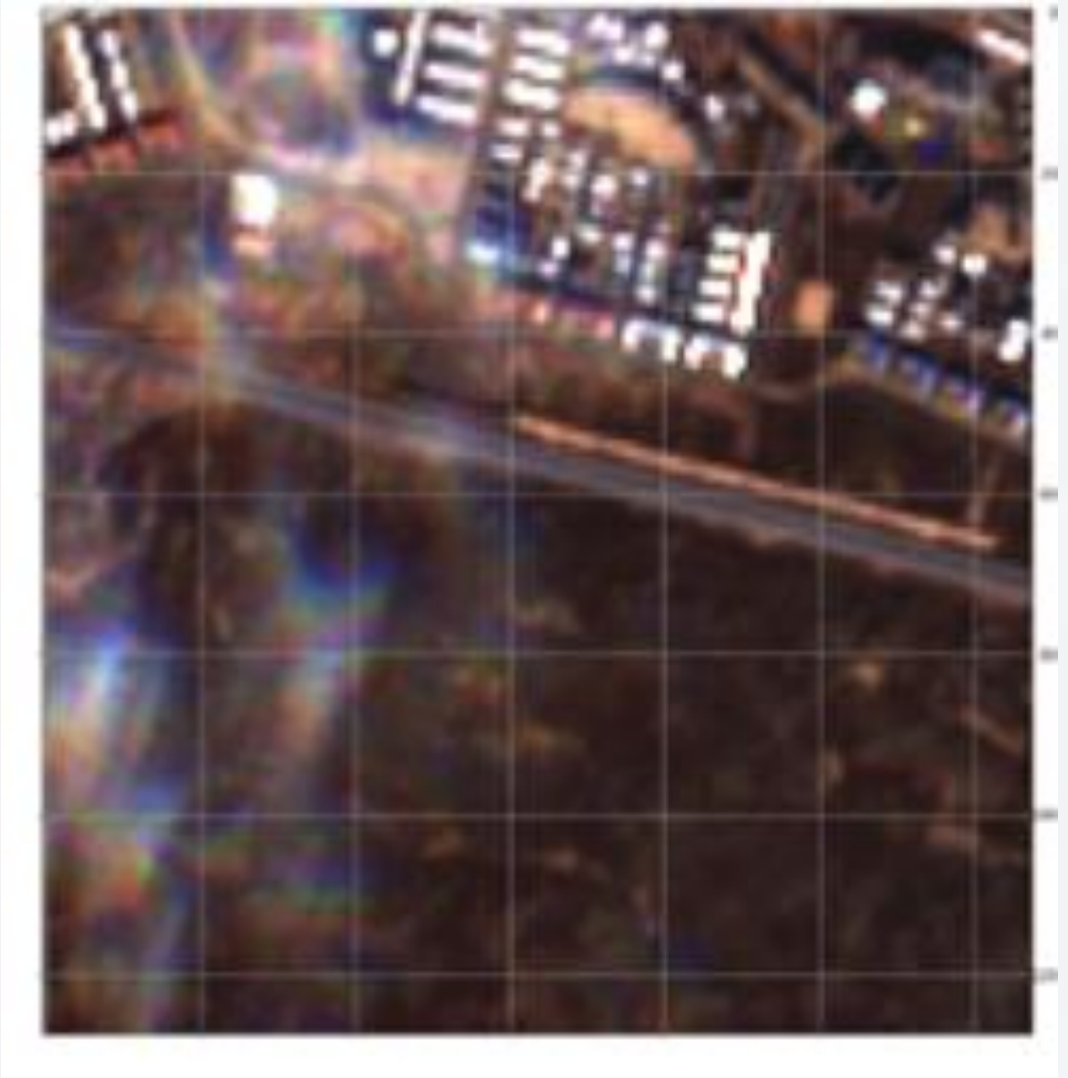
Importance and Motivation

- Data from *satellites* or aerial vehicles are most of the time unlabelled
    - Earth Observation (EO) satellites capture large amounts of *unlabelled* data
- Learning from unlabelled data is challenging
- **Lack** of annotated data
    - Labels: Need *time* & are expensive



http://arxiv.org/pdf/2405.04285

Domain adaptation for semantic segmentation of unlabelled data



Learning from unlabelled data

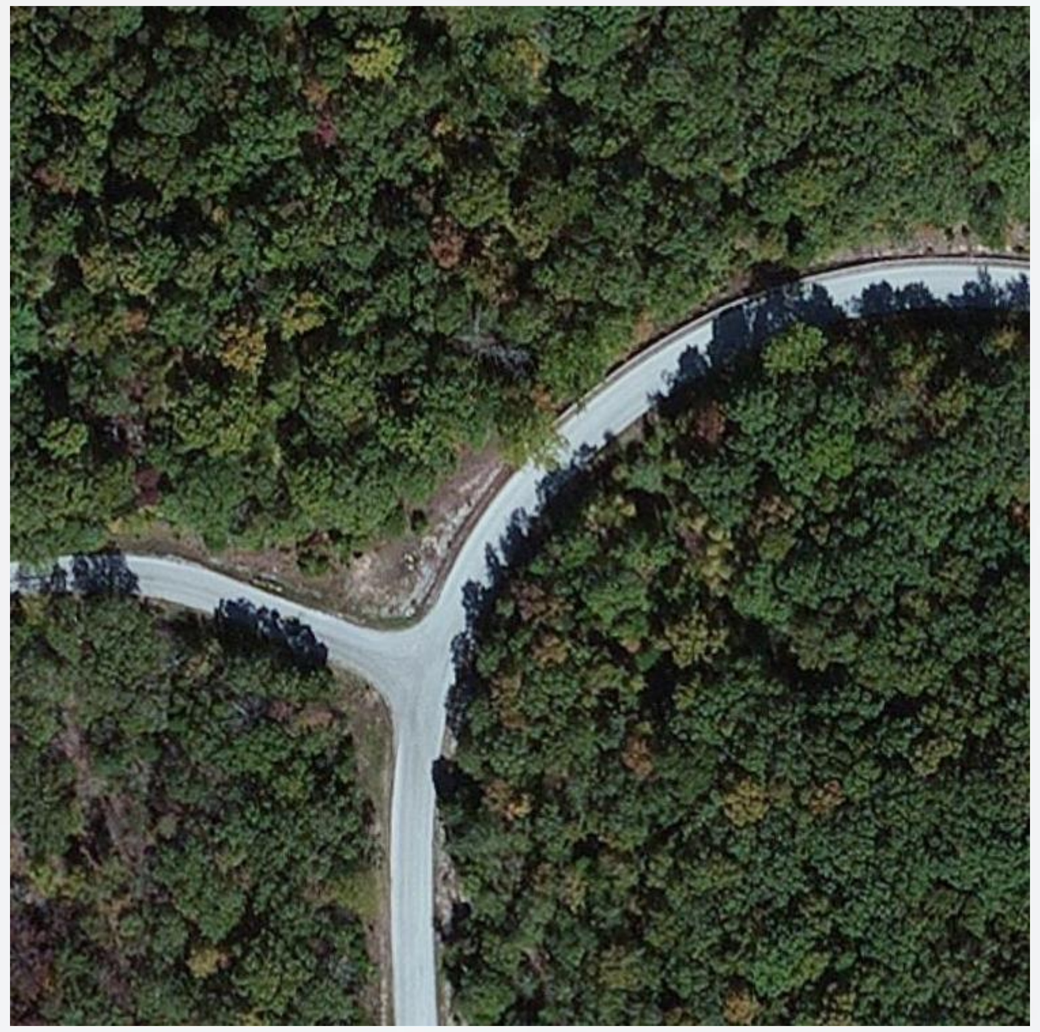
- For *semantic* segmentation
  - Within a semi-supervised learning framework

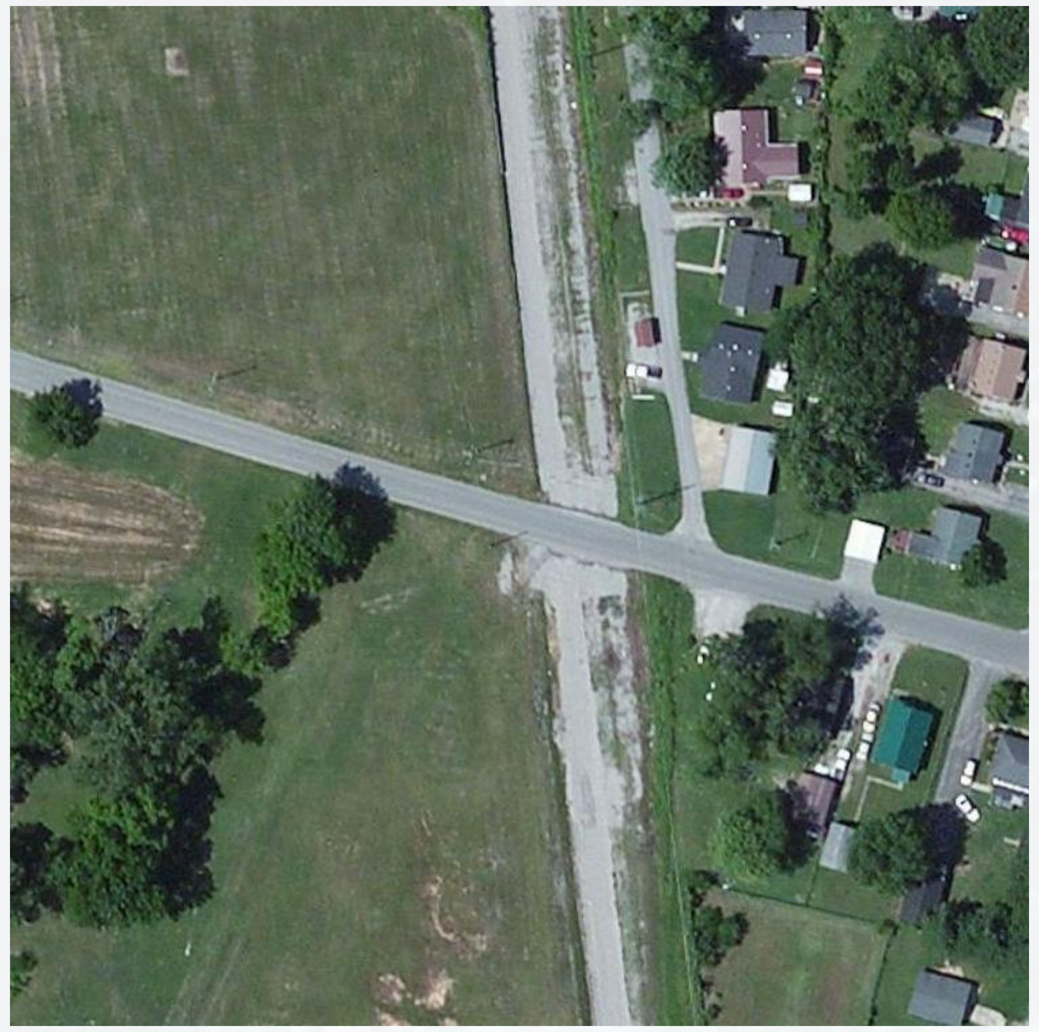
- Example: 1) *Segments*: The segmentation problem
  - Like instance segmentation
- 2) **Labels:** The classification problem
  - The labelling problem
  - Like semantic segmentation
- 3) Example: For *Cropland*: Crop type & yield
  - If segment wrong: Crop yield incorrect & *not* useful

- Labelling data *accurately*: Requires expertise
- Even if EO data were correctly labelled: Labels **change** over time

- Model for semantic segmentation of *unlabelled* data
  - The proposed Non-annotated Earth Observation Semantic Segmentation (NEOS) model

Learning from unlabelled data for semantic segmentation





Example images from the dataset CVUSA: *No annotations*

Objectives

- Importance: *Supervised* learning has shown good performance for classification & segmentation
- **However:** High-quality large labelled datasets

- Because of several satellites and aerial images and their *non-annotated* samples, it is challenging to use these data

The more general problem we want to solve

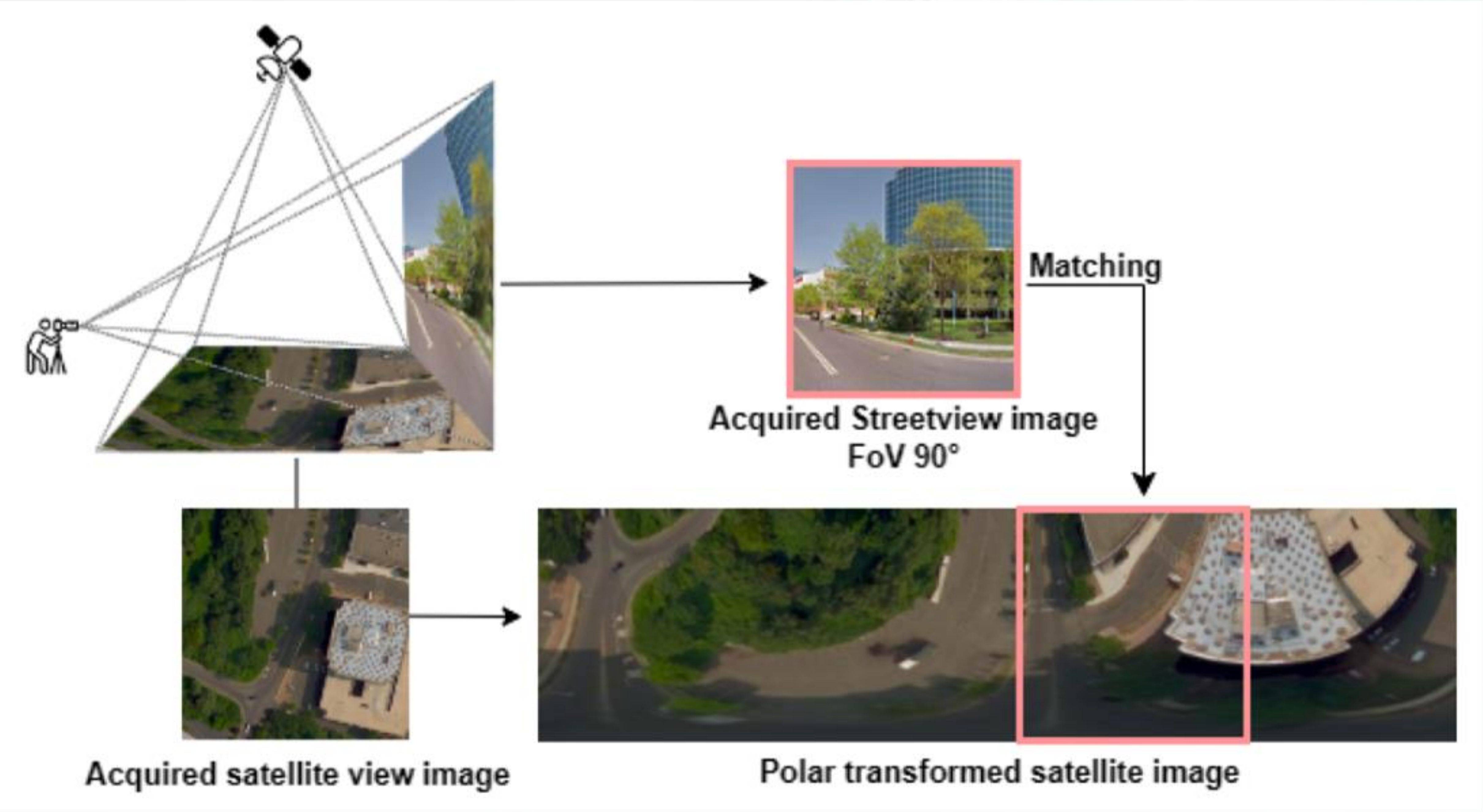
- General methodology
- Labelled and *unlabelled* data
- Semantic segmentation: Segments and labels

- Domain adaptation
  - Source domain & *Target* domain

- Our model **NEOS**:
  - Performs domain adaptation as the target domain does *not* have ground truth masks

Domain adaptation: Methodologies for non-annotated data

Non-annotated image data samples: *No* land cover labels, i.e. **semantic** segmentation masks



Examined *use case*: Matching of street-view and aerial images: Geo-localization

Using *Transformer*-based models

- A Transformer-based model, *SegFormer* for semantic segmentation [1]
  - A recent powerful architecture

- **Aim:** Accurate semantic segmentation
- Target domain: *Unlabelled* data
  - Source domain: Similar labelled data

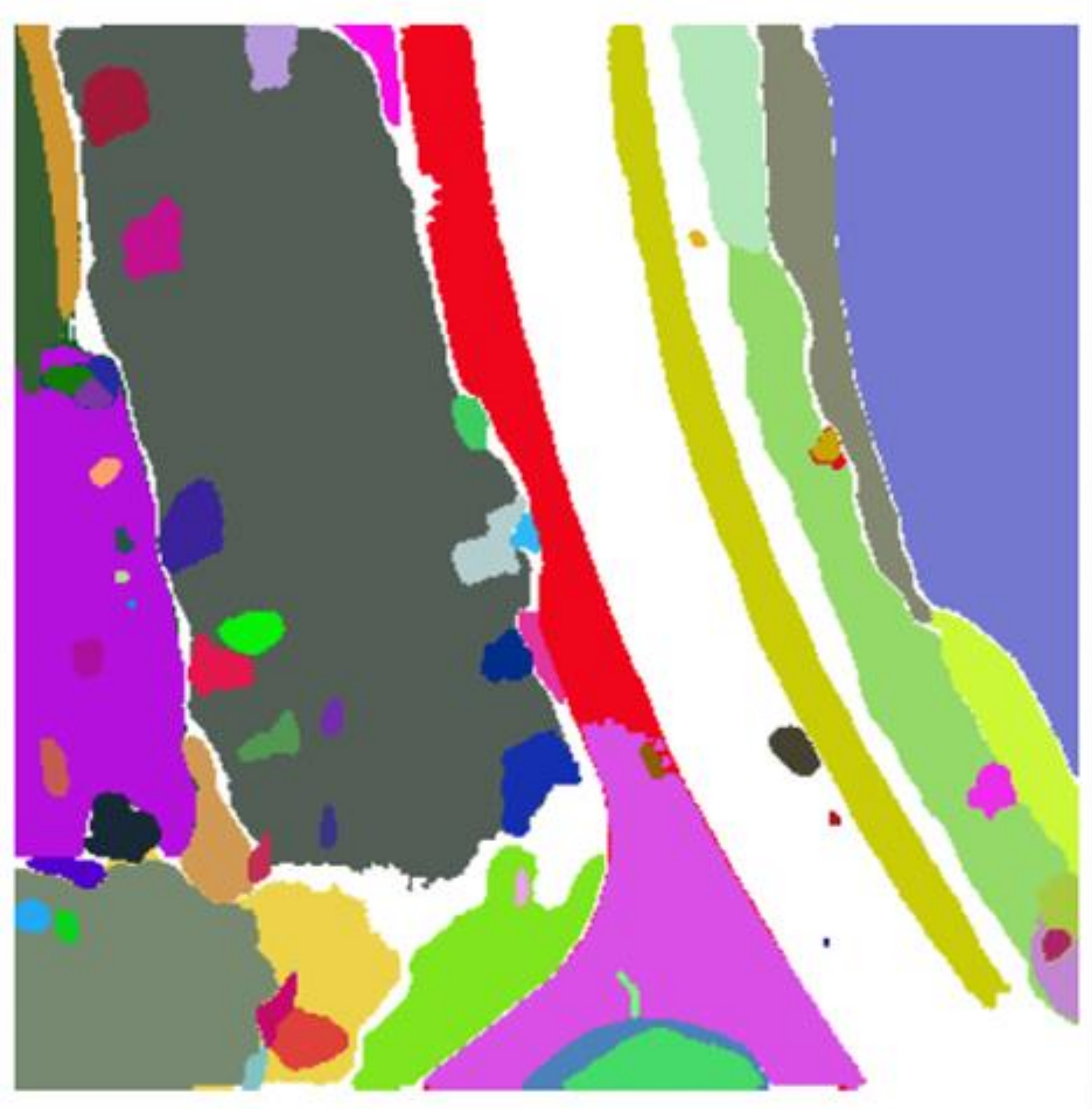
- The proposed model NEOS
  - *Aligns* the representations of the different domains to make them coincide

- **Application/ use case:** Cross-view matching
- *Semantic* segmentation masks: Additional information to **improve** performance

[1] E. Xie, et al., "*SegFormer: Simple and efficient design for semantic segmentation with Transformers*," NeurIPS, 2021
[2] F. Pro, N. Dionelis, et al., "*A semantic segmentation-guided approach for ground-to-aerial image matching*" IGARSS, 2024

Related Work

Unsupervised Domain Adaptation (UDA)



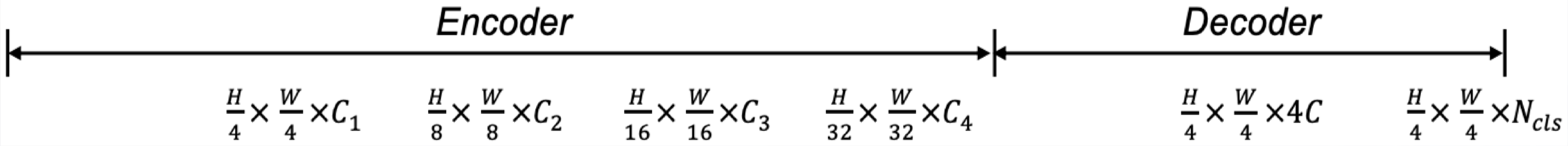
Left to right: Input, Instance segmentation by *SAM*

- Models trained on one domain might **not generalize** well on other domains
- Domain adaptation
  - Region change: *Shift* between the distributions in source & target domains
- Source and target domains with labelled and *unlabelled* data, respectively
- Aerial CVUSA dataset
  - No semantic segmentation model transferred well/ *accurately* in [3]
- **Architectures** for semantic segmentation:
  - Transformer: **SegFormer**
    - Multi-scale features/ representations
  - Encoder-decoder, U-Net: SegNet
- Segment Anything Model (SAM)

[3] R. Rodrigues, et al., “Are these from the same place? Seeing the unseen in *cross-view* image geo-localization,” WACV, 2021

Proposed Model

Semantic segmentation of unlabelled data using NEOS



SegFormer framework

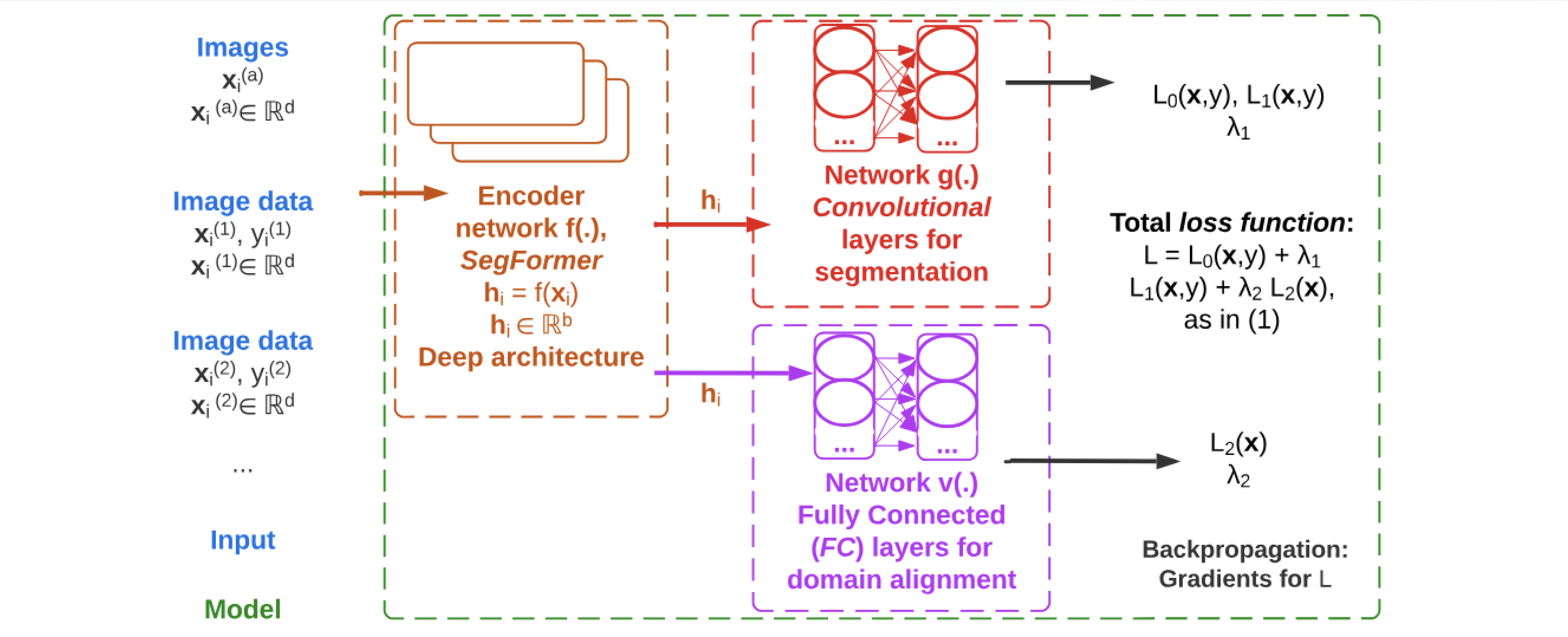


Fig. 1. Flowchart of NEOS for semantic segmentation using domain adaptation on datasets with no *ground truth* labels.

- The proposed Non-annotated EO Semantic Segmentation (**NEOS**) model
- NEOS: Based on *architecture* SegFormer B5
  - *Multi-level* feature map
- **Flowchart** diagram:
  - **Output:** *Semantic* segmentation mask
- **Second output head:** For features misalignment loss term - To perform *domain adaptation*
- NEOS *loss function*:
  - a) Cross-entropy loss
    - For *classification*
  - b) 1 - Dice score
    - For **segmentation**
  - c) Features *misalignment* loss
    - For domain adaptation

Unlabelled data domain adaptation - NEOS

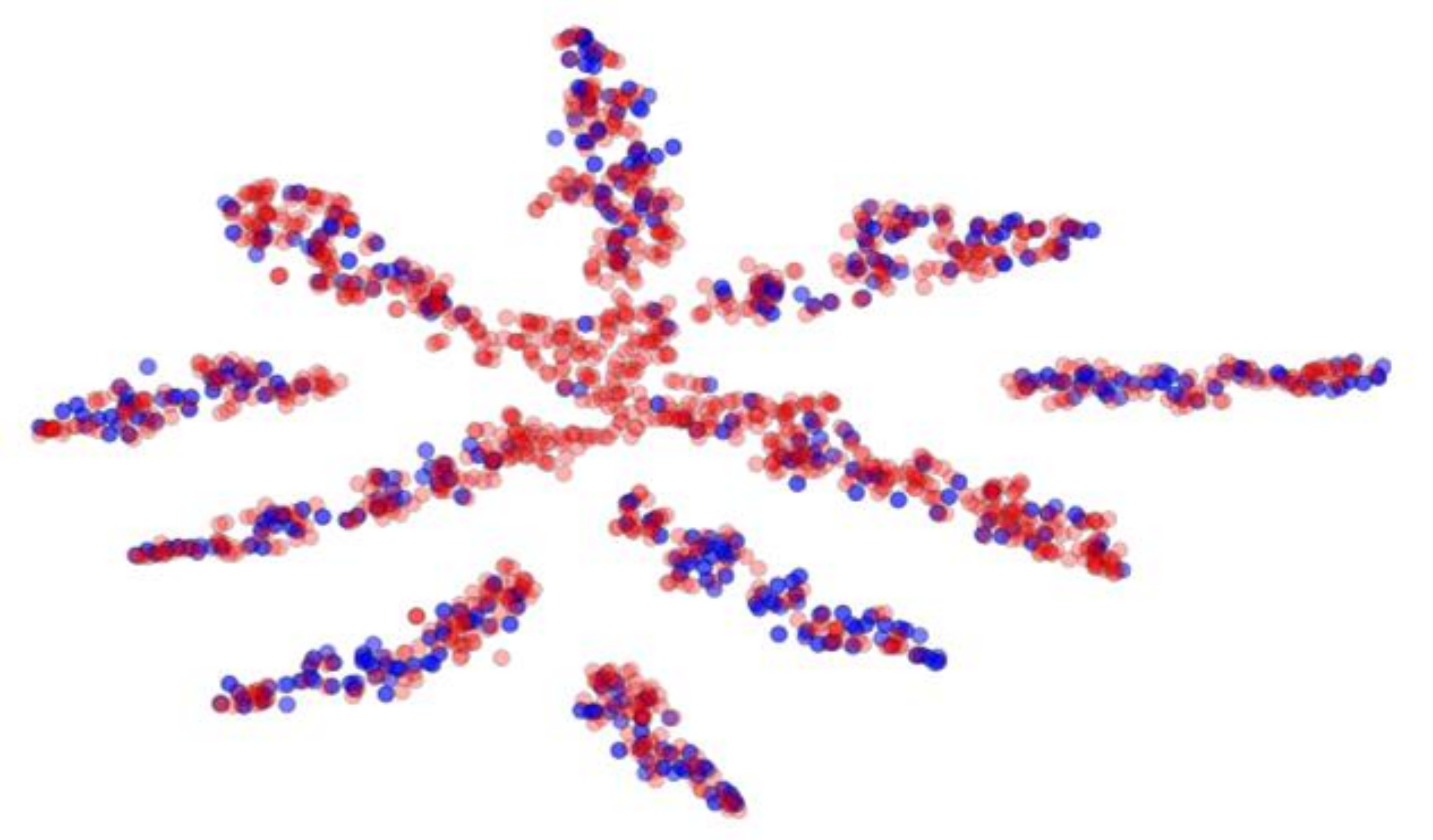
$$\operatorname{argmin}_f L = L_0(\mathbf{x},\mathbf{y}) + \lambda_1 L_1(\mathbf{x},\mathbf{y}) + \lambda_2 L_2(\mathbf{x}).$$

$$L_0 = -\frac{1}{NWH}\sum_{j=1}^{N}\sum_{i=1}^{W}\sum_{l=1}^{H}\log\frac{\exp(f_{y_{j,i,l}}(\mathbf{x}_j))}{\sum_{k=1}^{K}\exp(f_{k,i,l}(\mathbf{x}_j))}$$

$$1 - \frac{2\sum_{j=1}^{N}\sum_{i=1,l=1}^{W,H} g_{j,i,l}s_{j,i,l}}{\sum_{j=1}^{N}\sum_{i=1,l=1}^{W,H} g_{j,i,l} + \sum_{j=1}^{N}\sum_{i=1,l=1}^{W,H} s_{j,i,l}}$$

$$L_2 = \frac{1}{J}\sum_{j=1}^{J}\log\frac{\exp(f_{z_j}(\mathbf{x}_j))}{\sum_{m=1}^{M}\exp(f_m(\mathbf{x}_j))}$$

- **Minimization** of objective cost function
- NEOS: 3 loss terms
  - L_0 is: Cross-entropy loss
  - L_1 is: 1 - **Dice** score
  - L_2 is: Features misalignment loss
    - **Domain adaptation**
    - Features: Manifold *alignment* of the embeddings of different domains



Domain adaptation to make the *features* coincide

Evaluation and Results

NEOS model evaluation and main results

Numerical evaluation

Evaluation: Input images, *Estimated masks*

- Also: Without *ground truth* segmentation masks

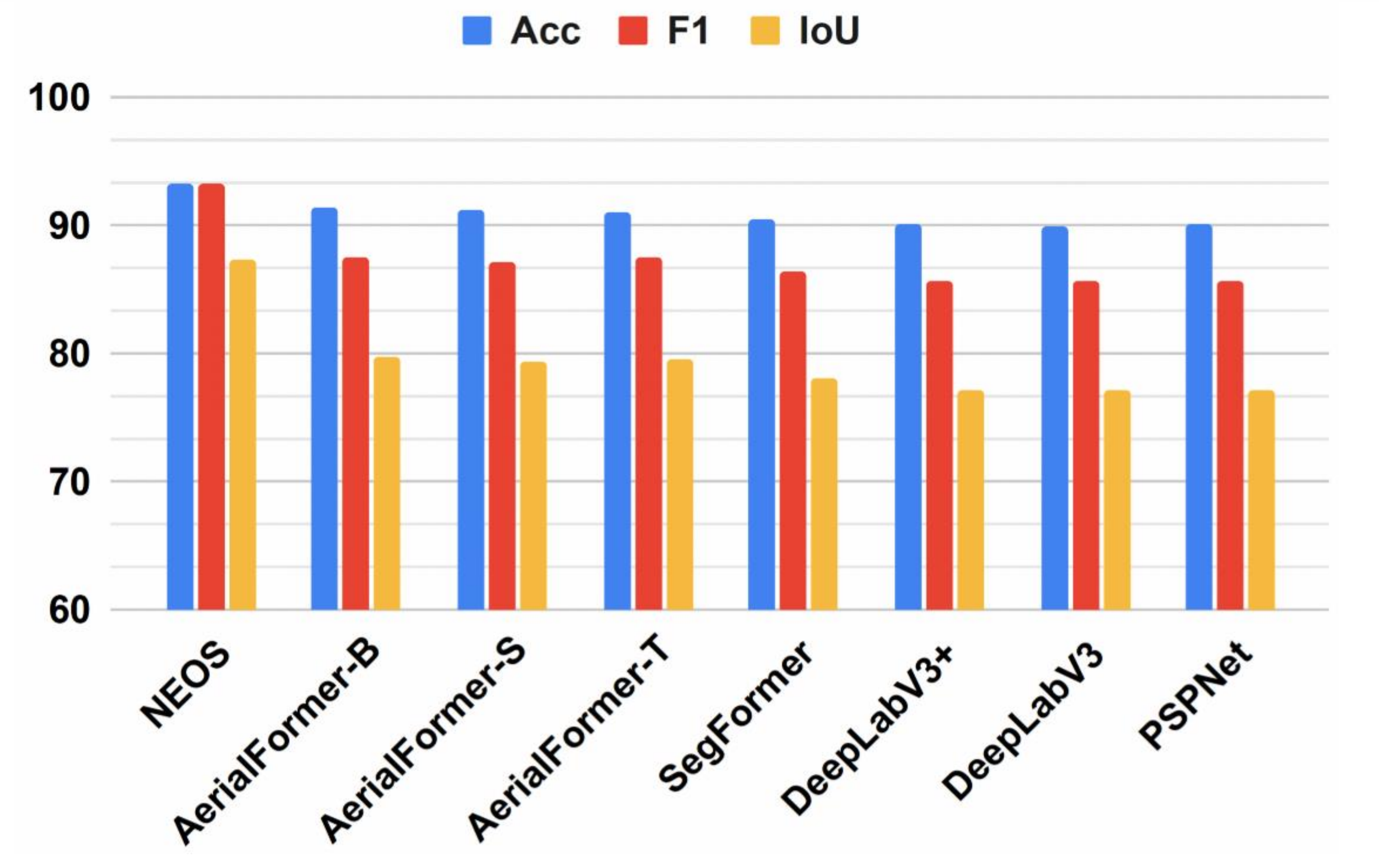


Fig. 2. Evaluation of NEOS in accuracy (Acc), F1-score (F1) and IoU on the dataset Potsdam with the class Clutter [12].

- **Evaluation** of the proposed model NEOS
- Aerial image datasets
  - Labelled:
    - Potsdam & Vaihingen
  - *Unlabelled*:
    - Top-view aerial CVUSA
    - **Domain adaptation**

- *Evaluate* NEOS on Potsdam and Vaihingen, as well as on CVUSA

- Our model outperforms *other* baseline models

- **Classes:**
  - Buildings (blue colour, image next slide)
  - Trees (green)
  - Cars (yellow)
  - Low vegetation (cyan)
  - Roads (white)
  - Clutter (red)

Semantic segmentation NEOS evaluation results

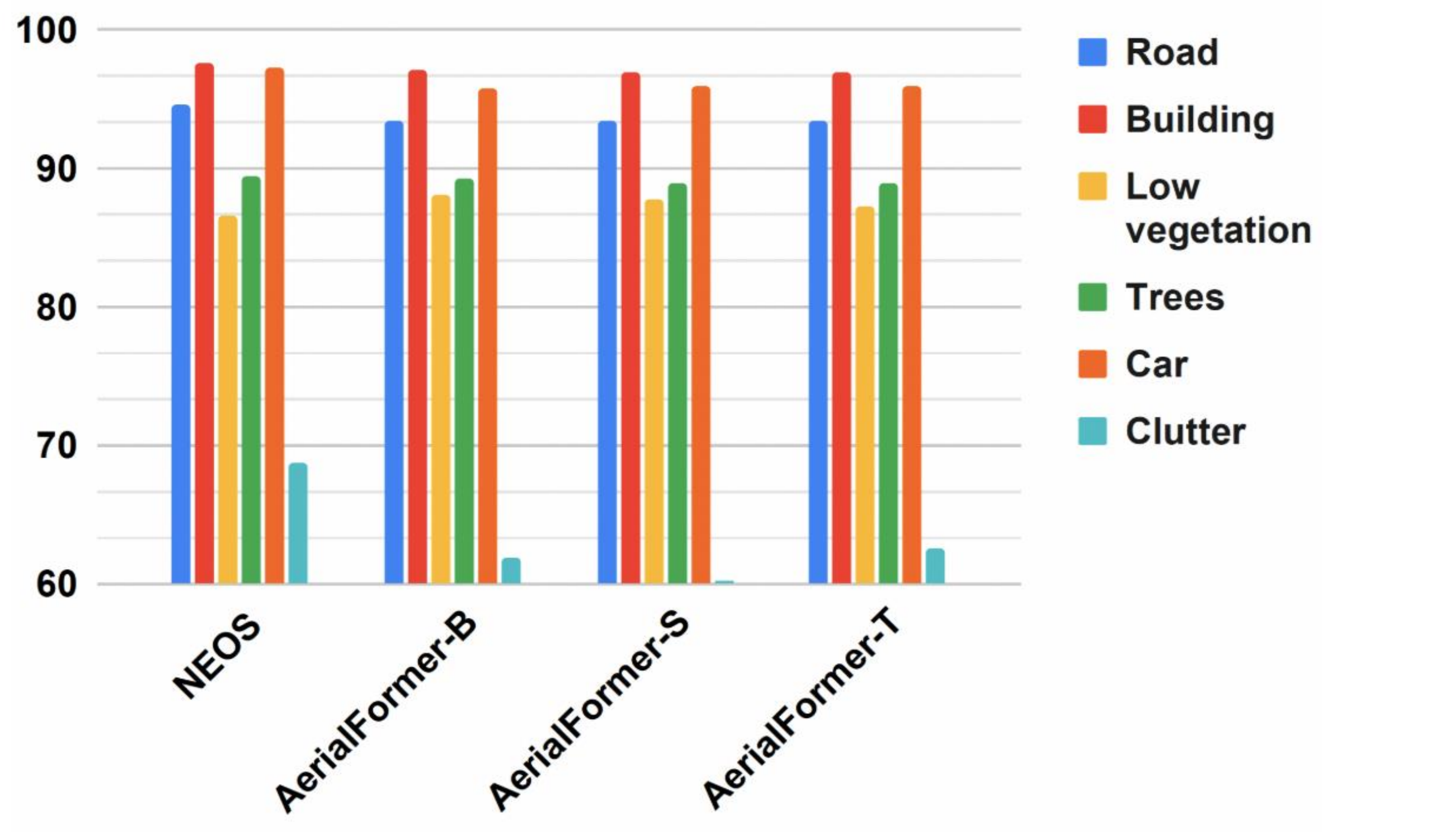
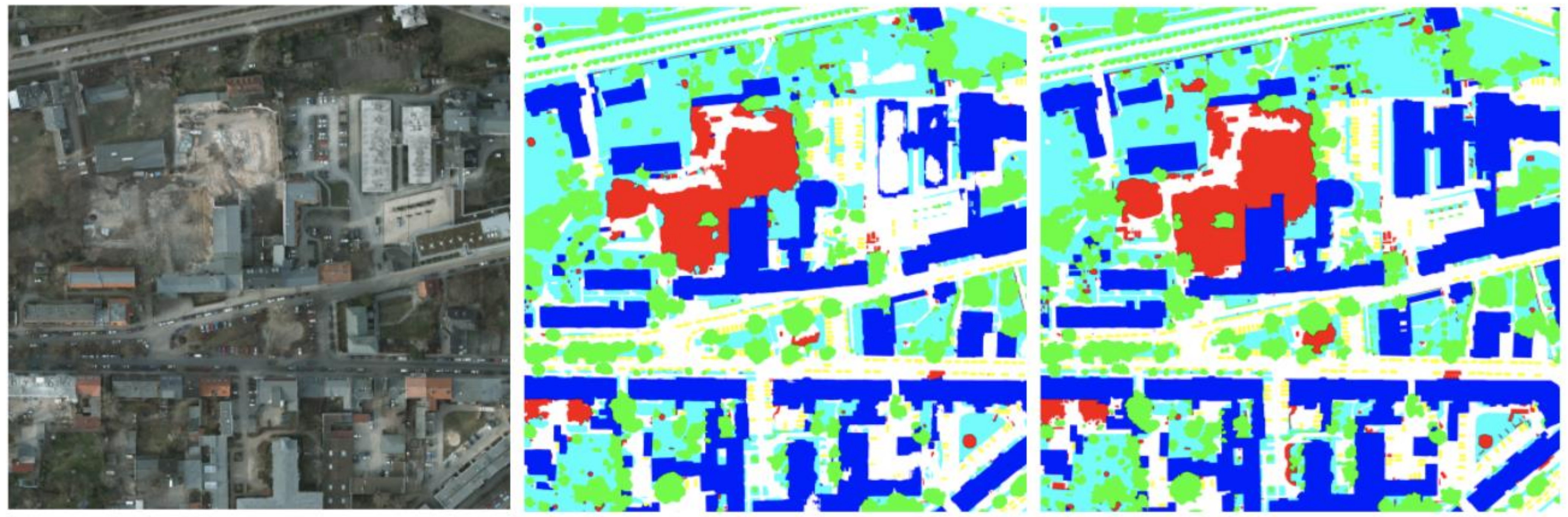


Fig. 3. Per-class *F1-score* evaluation (in %) of NEOS on the Potsdam dataset including the class Clutter in the evaluation.



a) Input b) NEOS (Ours) c) Ground truth

Fig. 4. Semantic segmentation masks by NEOS on Potsdam.

- **Evaluation of NEOS** on the dataset Potsdam
- Several evaluation metrics:
  - **F1-score**
  - Accuracy
  - Intersection over Union (IoU)

- Our model NEOS *outperforms* other baseline models

Results:
- 1) Average over the *classes*
- 2) **Per-class** results:
  - Evaluate per-class F1-score performance of NEOS on Potsdam

- Class Clutter: *Included* in the evaluation
  - 6 classes in total

- Qualitative results of our model NEOS

Evaluation of NEOS in accuracy, F-1 score and IoU

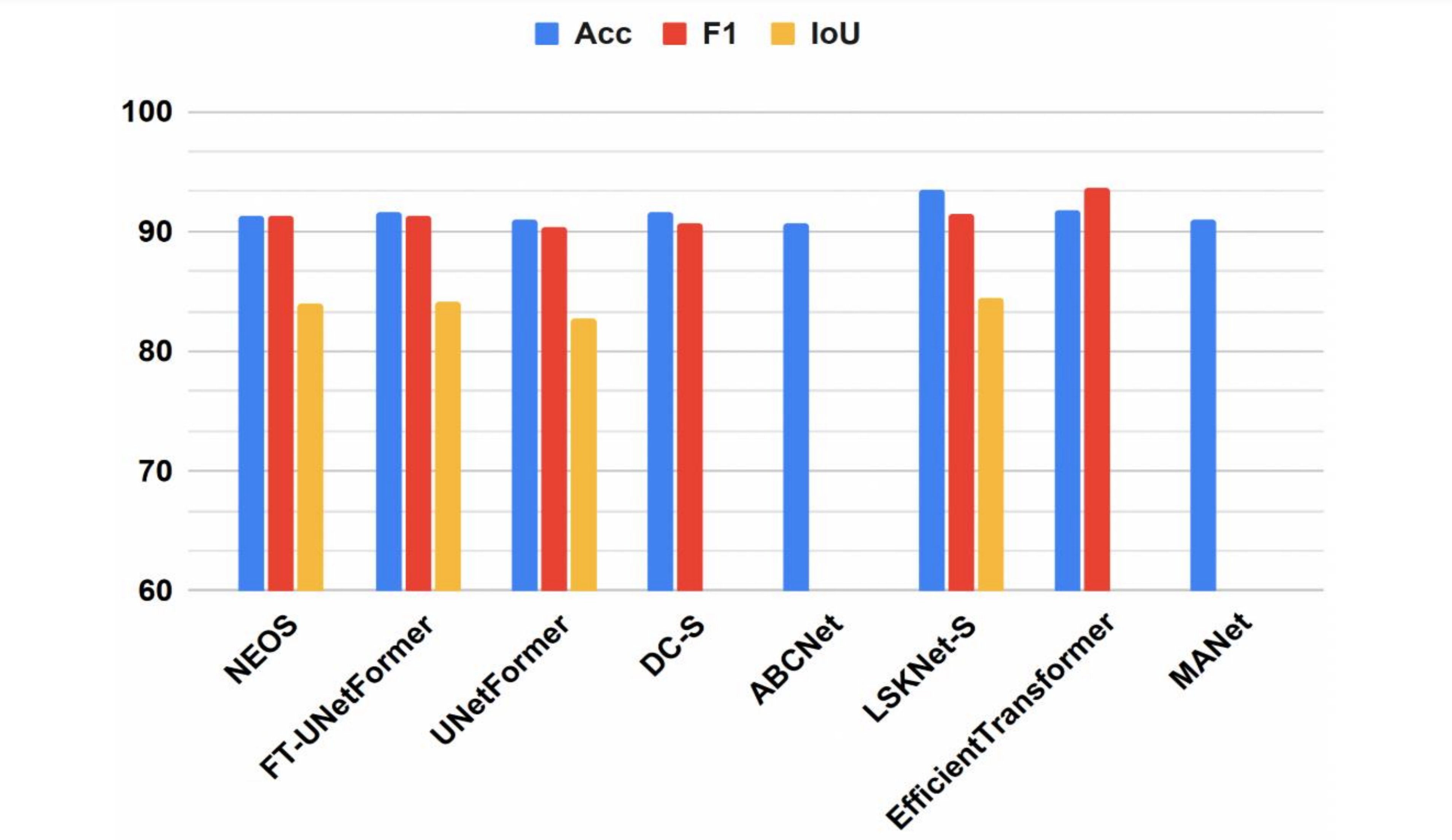


Fig. 5. Evaluation of NEOS on Vaihingen in Acc, F1 and IoU.

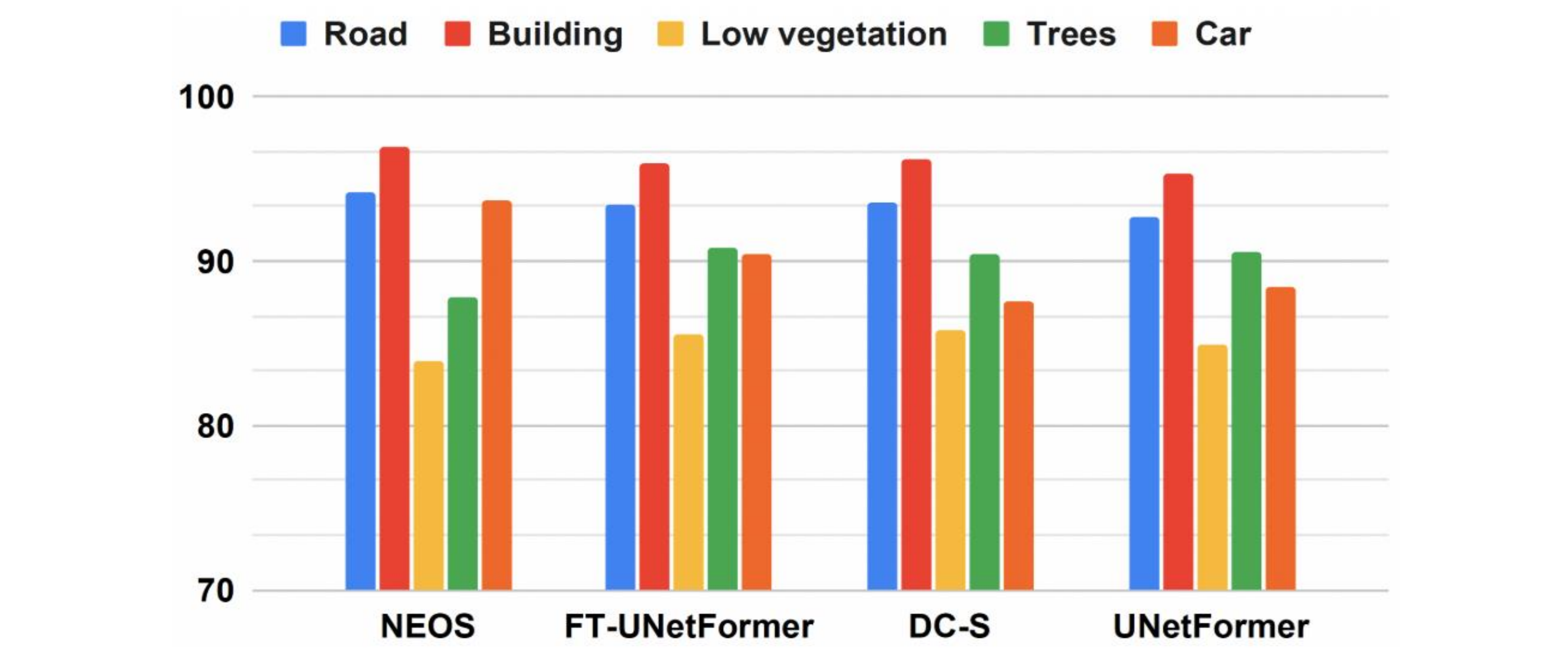


Fig. 6. Per-class F1-score evaluation of NEOS on Vaihingen.

- Evaluation of the **proposed model NEOS** on the dataset Vaihingen
- Unlabelled Domain Adaptation (UDA):
  - Evaluate NEOS on source domain data:
    - Important to achieve good performance in **both** the source and target domains
  - Source domain:
    - Labelled data:
      - Datasets Potsdam & Vaihingen
  - *Target* domain:
    - Unlabelled data:
      - Top-view aerial CVUSA

- Evaluate NEOS in *accuracy*, F1-score and IoU
- Examine the F1-score of NEOS for **each class**
  - Per-class results

- Our model NEOS on Vaihingen outperforms other baseline models: For *Roads*, Buildings and Cars

Evaluating NEOS on labelled and *unlabelled* datasets

Results

Semantic segmentation on the dataset CVUSA

- Good qualitative and *quantitative* evaluation results

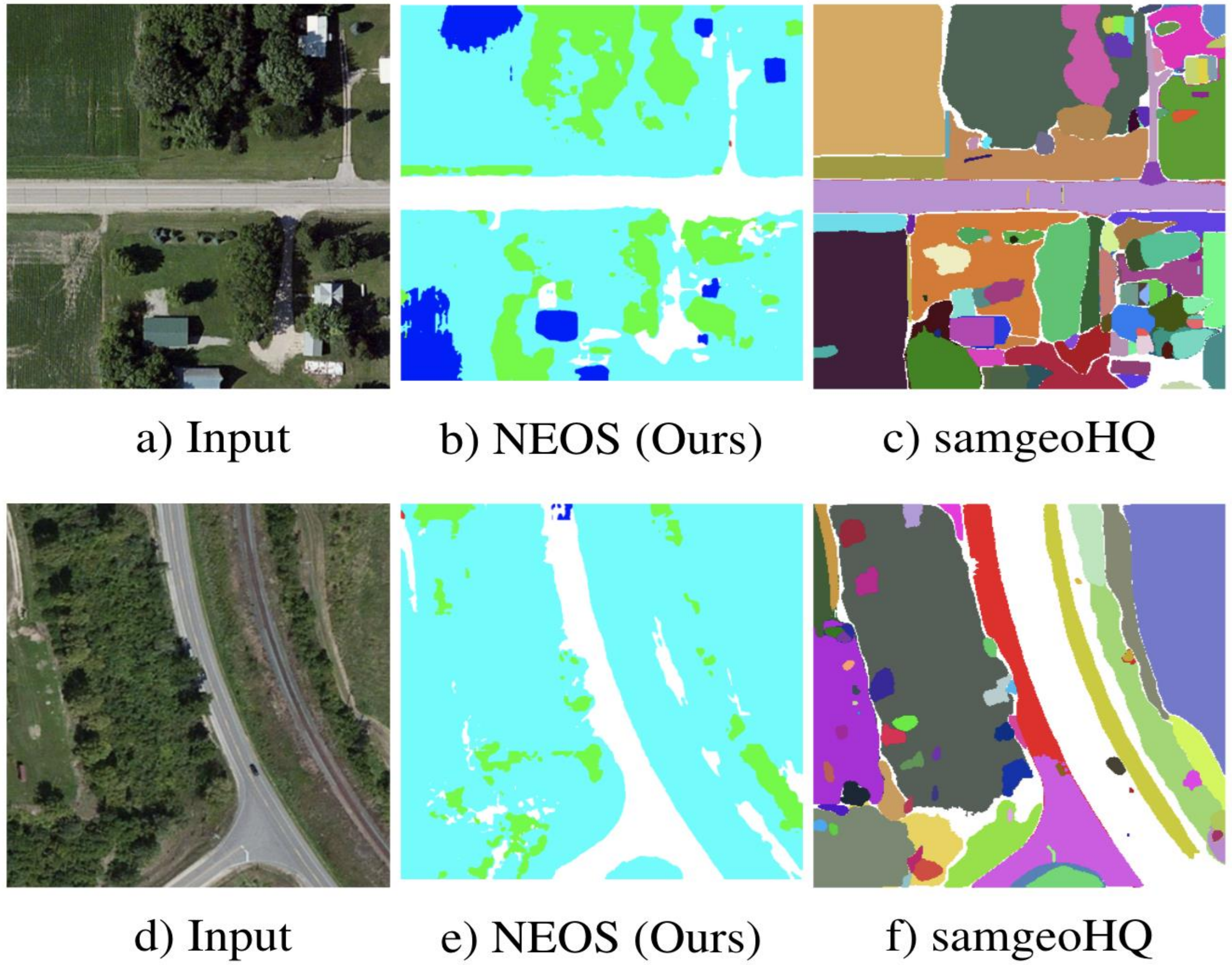
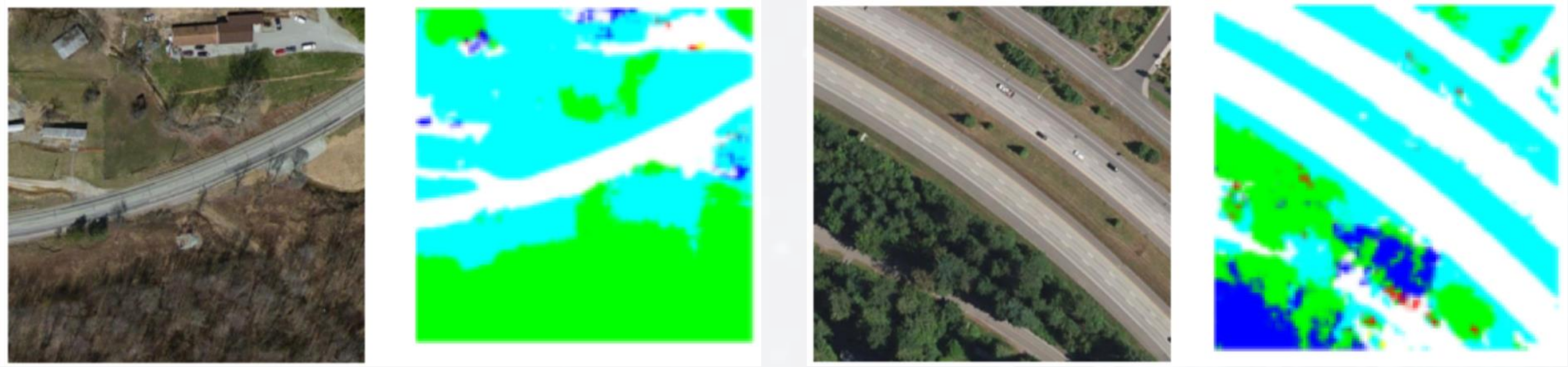


Fig. 7. Qualitative evaluation of NEOS on the unlabelled CVUSA aerial dataset, and comparison to samgeoHQ [25].

- Evaluation on the **unlabelled** dataset CVUSA
  - *Top-view* aerial CVUSA
- Testing: On a dataset with *no* ground truth masks
- **Qualitative** evaluation results of NEOS
  - Estimated *semantic* segmentation masks
- **Evaluation results:** NEOS is effective for semantic segmentation of unlabelled data
  - NEOS can outperform other alternative models
- From left to right: Input image
  - **Semantic segmentation** by our model NEOS
  - *Instance* segmentation by SAM
- In (b) and (e):
  - NEOS performs semantic segmentation
  - **Recognize** classes: Roads, Low vegetation

Qualitative and numerical evaluation of NEOS on CVUSA



Qualitative evaluation: *Accurate* segmentation and classification of roads (white)

$$\text{SPIE} = \frac{1}{R} \sum_{j=1}^{R} g(f(\mathbf{x}_j)) - g(\mathbf{x}_j)$$

Table 1. Evaluation of NEOS on the CVUSA dataset, on both the aerial (Aer) and street (Str), and the improvement (I) over the base model.

SPIE for aerial & street	Aer	I Aer	Str	I Str
NEOS (Ours)	0.047	32%	0.041	21%
Base model, SegFormer	0.069	N/A	0.052	N/A
CNN-based using Eq. (1)	0.064	7.2%	0.049	5.8%

- NEOS: Performs **semantic** segmentation
- SAM: Performs **instance** segmentation
  - *No* classification, i.e. without classes
  - Random colours for **segmentation masks**

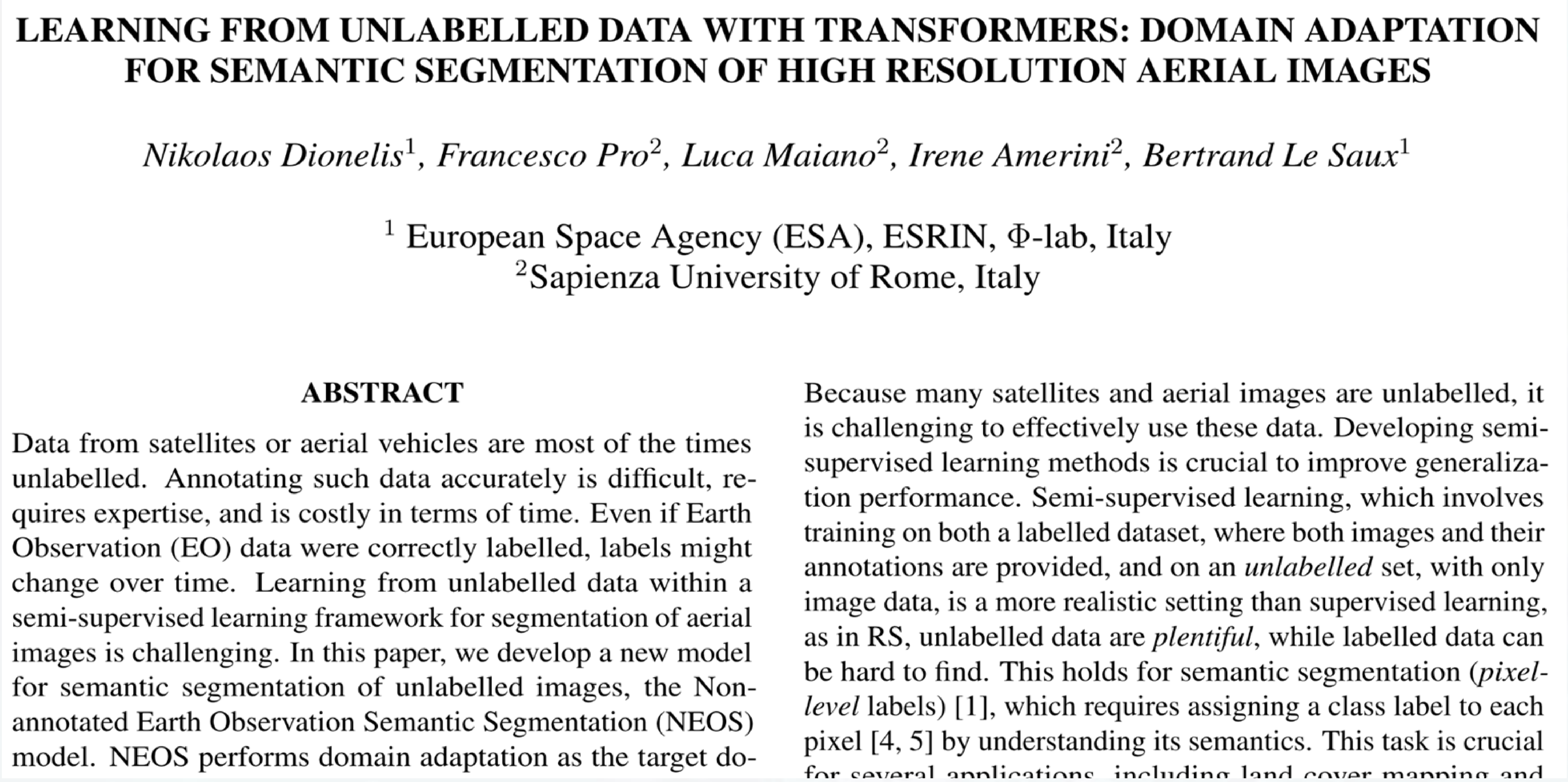
- **Qualitative evaluation** of the proposed model NEOS on unlabelled data
- **Numerical**/ quantitative evaluation of NEOS on the unlabelled dataset CVUSA

- Examine the images/ qualitative results of NEOS
- We also do this at a **large scale**
  - *Automate* the process
- We evaluate NEOS numerically

- Quantitative evaluation on **unlabelled data**
  - For semantic segmentation
  - In the *absence* of the ground truth
- Segments of Predictions and Inputs Error (SPIE)

Conclusion

esa



LEARNING FROM UNLABELLED DATA WITH TRANSFORMERS: DOMAIN ADAPTATION FOR SEMANTIC SEGMENTATION OF HIGH RESOLUTION AERIAL IMAGES

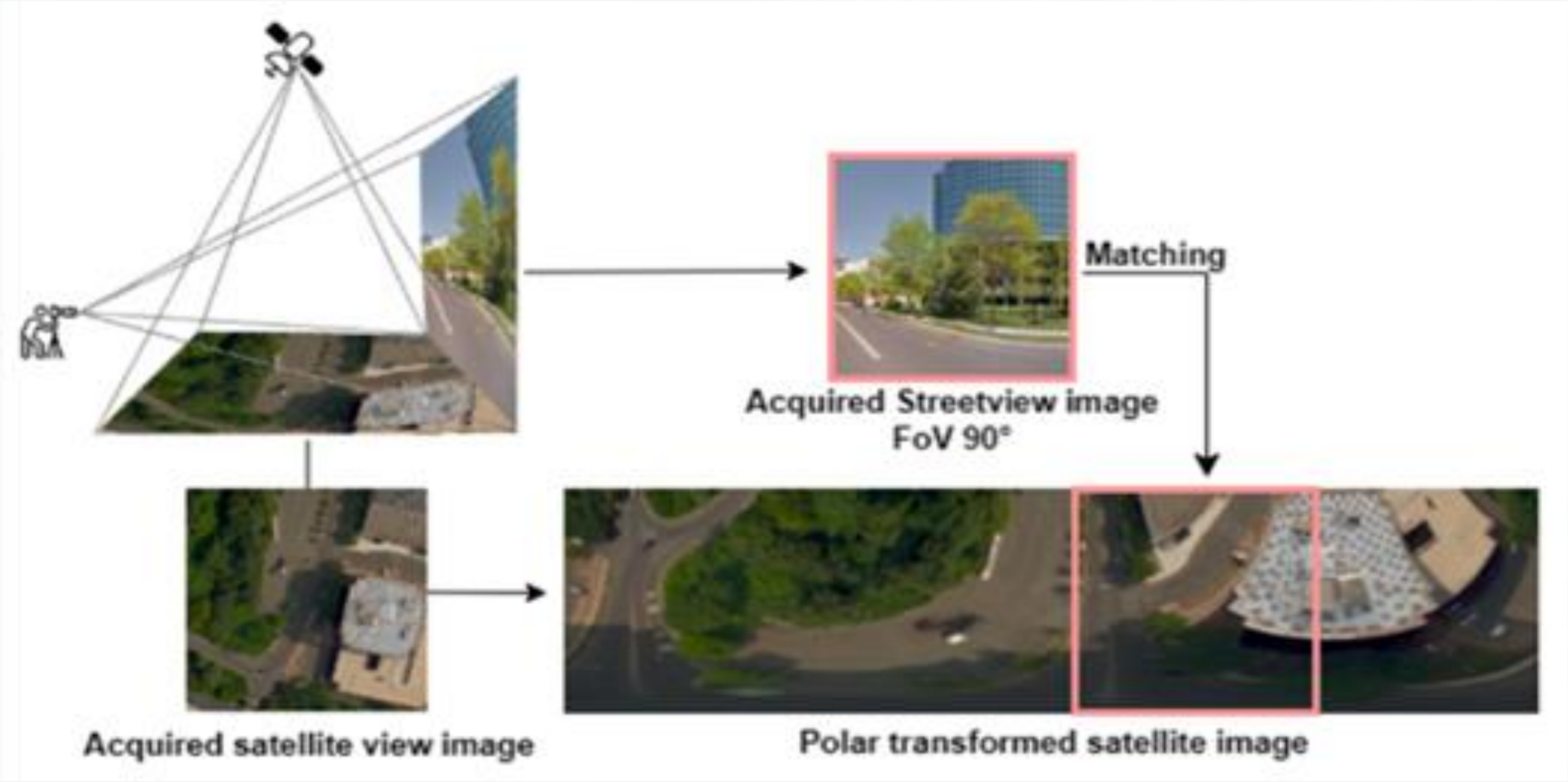
Nikolaos Dionelis[1], *Francesco Pro*[2], *Luca Maiano*[2], *Irene Amerini*[2], *Bertrand Le Saux*[1]

[1] European Space Agency (ESA), ESRIN, Φ-lab, Italy
[2] Sapienza University of Rome, Italy

ABSTRACT

Data from satellites or aerial vehicles are most of the times unlabelled. Annotating such data accurately is difficult, requires expertise, and is costly in terms of time. Even if Earth Observation (EO) data were correctly labelled, labels might change over time. Learning from unlabelled data within a semi-supervised learning framework for segmentation of aerial images is challenging. In this paper, we develop a new model for semantic segmentation of unlabelled images, the Non-annotated Earth Observation Semantic Segmentation (NEOS) model. NEOS performs domain adaptation as the target do-

Because many satellites and aerial images are unlabelled, it is challenging to effectively use these data. Developing semi-supervised learning methods is crucial to improve generalization performance. Semi-supervised learning, which involves training on both a labelled dataset, where both images and their annotations are provided, and on an *unlabelled* set, with only image data, is a more realistic setting than supervised learning, as in RS, unlabelled data are *plentiful*, while labelled data can be hard to find. This holds for semantic segmentation (*pixel-level* labels) [1], which requires assigning a class label to each pixel [4, 5] by understanding its semantics. This task is crucial for several applications, including land cover mapping and



Examined *use case*: Matching of street-view and aerial images: Geo-localization

NEOS

- *Semantic* segmentation
- Unlabelled data

- **Unlabelled** dataset CVUSA
- Labelled data from Potsdam & Vaihingen

- **Results** and Evaluation of NEOS:
  - The proposed model is effective & can outperform other baseline models

- We have also *used* the results:
  - The **semantic segmentation masks**:
  - For: Cross-view geo-location matching

Application/ use case

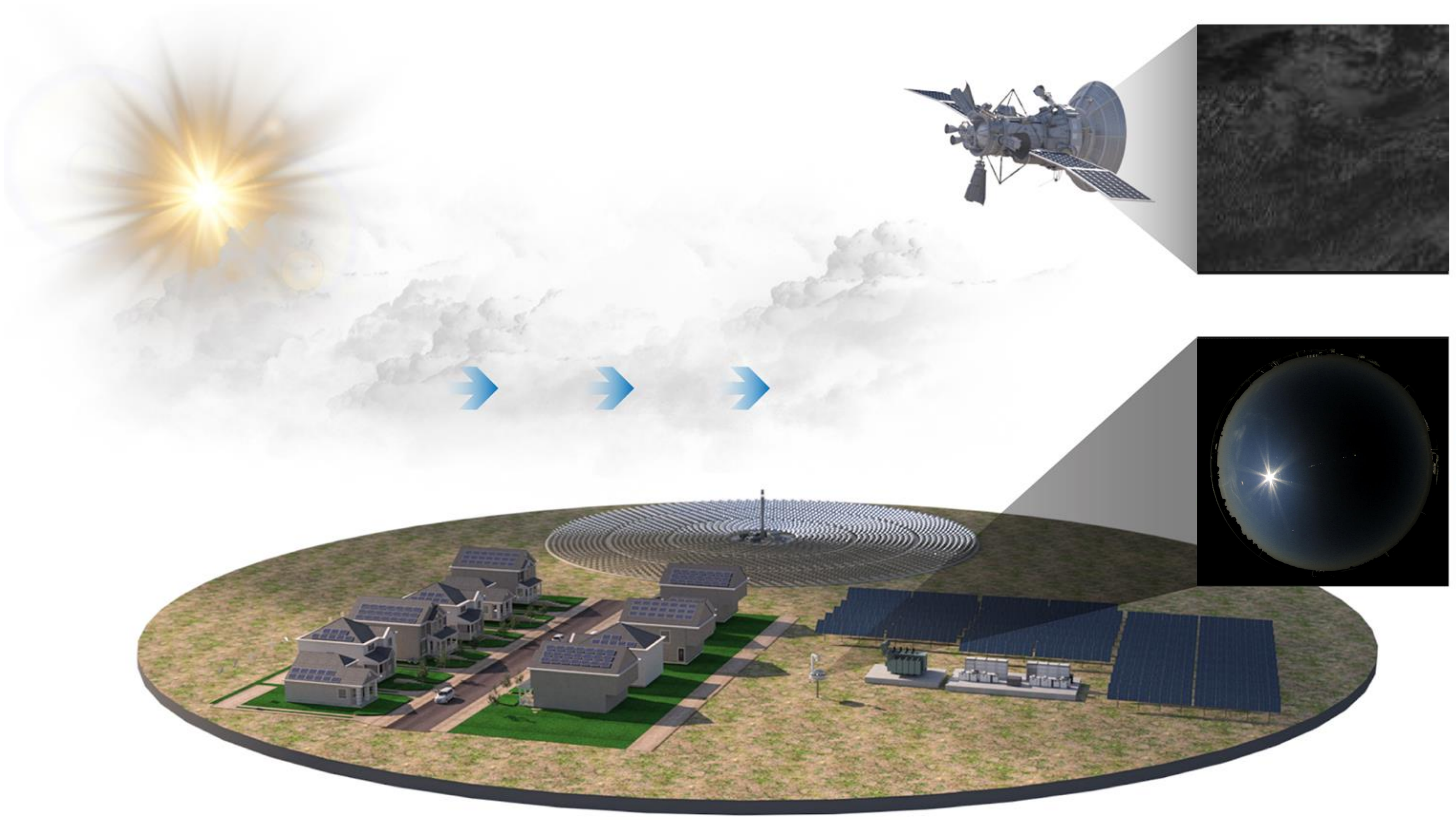
Cross-view matching of street-view and aerial images

- Using the semantic segmentation masks as additional information to improve performance
- Polar transformation for satellite images

Weather forecasting for solar energy

Φ-lab: Quentin Paletta

Short-term solar forecasting from cloud cover observations

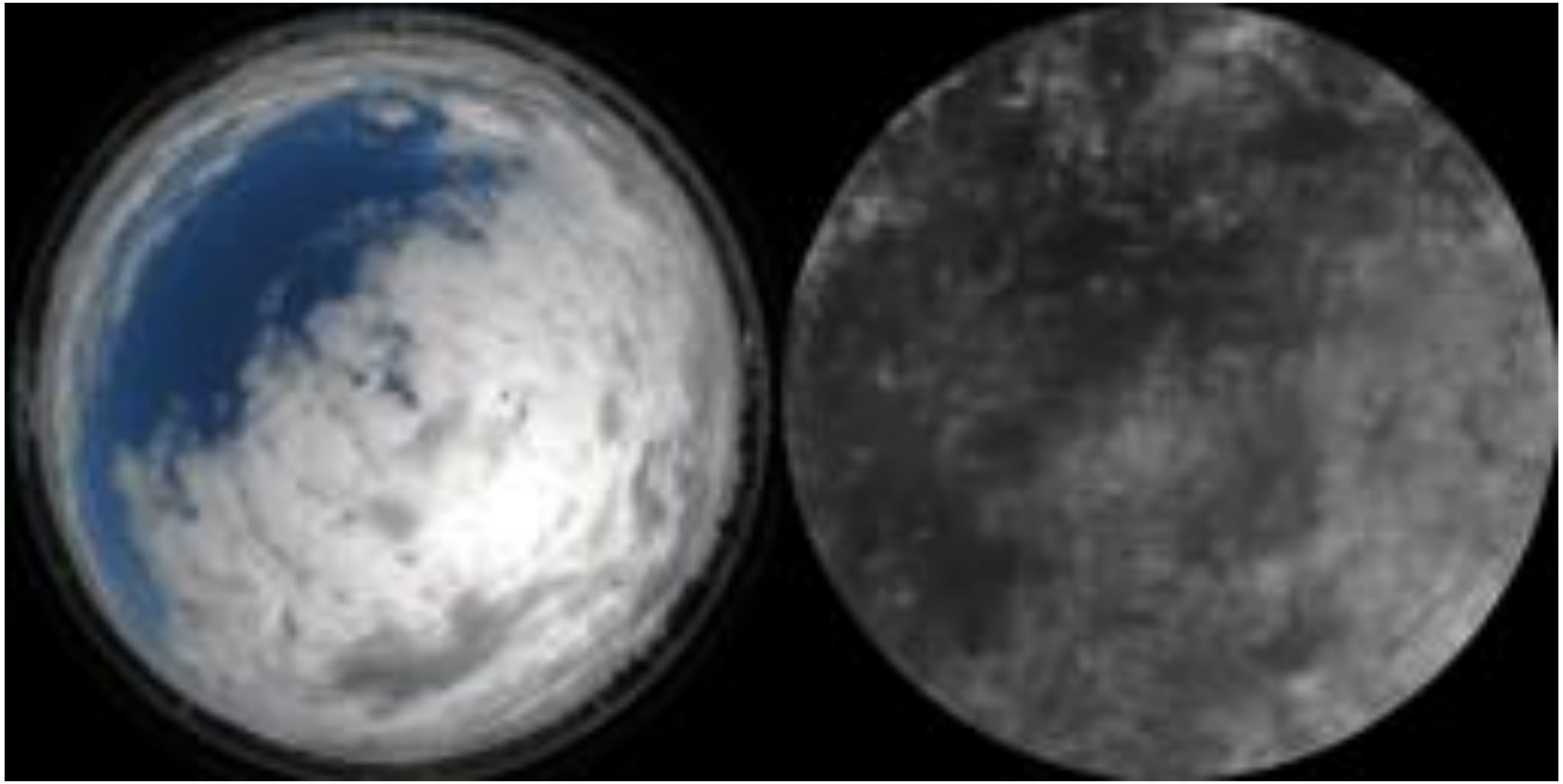


Paletta et al., Advances in solar forecasting: computer vision with deep learning, Advances in Applied Energy (2023)


Sky camera
network

Preliminary results

Input	Prediction



Ocean Topography with Implicit Neural Representation

Φ-lab: Peter Naylor, Bertrand Le Saux
Science Hub: Florian Le Guillou, Marie-Hélène Rio

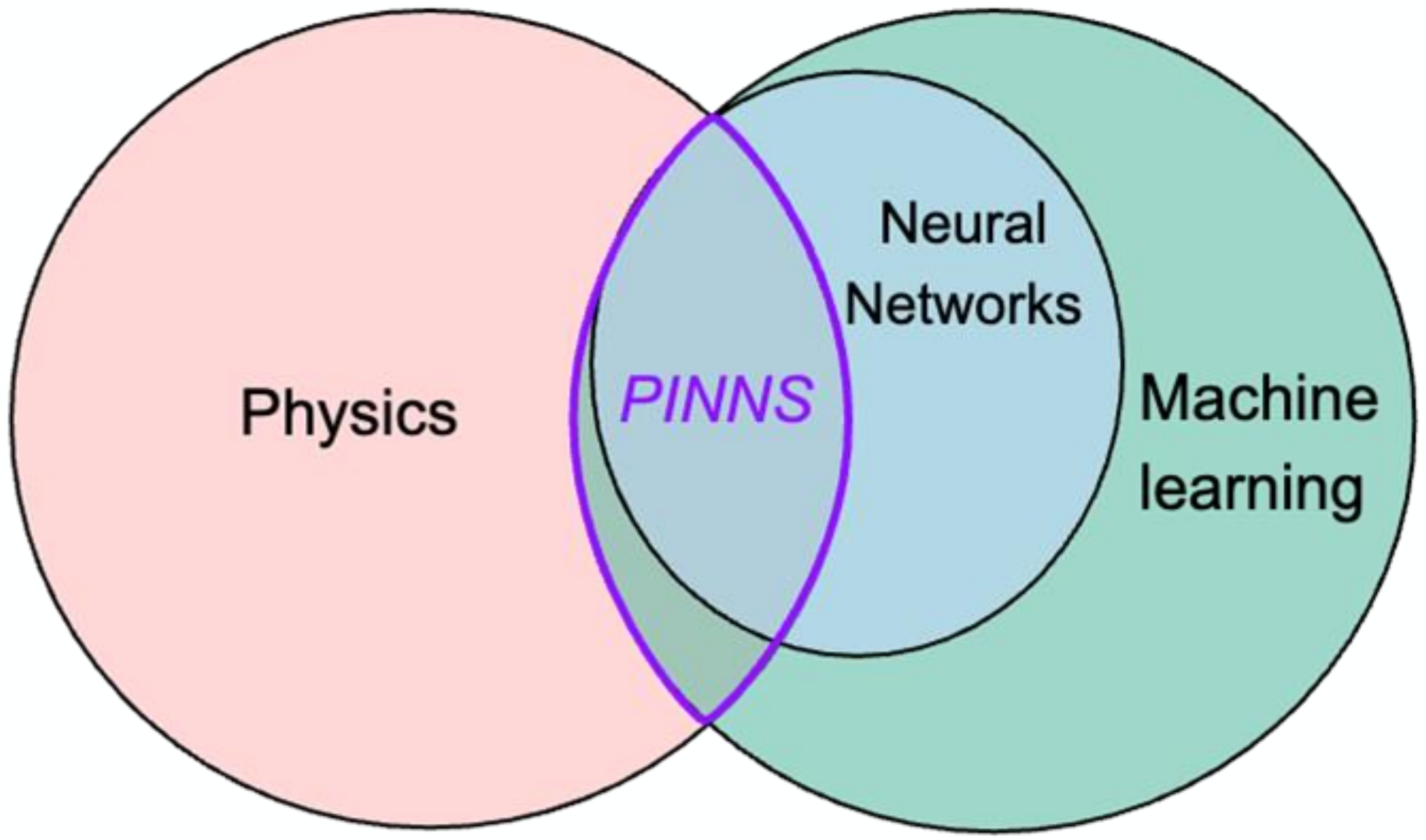
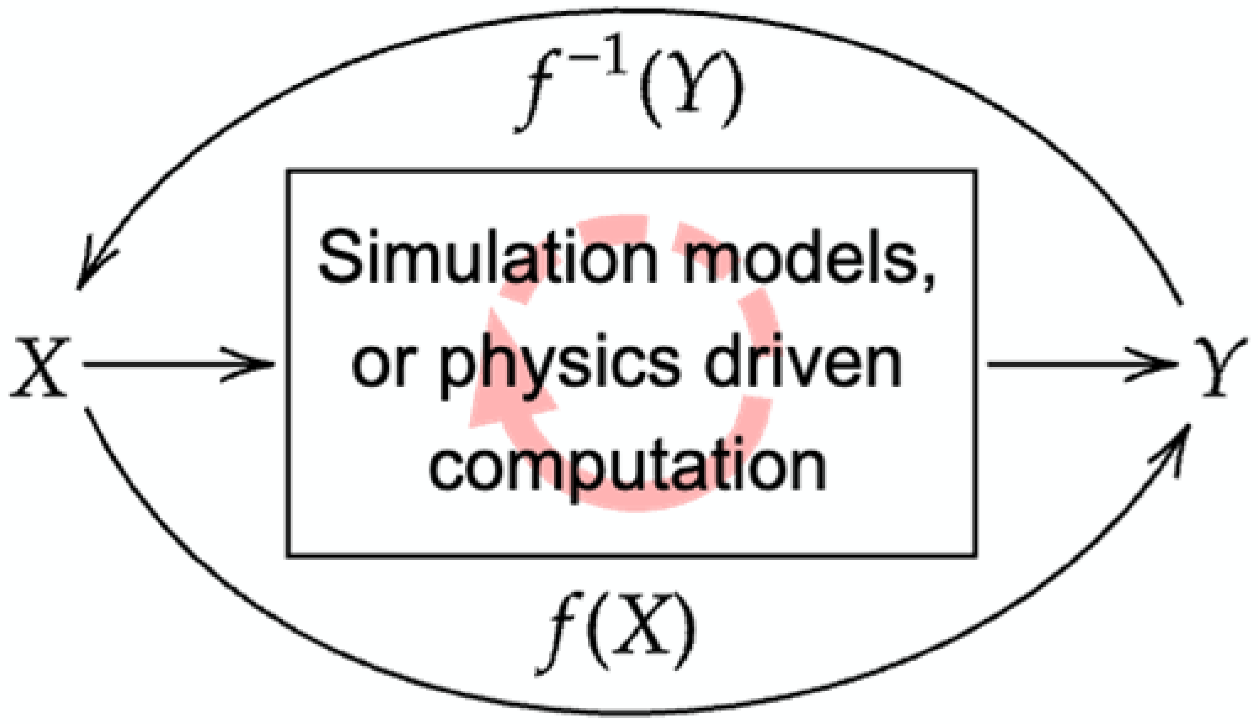


Fig. PINNs



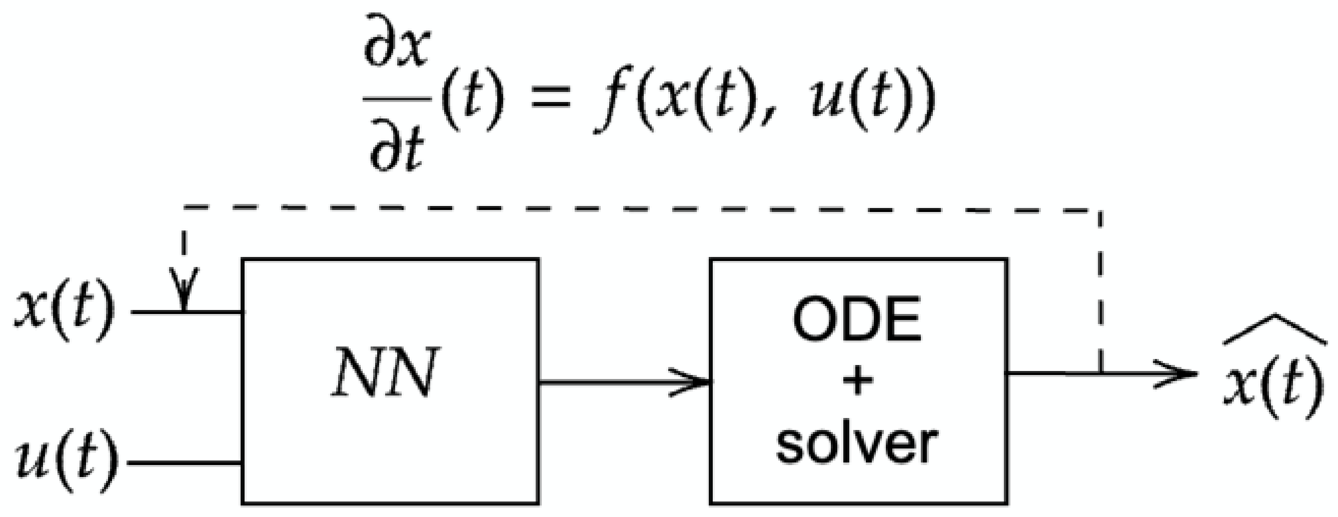


Fig. Neural ODE

Fig. Function Approximation

How do we incorporate physics into Machine learning?

PINNS:

- <u>Definition</u>: Neural networks that incorporate physical knowledge
- <u>Idea</u>: In low data availability setting, enables interpolation and extrapolation of data in a smart way, i.e. by respecting the underlying physics

Function Approximation

- Uses physics simulation, or solvers to build dataset pairs for training ML models. Can help accelerate models and overcome noise
- We approximate the real physics through image generation

Neural ODE

- Uses a neural network to solve the differential equations
- A single variable, and a single derivative

As a foundation: the universal approximation theorem of neural networks

PINNs for Earth observation: Some applications

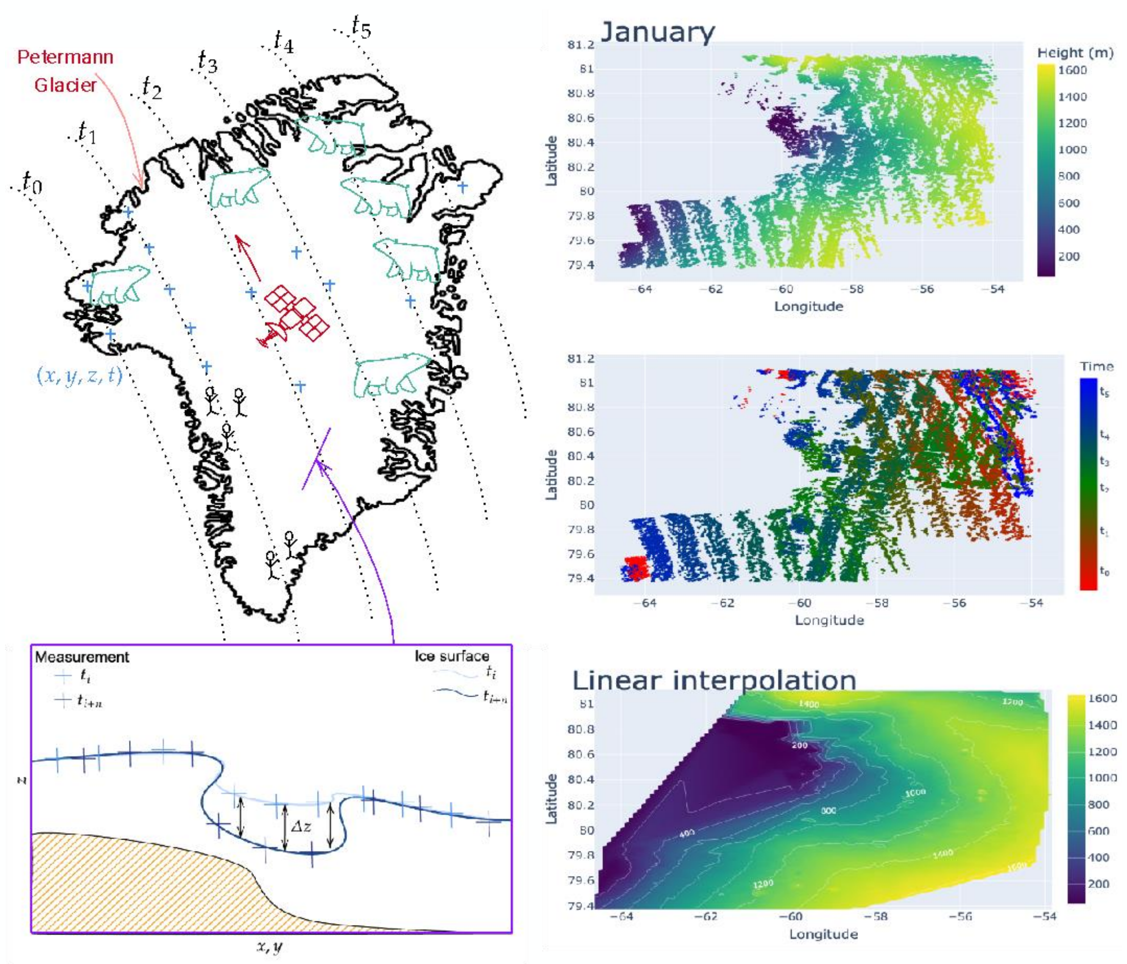


Fig. Ice Sheet monitoring

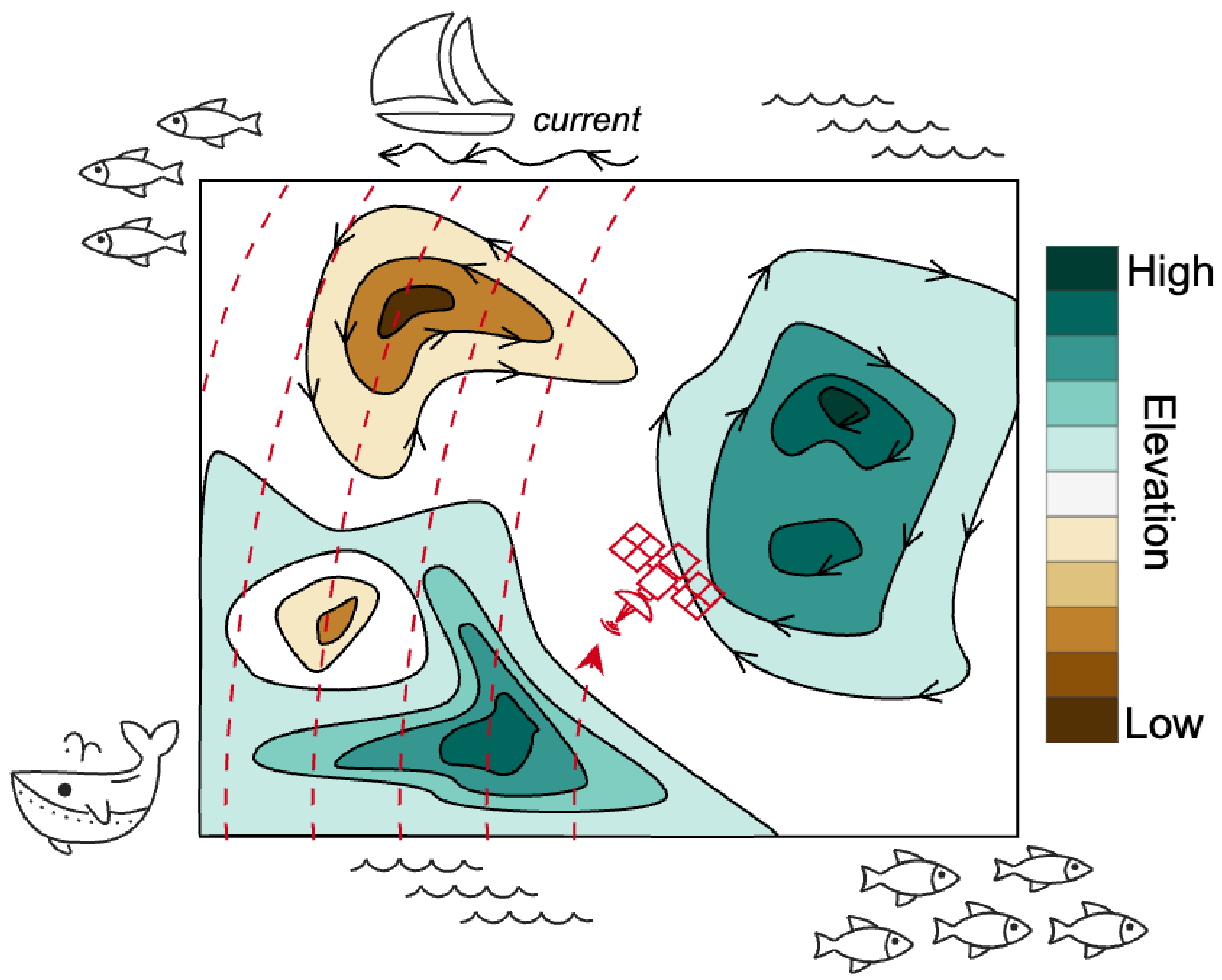


Fig. Ocean Topography

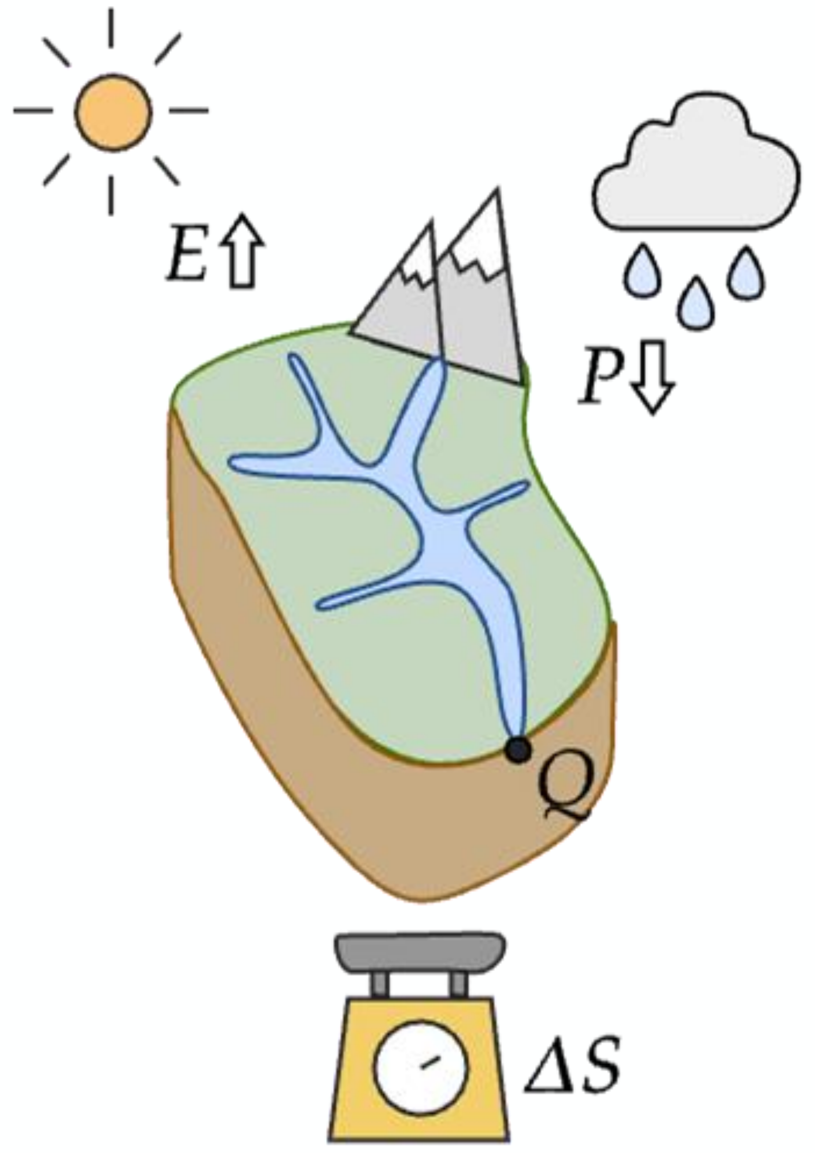


Fig. Hydrology

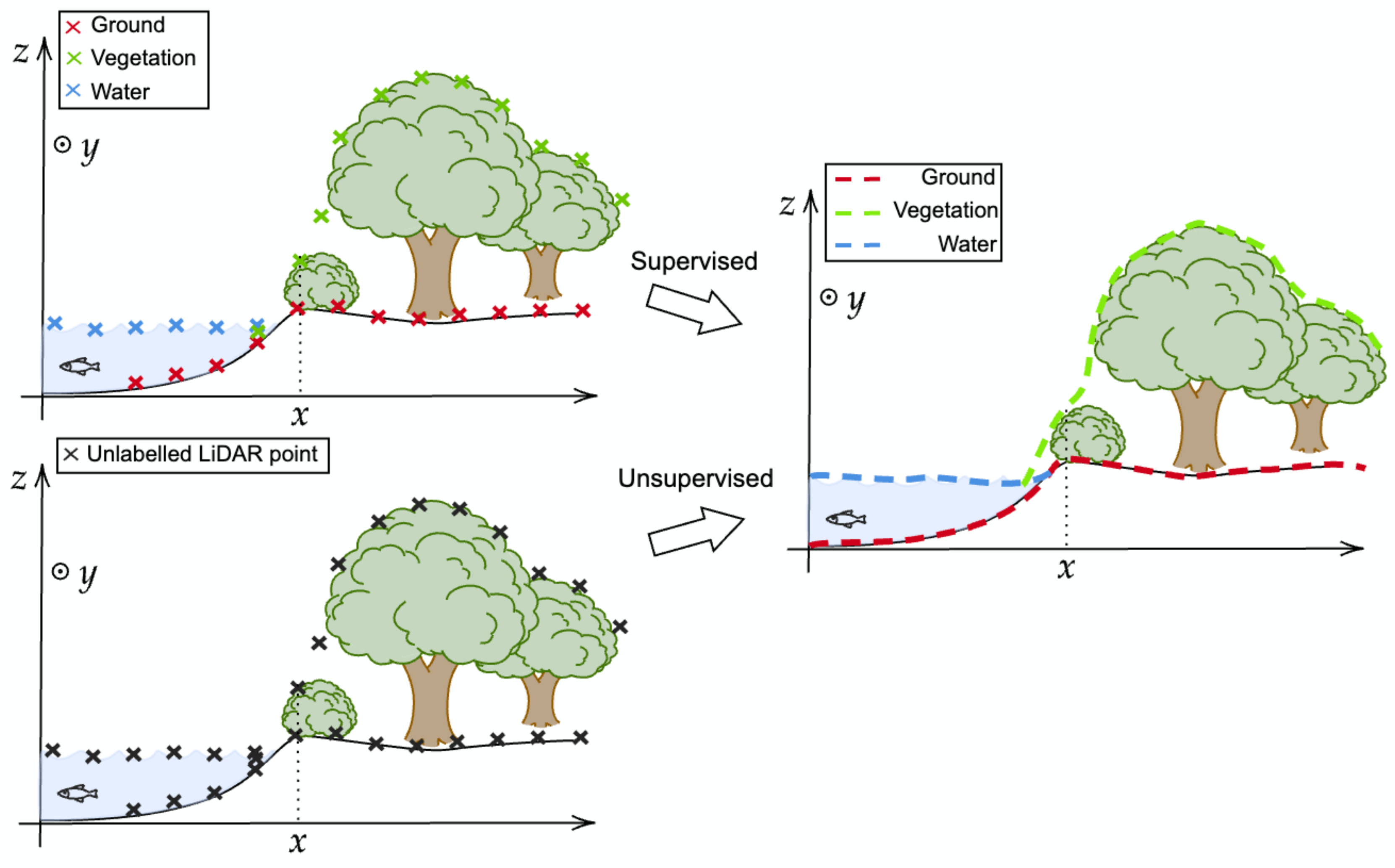


Fig. Riverbed reconstruction

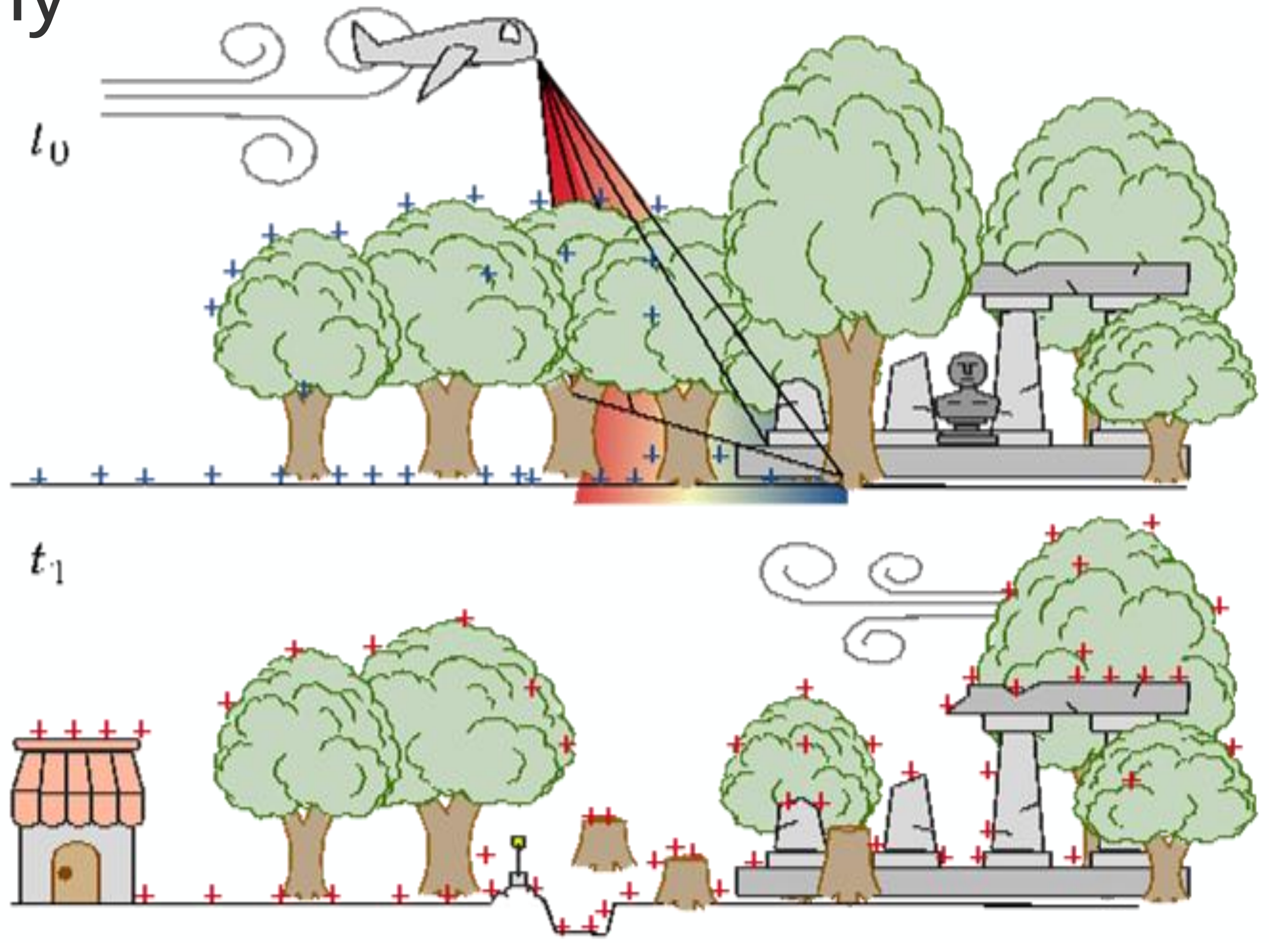


Fig. Looting detection

Ocean Topography - PINNs

esa
Φ-lab

Project objectives:

Encoding Ocean Dynamics

- Implicit Neural Representation (INR) for encoding the Ocean surface's height over a year.
- Encoding Ocean Dynamics.
- Surpassing other approaches

Technical objectives:

- INR have been applied but without Physics-informed Neural Networks (PINNs)
- The training of these models (third derivative computation)

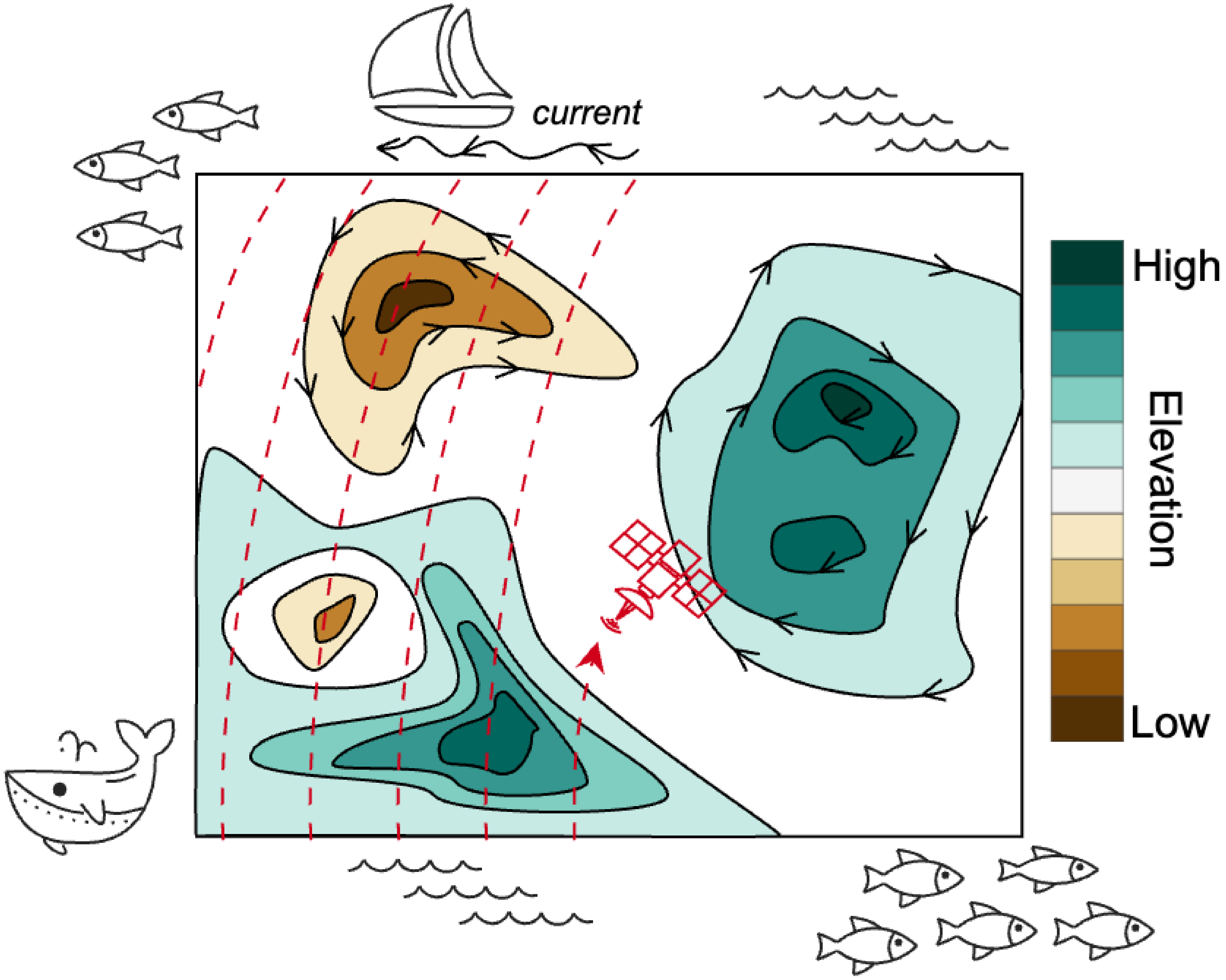
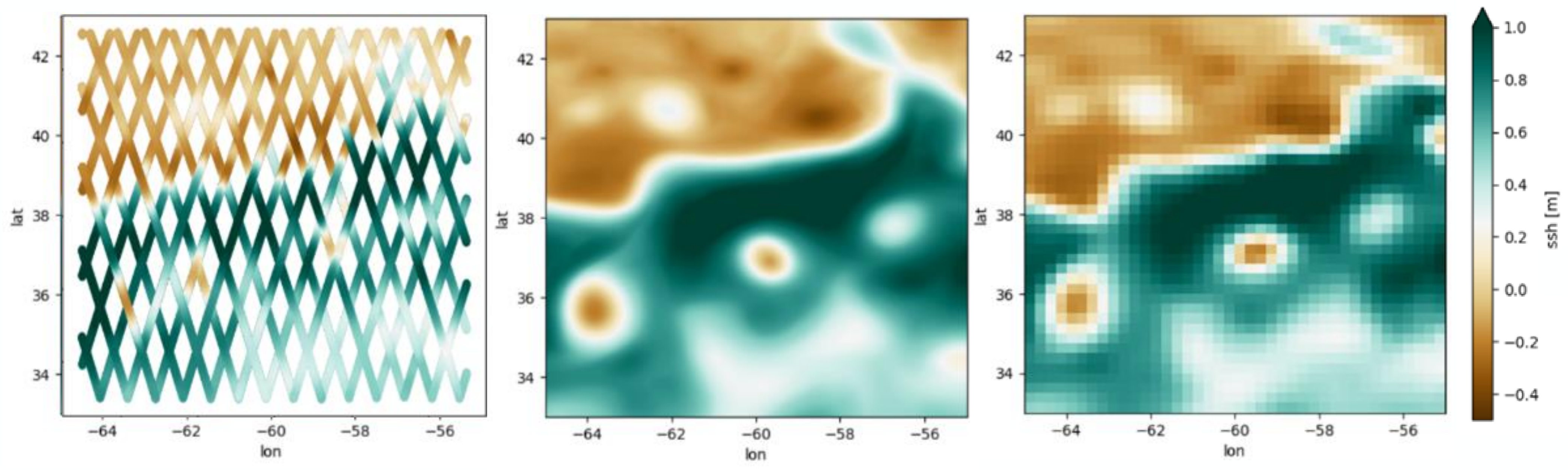


Fig. Sea Surface Height measurements from altimetry satellites



(a) Altimetry data **(b)** Reference **(c)** DUACS

Fig. Sea Surface Height

→ THE EUROPEAN SPACE AGENCY

Conclusion

- **ESA Φ-lab:**
  - **AI** for Earth Observation (EO)

- The ESA Φ-lab **Explore and Invest Offices**

- **ESA Φ-lab collaborations** and partnerships

- **Current projects at the Φ-lab**
  - PhilEO: EO Foundation Model and Evaluation Framework
  - Major TOM: Expandable **Datasets** for EO and Remote Sensing

  - Learning from unlabelled data: Domain adaptation
  - Weather forecasting for **solar energy**
  - **PINNs:** Physics Informed Neural Networks

Thank you very much for your attention!

Questions?

Nikolaos Dionelis

European Space Agency (ESA), Φ-lab, ESRIN, Italy

Nikolaos.Dionelis@esa.int