UNIVERSITÄT BONN

CRC 1502
DETECT

LARGE-SCALE LAND COVER CLASSIFICATION FROM SPARSE LABELS USING OBJECT-BASED DEEP LEARNING METHODS

JOHANNES LEONHARDT

DGK PHD COLLOQUIUM

SEPTEMBER 15, 2025

WHAT DO YOU USE LAND COVER MAPS FOR?

WHAT DO YOU USE LAND COVER MAPS FOR?

I use land cover data to identify how **urban areas** have developed over time.

WHAT DO YOU USE LAND COVER MAPS FOR?

I use land cover data to identify how **urban areas** have developed over time.

I need to know where **fields** are to identify which **crops** are grown there.

WHAT DO YOU USE LAND COVER MAPS FOR?

I use land cover data to identify how **urban areas** have developed over time.

I need to know where **fields** are to identify which **crops** are grown there.

Land cover maps help me monitor **protected areas**.

WHAT DO YOU USE LAND COVER MAPS FOR?

I use land cover data to identify how **urban areas** have developed over time.

I need to know where **fields** are to identify which **crops** are grown there.

Land cover maps help me monitor **protected areas**.

I study the impact of **land cover change** on the **regional climate**.

WHAT DO YOU USE LAND COVER MAPS FOR?

I use land cover data to identify how **urban areas** have developed over time.

I need to know where **fields** are to identify which **crops** are grown there.

Land cover maps help me monitor **protected areas**.

I study the impact of **land cover change** on the **regional climate**.

→ **Need for accurate large-scale land cover maps**

LARGE-SCALE LAND COVER CLASSIFICATION

What do we need?

LARGE-SCALE LAND COVER CLASSIFICATION

What do we need?

Satellite imagery

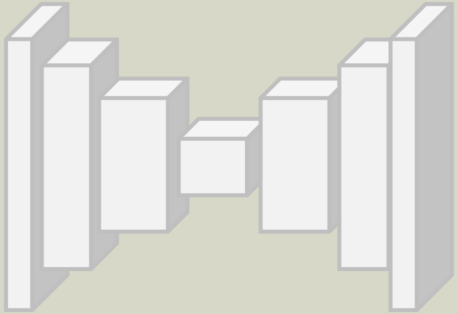
Medium resolution:
Sentinel-2, Landsat, …

LARGE-SCALE LAND COVER CLASSIFICATION

What do we need?

Satellite imagery

Medium resolution:
Sentinel-2, Landsat, …



Machine learning model

Random forest,
neural networks, …

LARGE-SCALE LAND COVER CLASSIFICATION

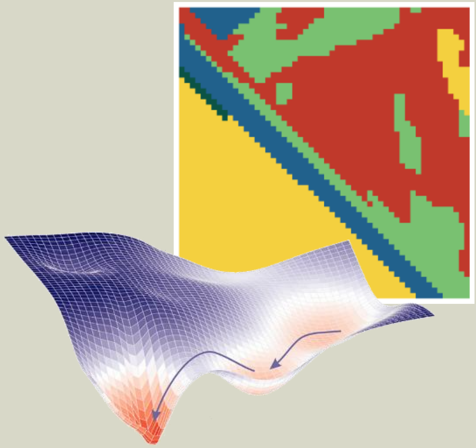
What do we need?

Satellite imagery

Medium resolution:
Sentinel-2, Landsat, …

Machine learning model

Random forest,
neural networks, …



Optimization

Ground truth labels +
cross entropy, …

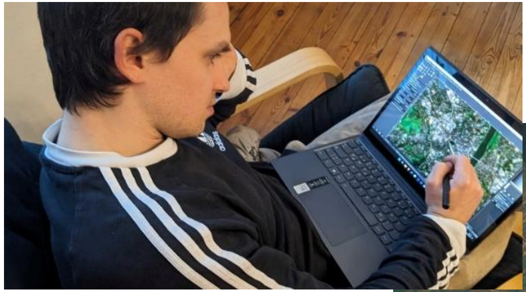
LABELS FOR LARGE-SCALE LAND COVER CLASSIFICATION

Annotations – Images are manually interpreted by annotators

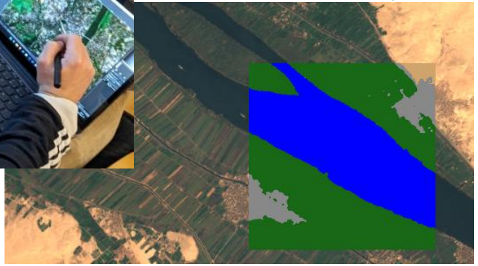
\+ Dense labels

\- Expensive to obtain for large scales

\- Subject to label errors



(Image: Alemohammad & Booth, 2020)



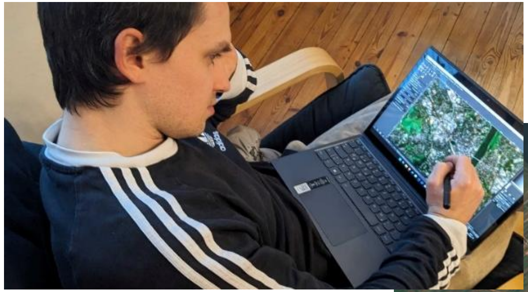
LABELS FOR LARGE-SCALE LAND COVER CLASSIFICATION

Annotations – Images are manually interpreted by annotators

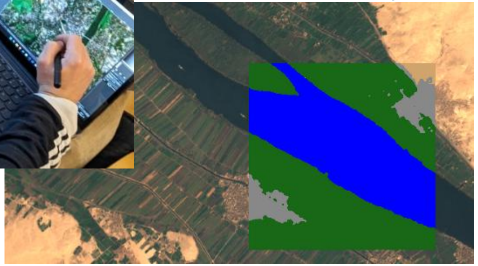
+ Dense labels

- Expensive to obtain for large scales

- Subject to label errors



(Image: Alemohammad & Booth, 2020)



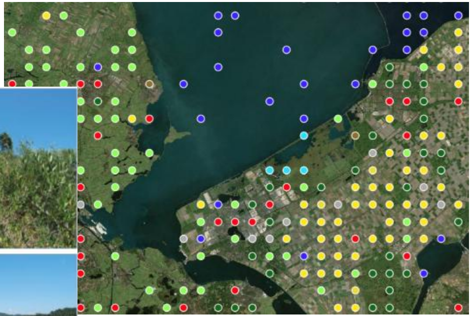
In-situ – Data from surveys is harmonized with images and used as labels

+ Readily available

+ High quality

- Sparse labels

(Image: Eurostat)

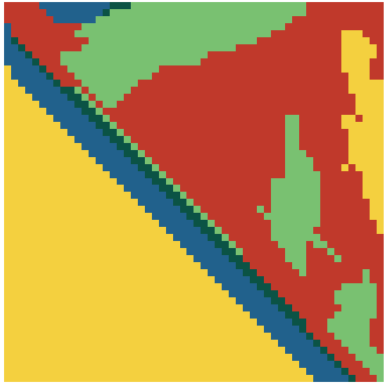


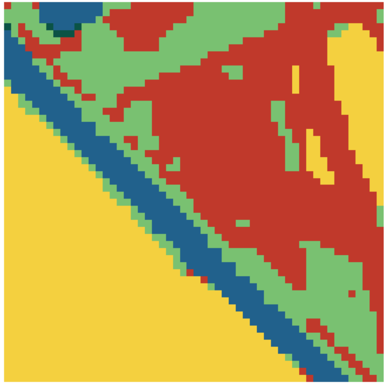
EXISTING METHODS USING SPARSE LABELS

- Pixel-wise or **object-wise** classification with rule-based or simple deep learning methods
  (e.g., Venter et al., 2021; Mirmazloumi et al., 2022)

- **Deep learning** methods from computer vision (UNet, vision transformers, ...) and semi-supervised learning
  (e.g., Galatola et al., 2023; Sharma et al., 2024)


Image


Object-wise MLP


UNet

→ Tradeoff between **accuracy** and **fragmentation**

EXISTING METHODS USING SPARSE LABELS

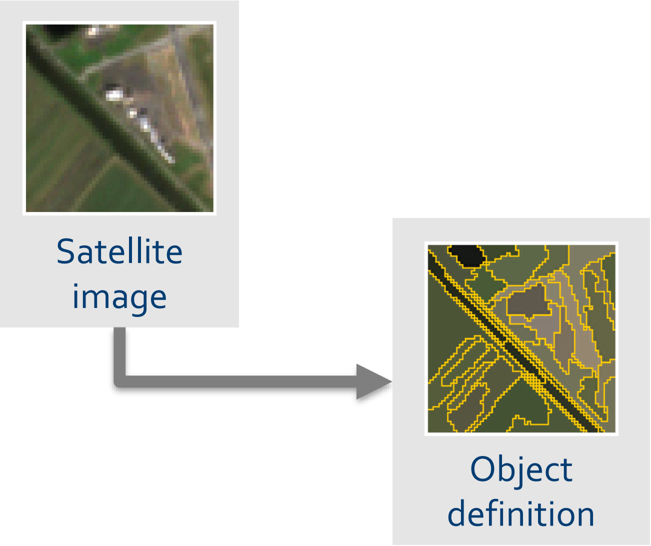
- Pixel-wise or **object-wise** classification with rule-based or simple deep learning methods
  (e.g., Venter et al., 2021; Mirmazloumi et al., 2022)

- **Deep learning** methods from computer vision (UNet, vision transformers, …) and semi-supervised learning
  (e.g., Galatola et al., 2023; Sharma et al., 2024)

Our approach:

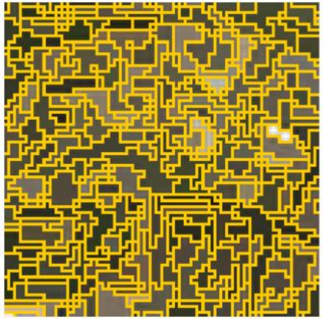
Object-based deep learning

OUR FRAMEWORK

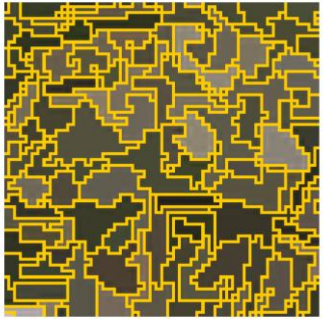


OBJECT DEFINITION

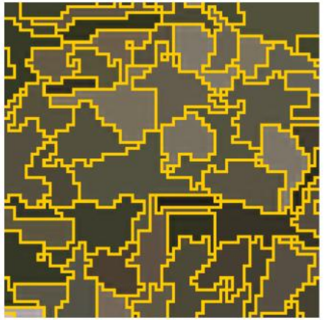
- Objects defined by graph-based **Felzenszwalb-Huttenlocher** algorithm for image segmentation which groups spectrally and spatially nearby pixels
- **Minimum mapping unit (MMU)** can be defined by min. area parameter
- MMU=1px → Each pixel is an object



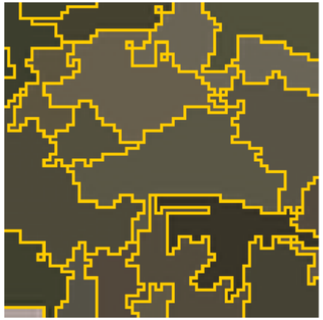
MMU=5px



MMU=10px

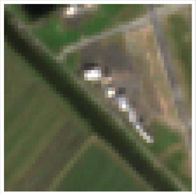


MMU=20px

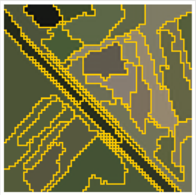


MMU=40px

OUR FRAMEWORK

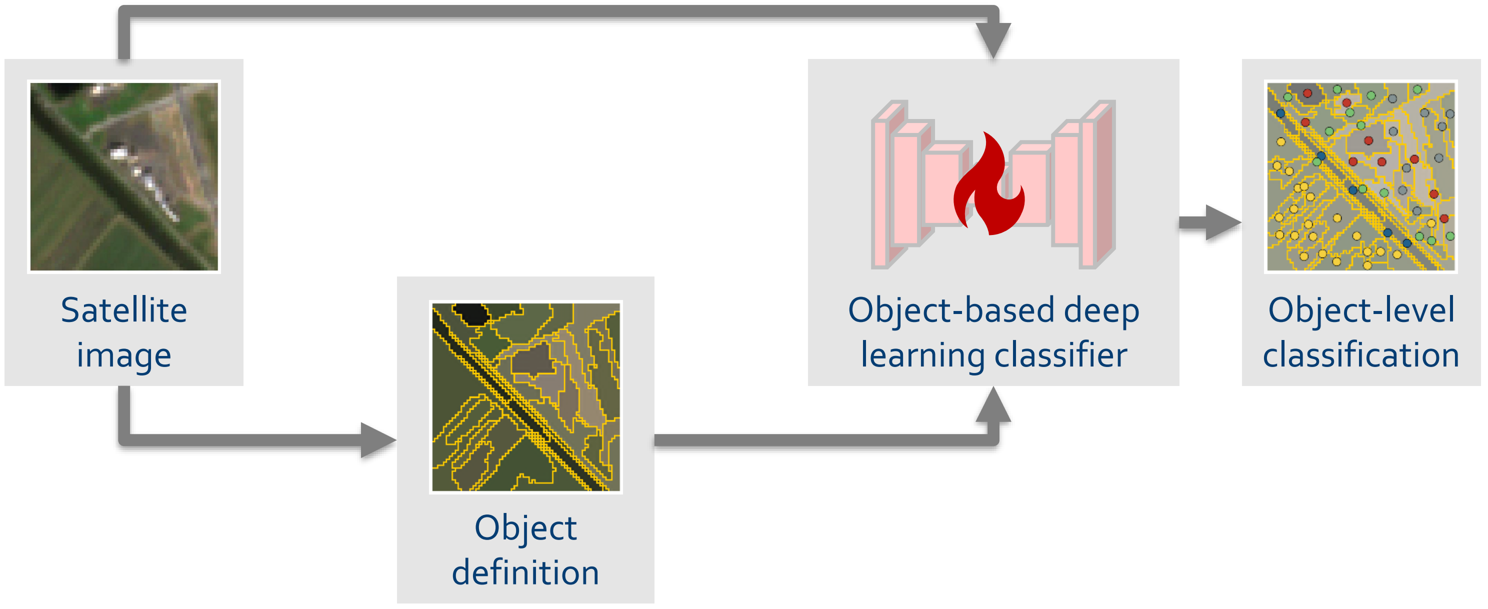


Satellite image



Object definition

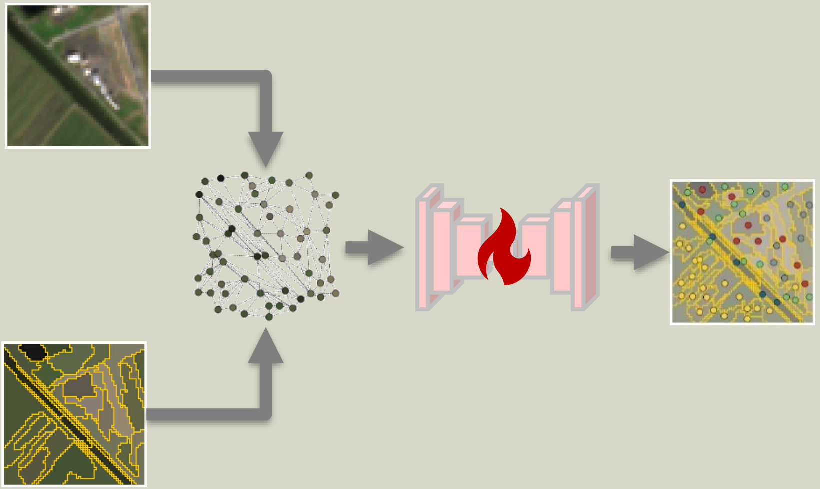
OUR FRAMEWORK



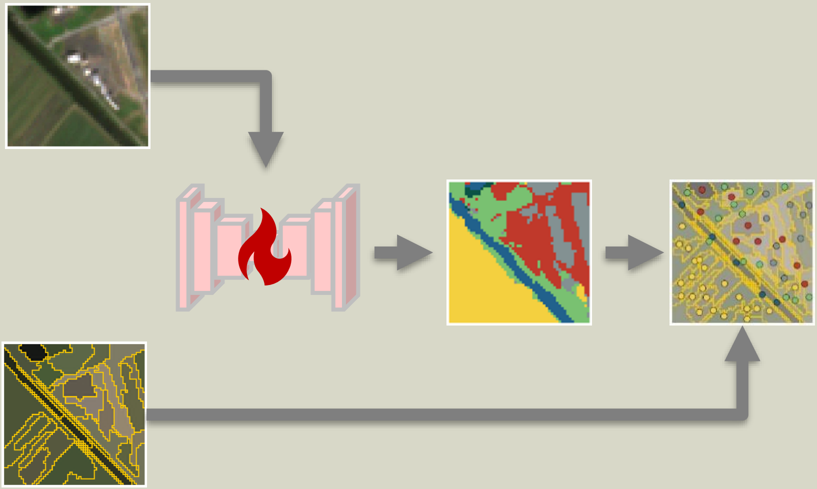
OBJECT-BASED DEEP LEARNING CLASSIFIER

Two approaches

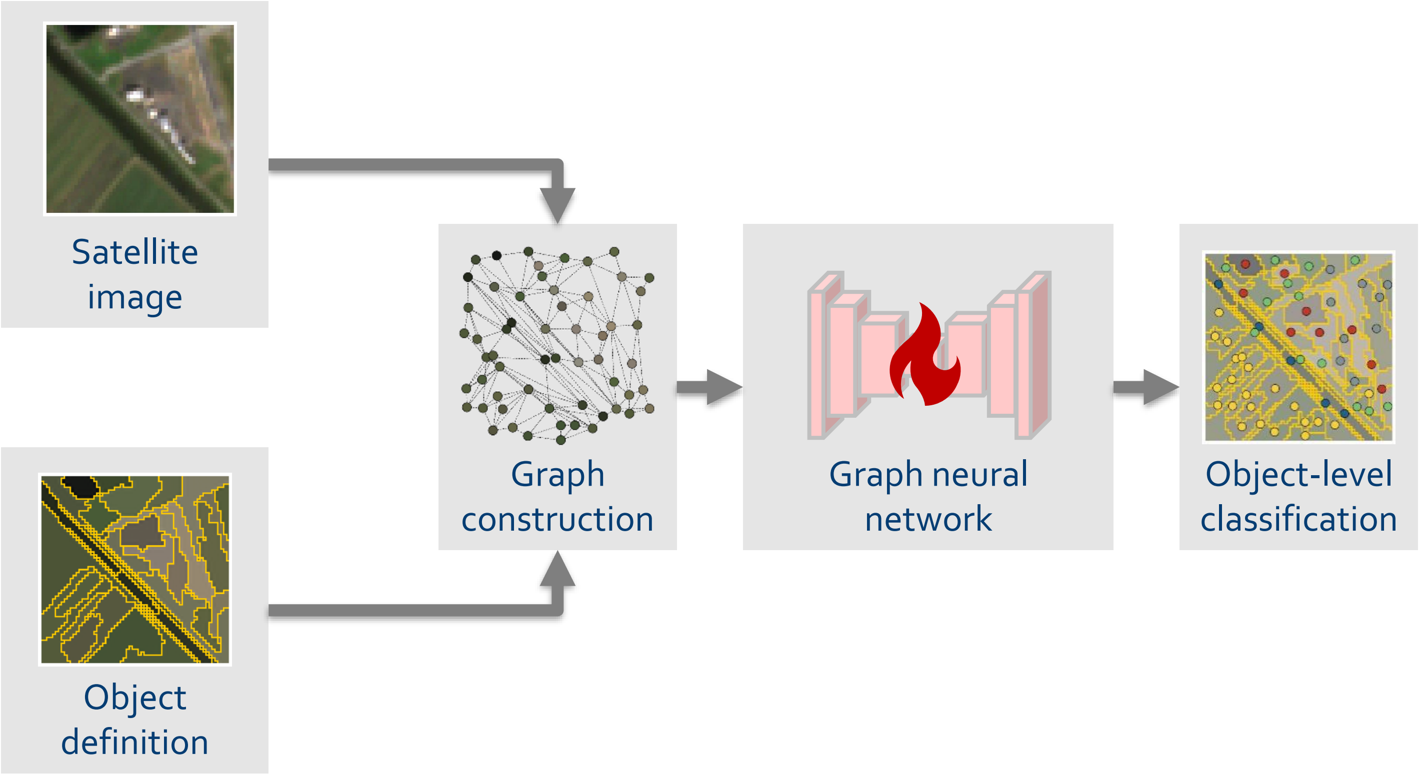
Input-level object aggregation



Output-level object aggregation



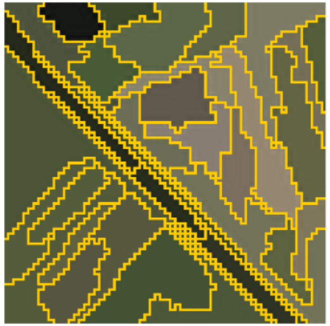
INPUT-LEVEL OBJECT AGGREGATION

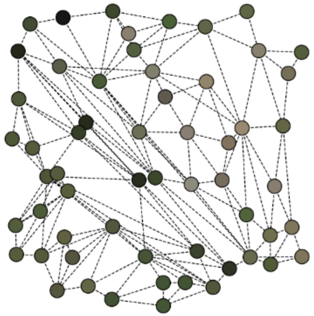


INPUT-LEVEL OBJECT AGGREGATION

Graph construction

- Each object/segment is represented as a node
- **Node features**
  - Mean intensities
  - Variability: Minimum, maximum, and standard deviation of intensities
  - Geometry: Size, radial dispersion, …
- **Edges** are based on region adjacency





INPUT-LEVEL OBJECT AGGREGATION

Graph neural networks

- Similar to **convolutional neural networks**, but **operate on graphs** instead of images
- Different **graph convolution operators** for aggregating data across neighborhoods
  - GCN, GraphSAGE, GAT, Transformer
- Different **architectures**
  - BaseGNN, Graph UNet

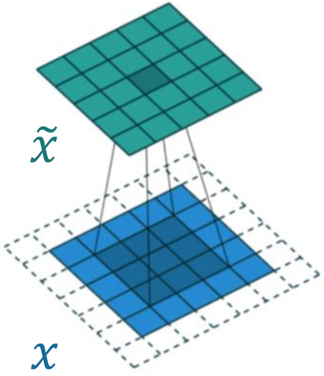
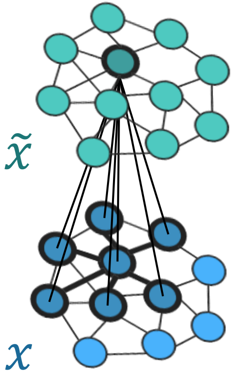
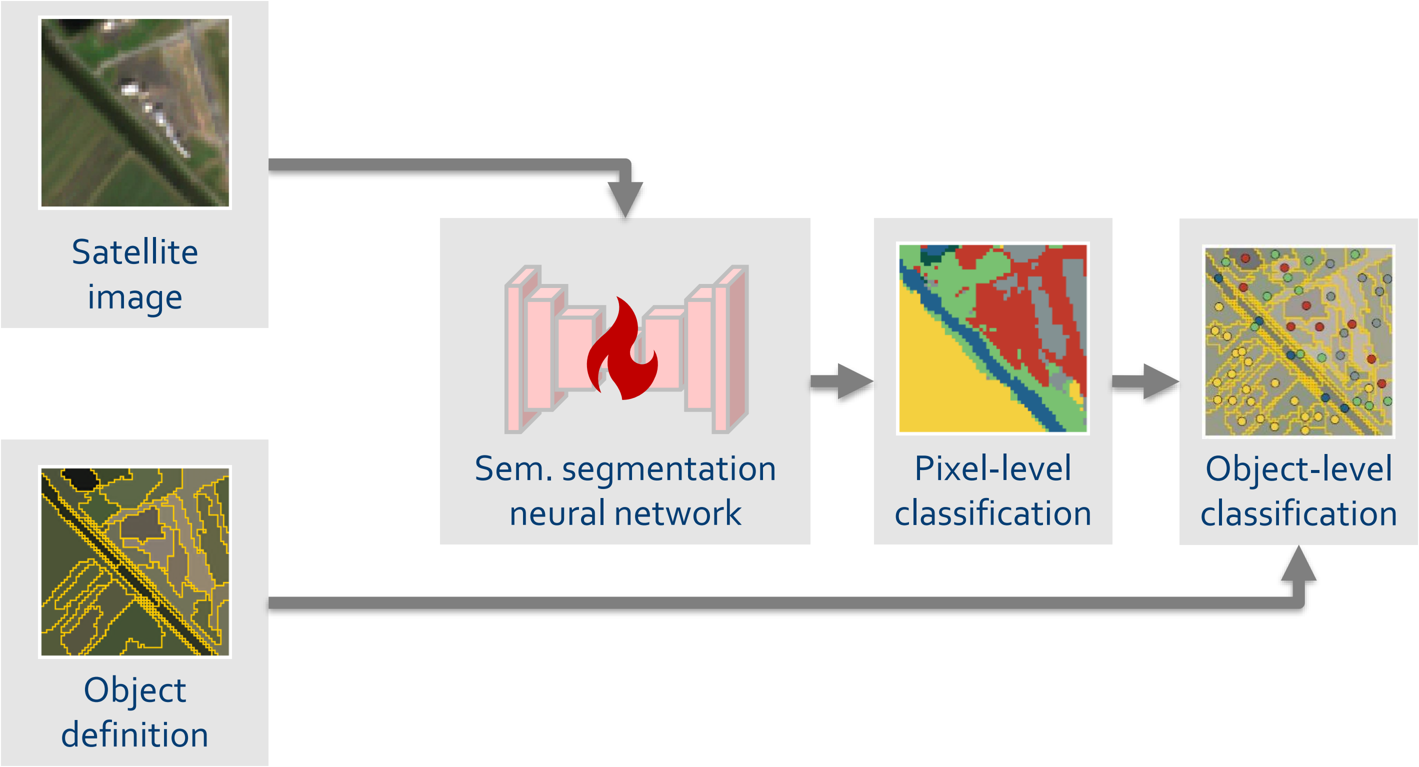


Image convolution
(Image: Prof. Dr. Johannes Maucher)



Graph convolution

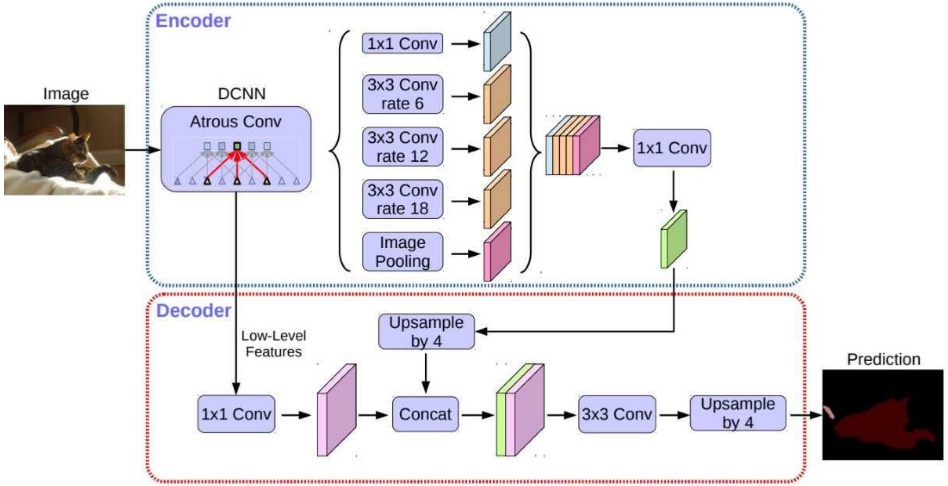
OUTPUT-LEVEL OBJECT AGGREGATION



OUTPUT-LEVEL OBJECT AGGREGATION

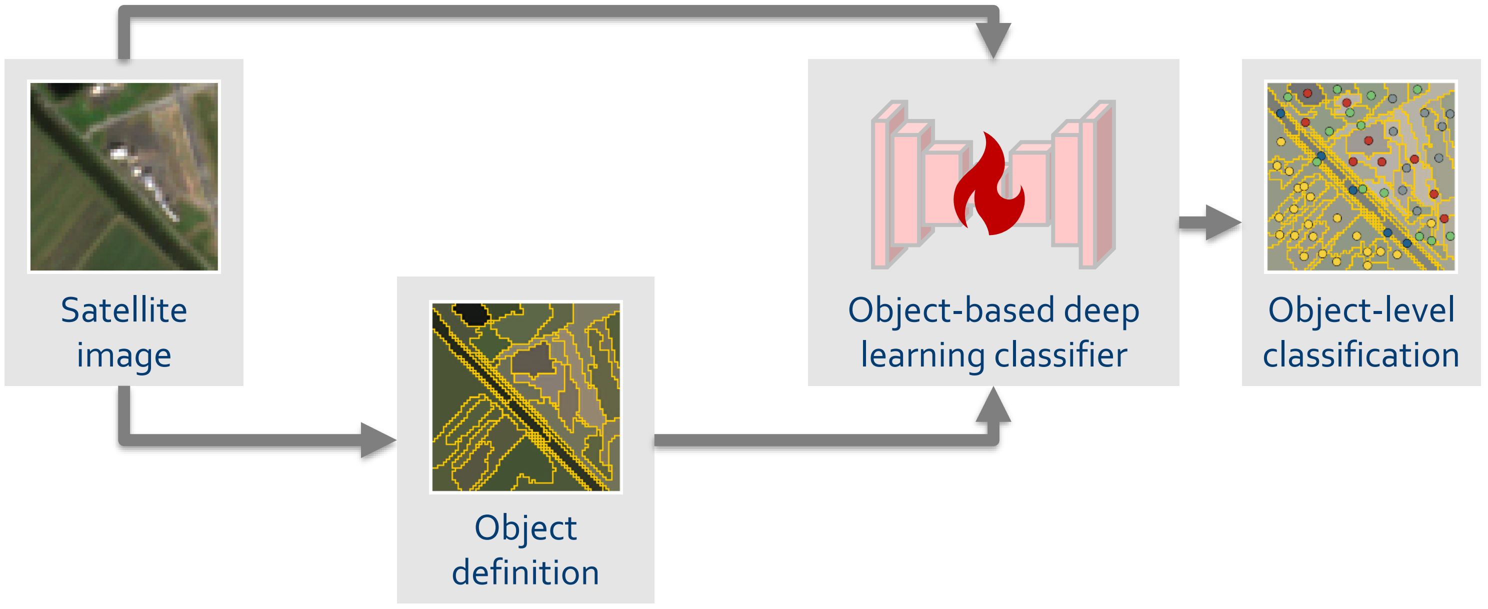
Semantic segmentation neural networks

- BaseCNN, Unet, UNet++, DeepLabV3, Segformer
- Important detail: Images need to be **upscaled** prior to classifications

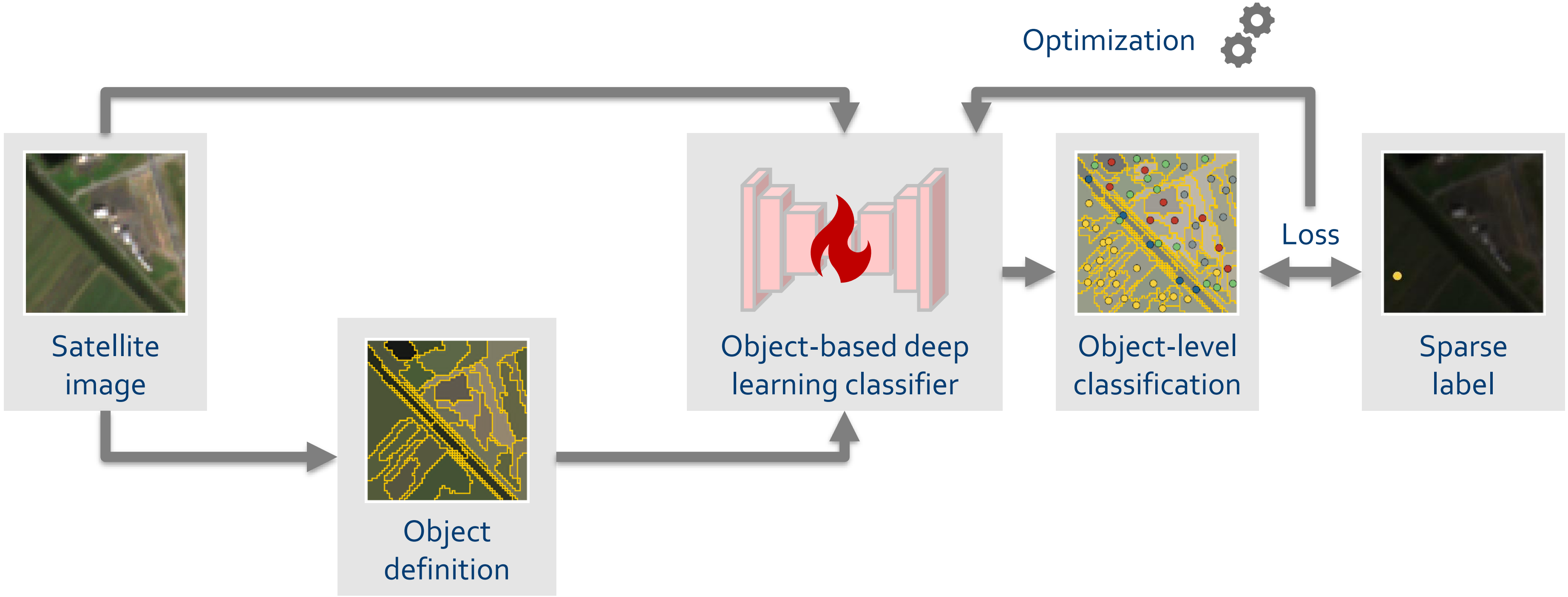


DeepLabV3 (Chen et al., 2017)

OUR FRAMEWORK



OUR FRAMEWORK



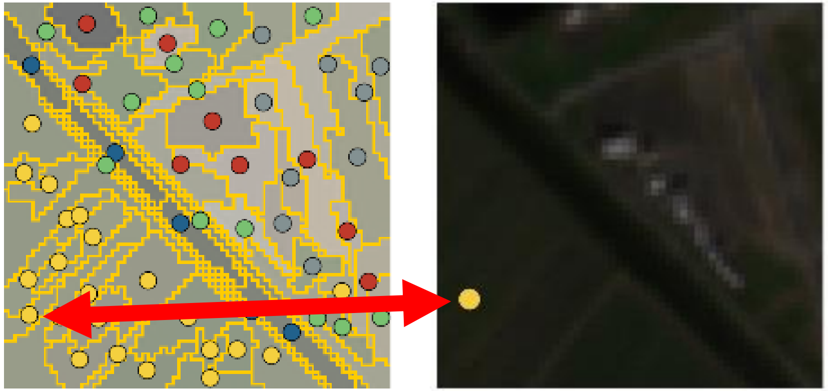
OPTIMIZATION WITH SPARSE LABELS

Sparse cross entropy loss

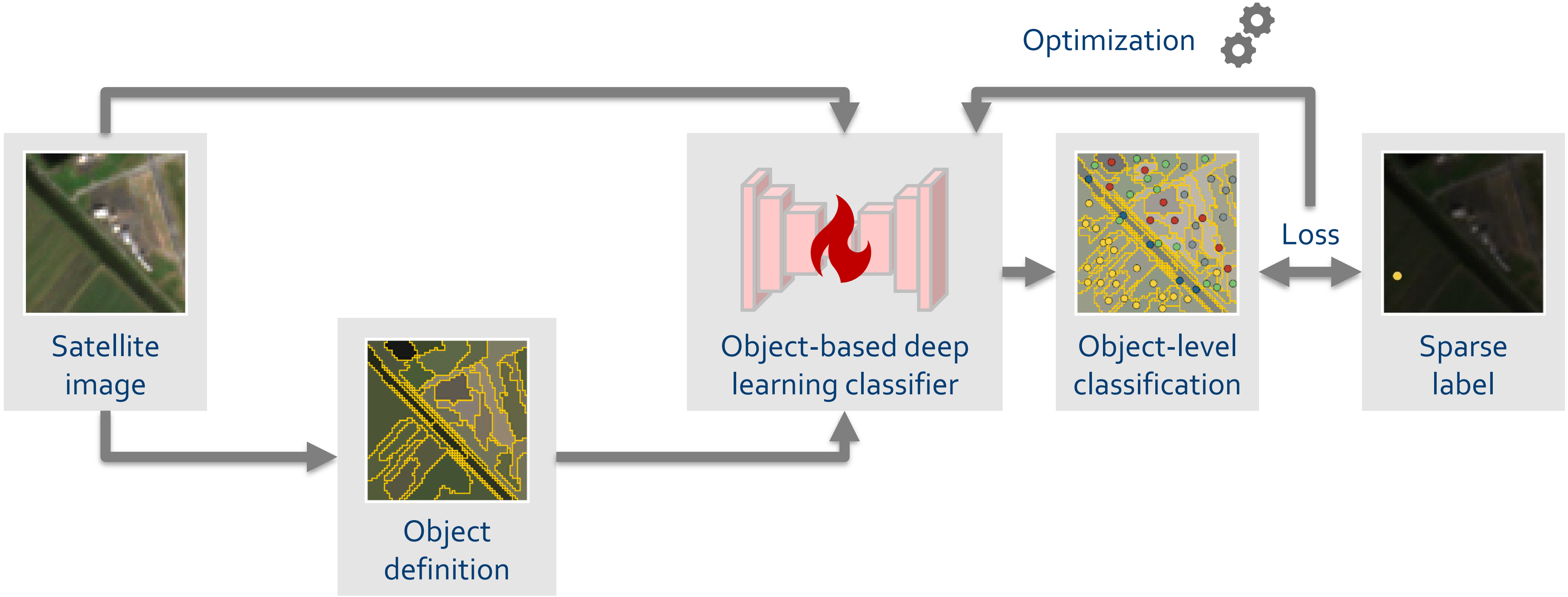
- Uses standard classification cross entropy loss

$$L(\hat{y}, y) = -\sum_{k \in K} y_k \log \hat{y}_k$$

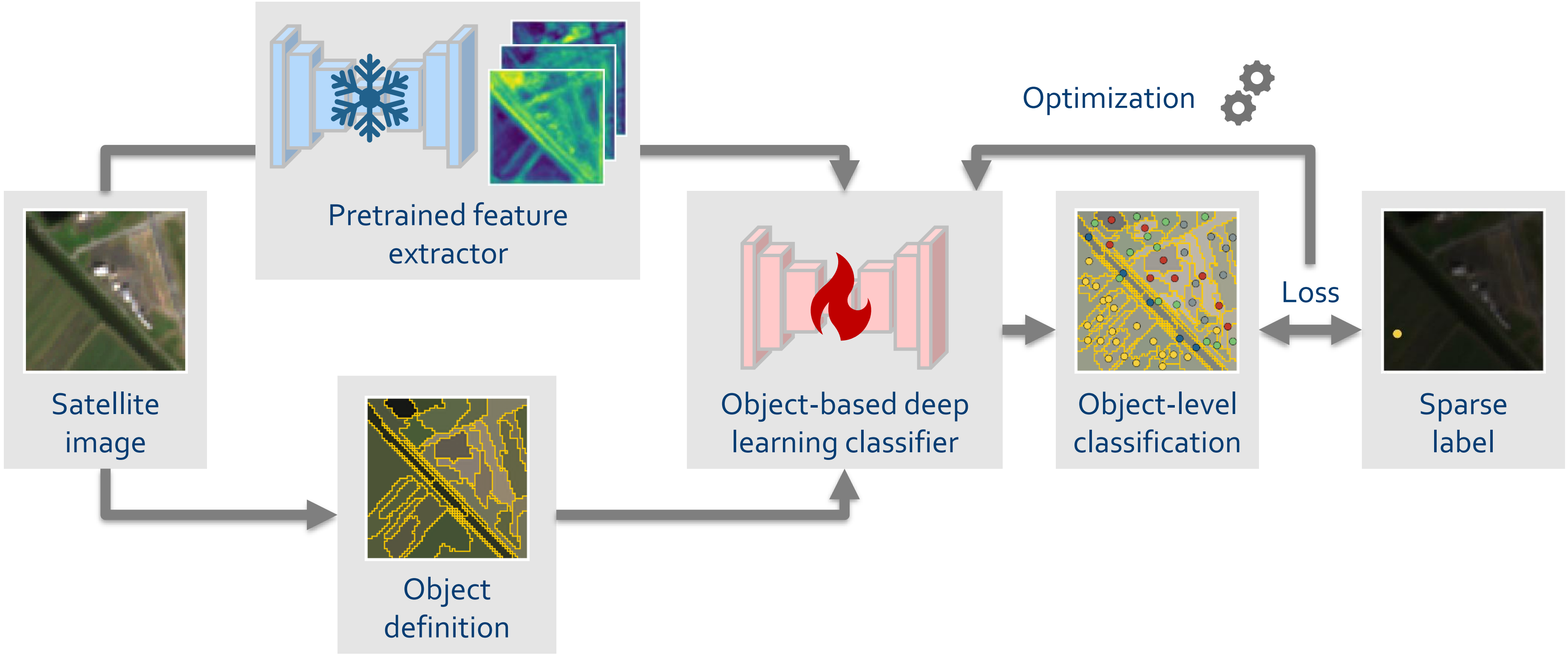
- Only evaluate the prediction at pixels where ground truth labels are given
- Object aggregation leads to effective increase of **portion of labeled data**



OUR FRAMEWORK

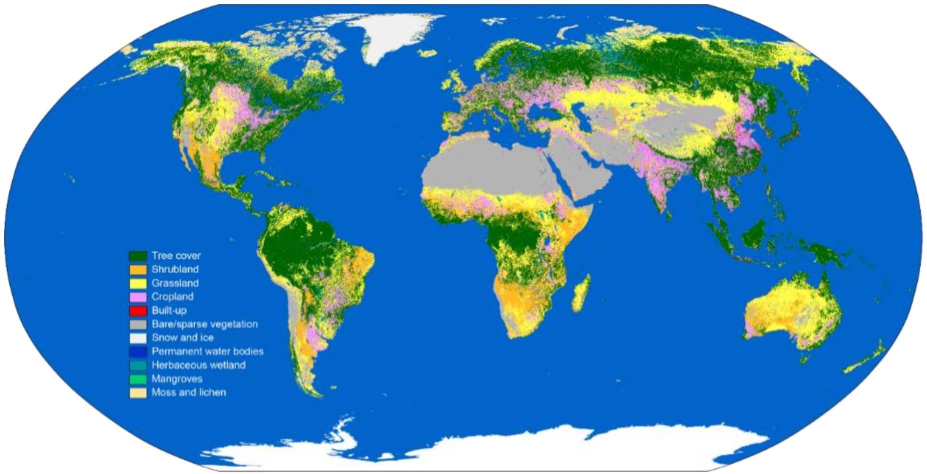


OUR FRAMEWORK



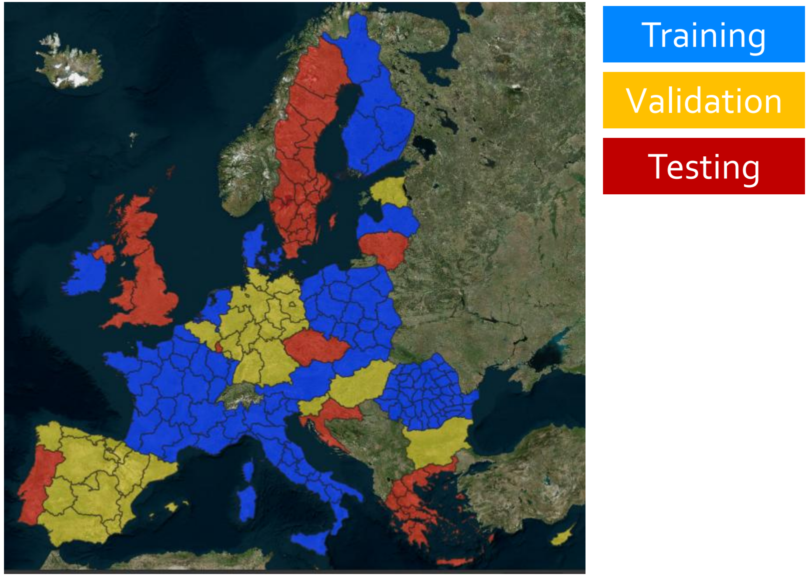
PRETRAINED FEATURE EXTRACTOR

- Use a **pretrained neural network** to extract meaningful features from the satellite images
- Here, pixel-wise model is **pretrained** with **ESA World Cover** data as pseudo-labels
- Use feature maps from **penultimate** layer as inputs to the object-based deep learning classifier

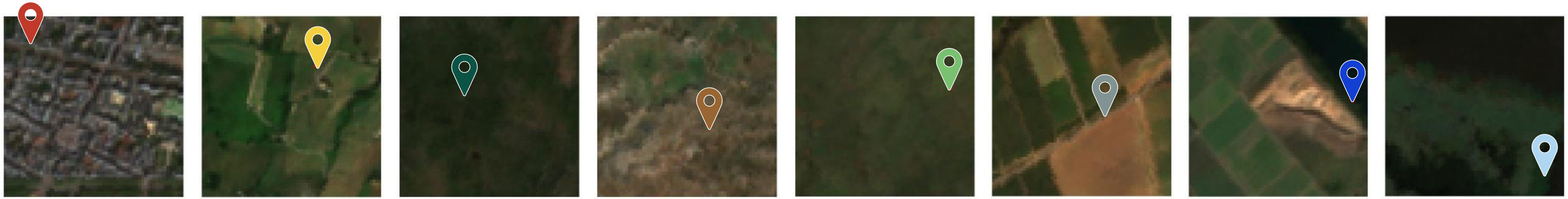


Zanaga et al. (2021)

OUR DATASET

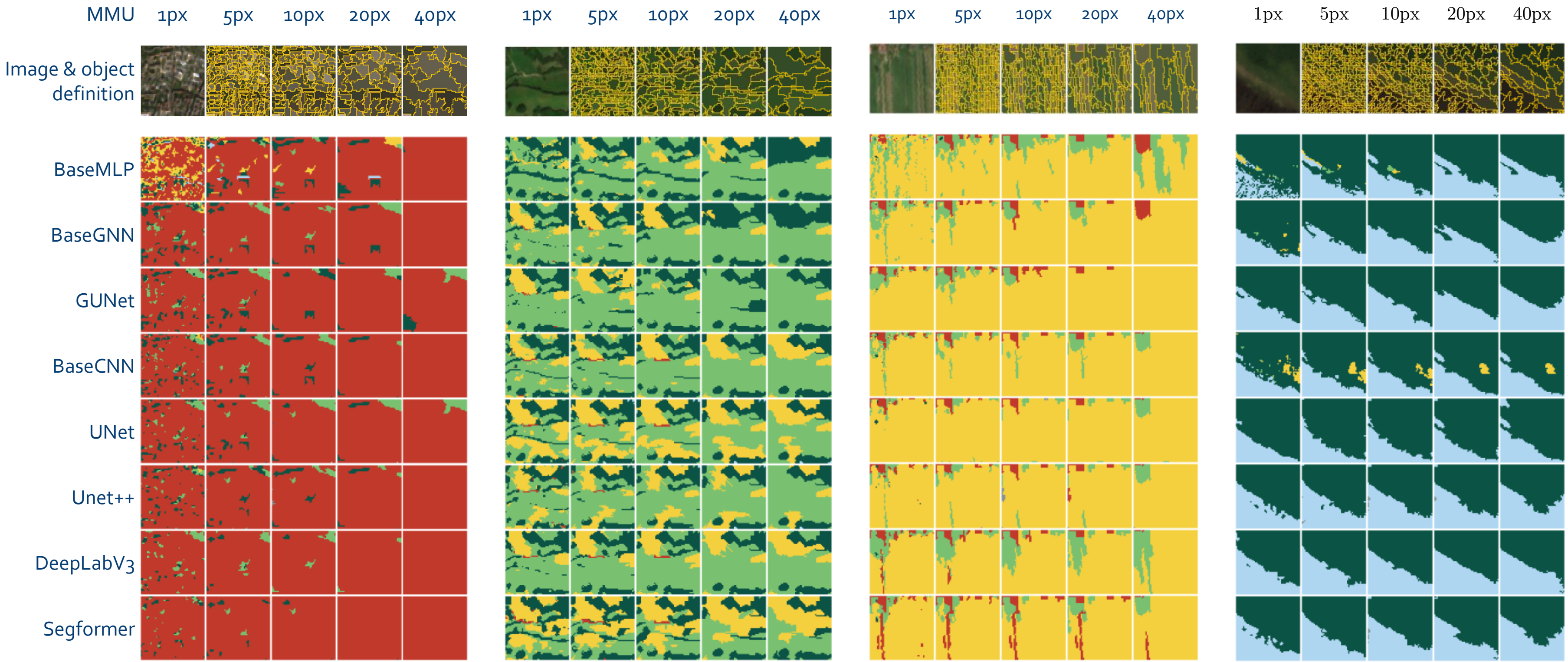


- **Images:** Sentinel-2 (R-G-B-NIR)
- **Sparse labels:** Eurostat's Land Use/Cover Area frame Survey (LUCAS)
- **337854 samples** from **2018**



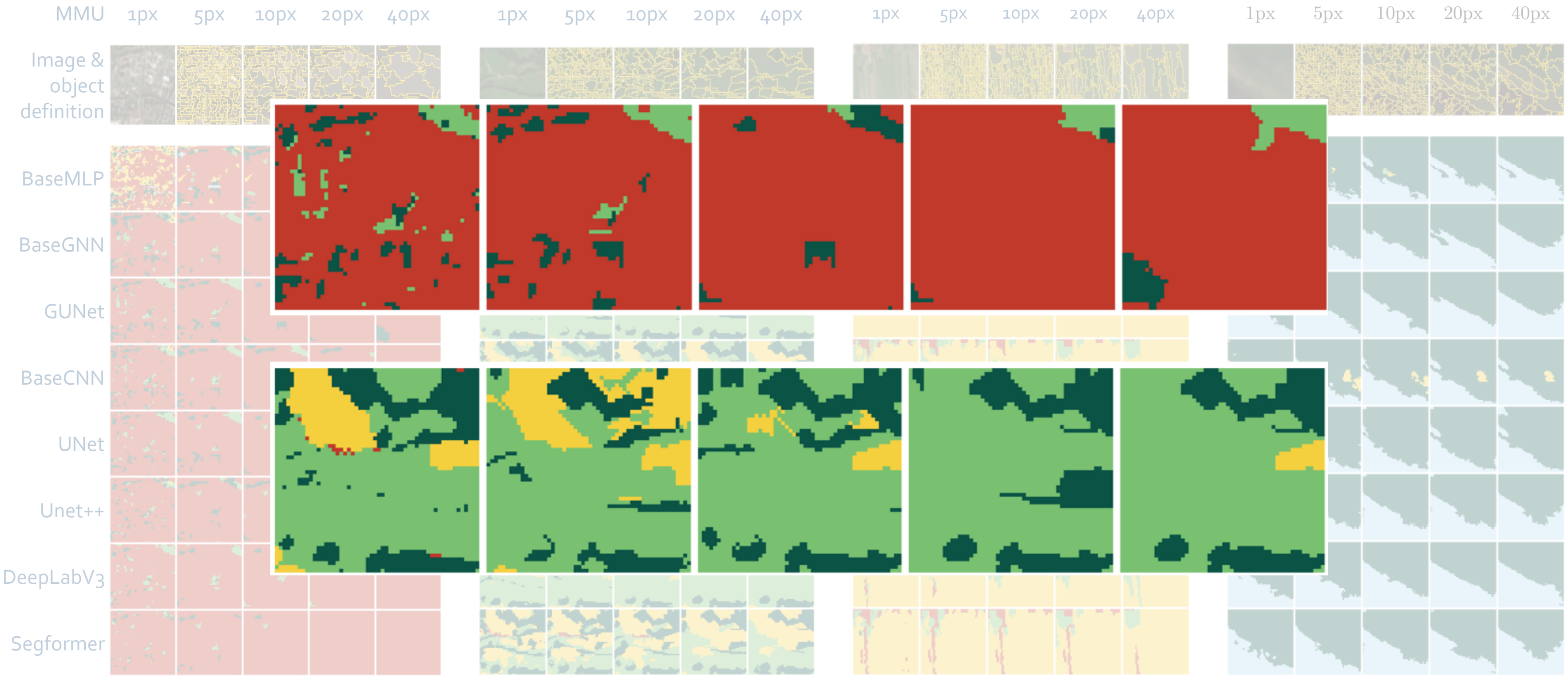
RESULTS

Effect of MMU



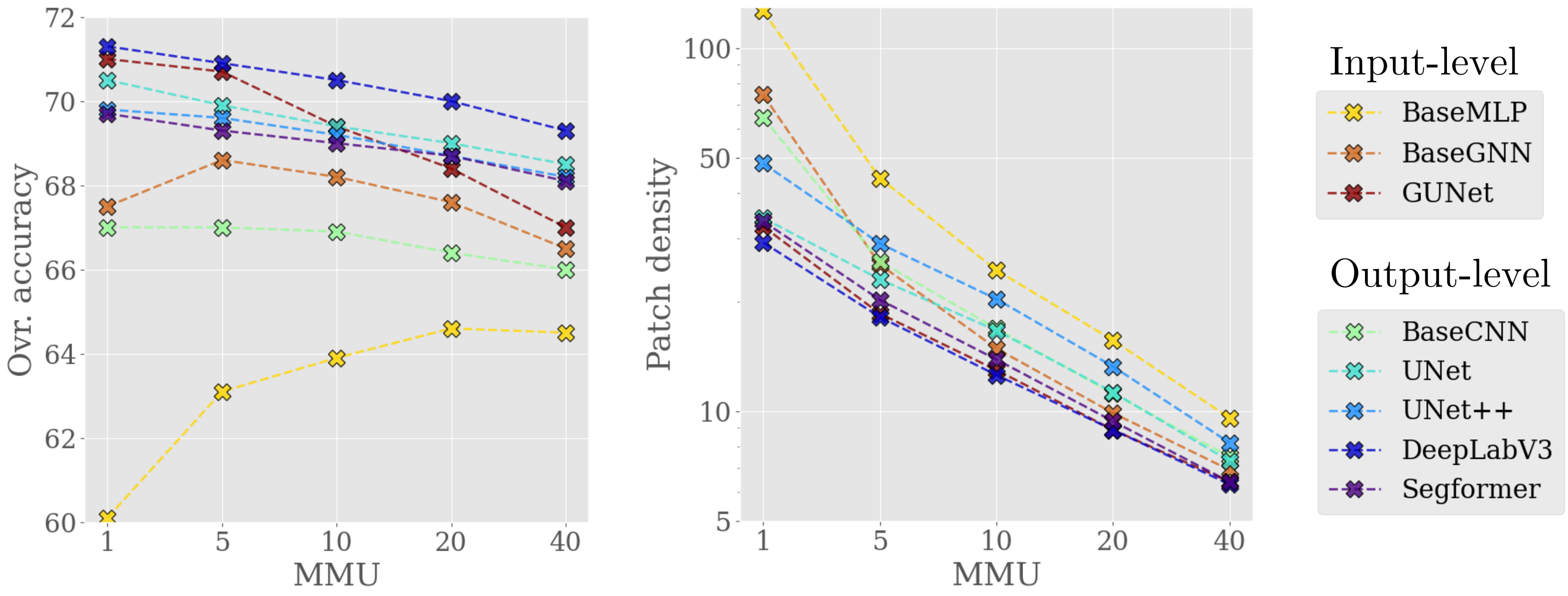
RESULTS

Effect of MMU



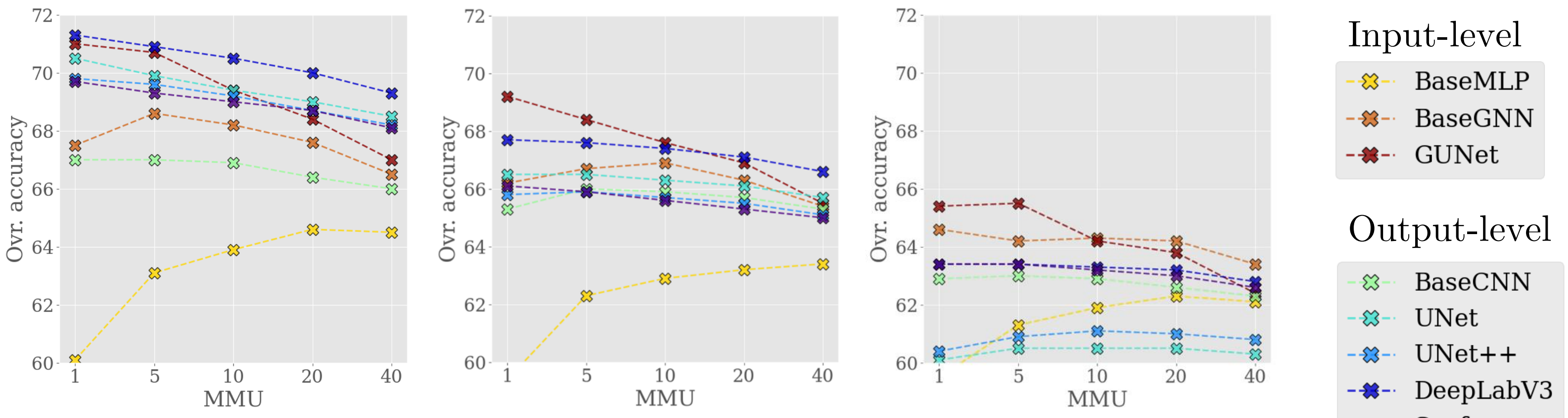
RESULTS

Effect of MMU



RESULTS

Effect of dataset size

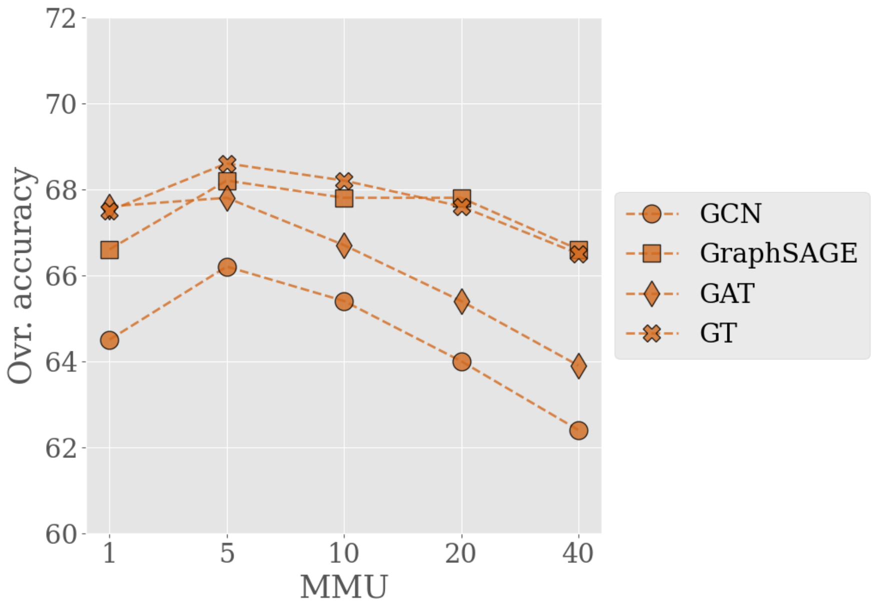


Full dataset

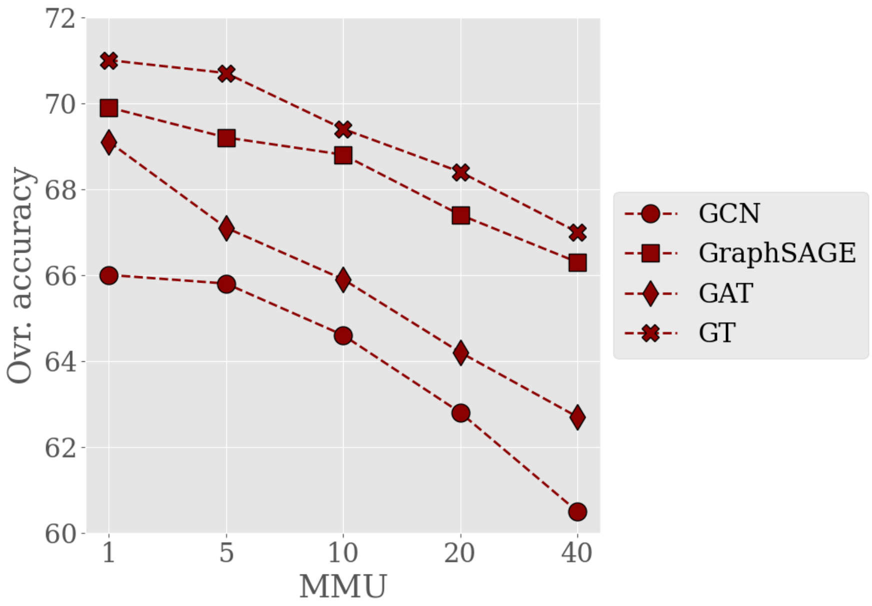
1/4 dataset

1/16 dataset

Ablation studies – Input-level object aggregation



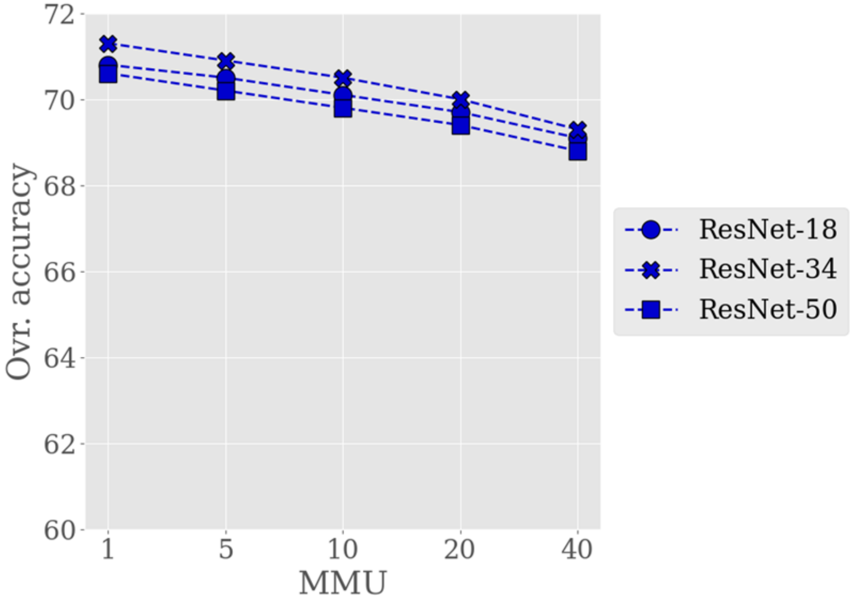
BaseGNN



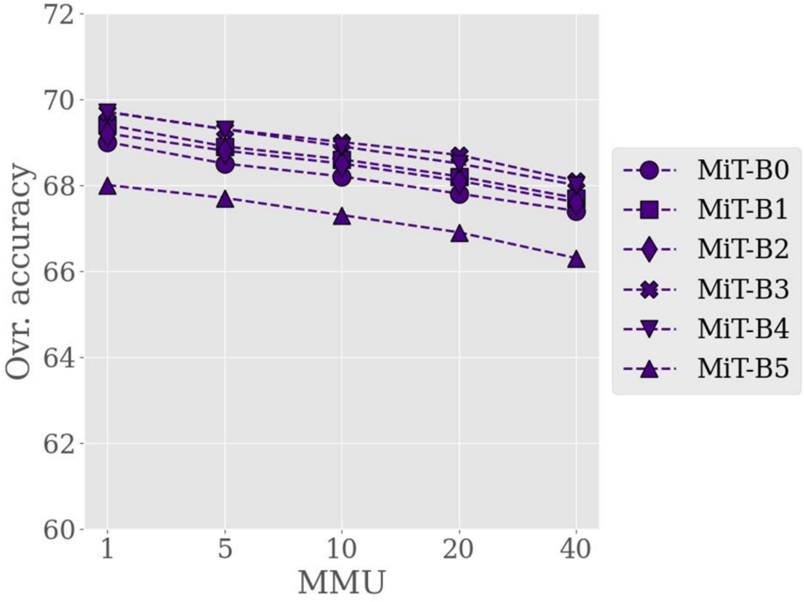
GUNet

RESULTS

Ablation studies – Output-level object aggregation



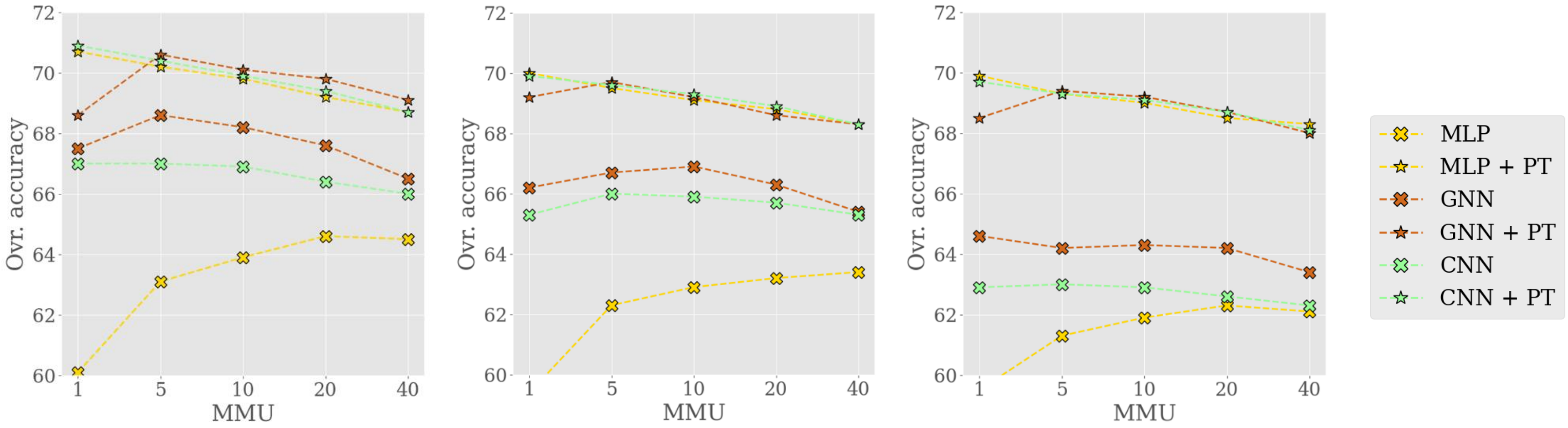
DeepLabV3



Segformer

RESULTS

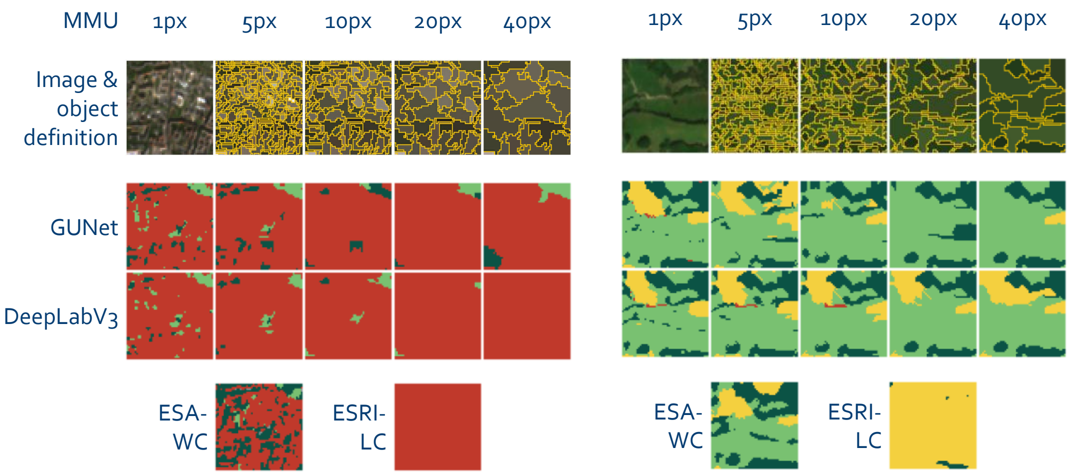
Integration of features from pretrained model

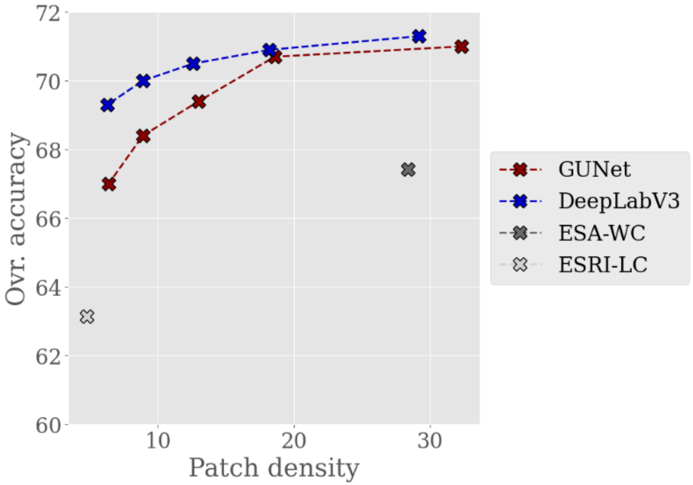


Full dataset 1/4 dataset 1/16 dataset

RESULTS

Comparison to third-party land cover products



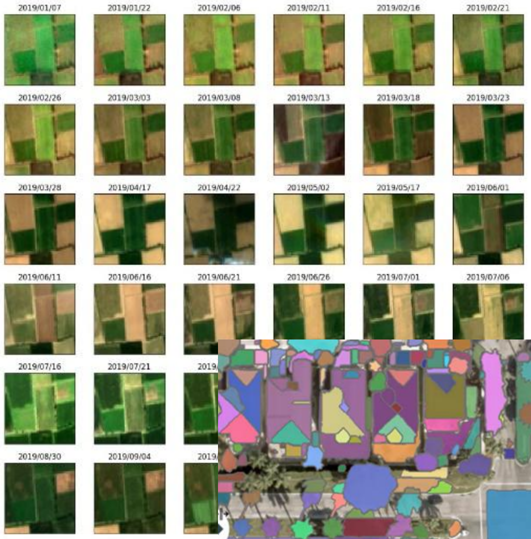


KEY TAKEAWAYS

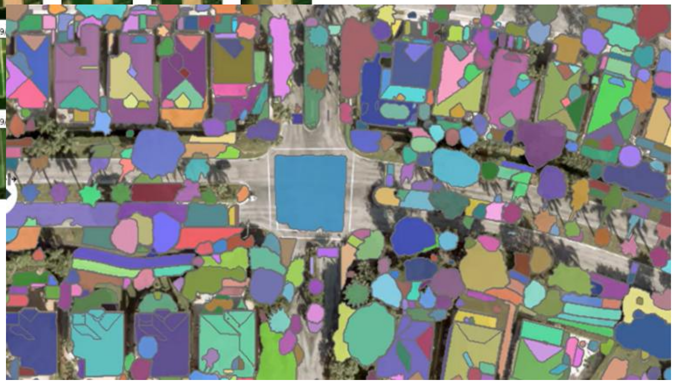
- **Object-based deep learning** methods provide more **coherent** maps with similar **accuracies** as pixel-wise methods.
- **Input-level aggregation** works better for small datasets; **output-level aggregation** works better for large datasets.
- **Features from pretrained models** improve accuracy and reduce data demand.
- Produced maps are more accurate than **existing large-scale products**.

FUTURE WORK

- Generalization of the object-based classification framework to **satellite image time series**
- Assess the suitability of **foundation models** into the framework
  - **Feature extractors** (e.g. Prithvi-EO, AlphaEarth)
  - **Object definition** (SAM)



(Image: Mattia Gatti)

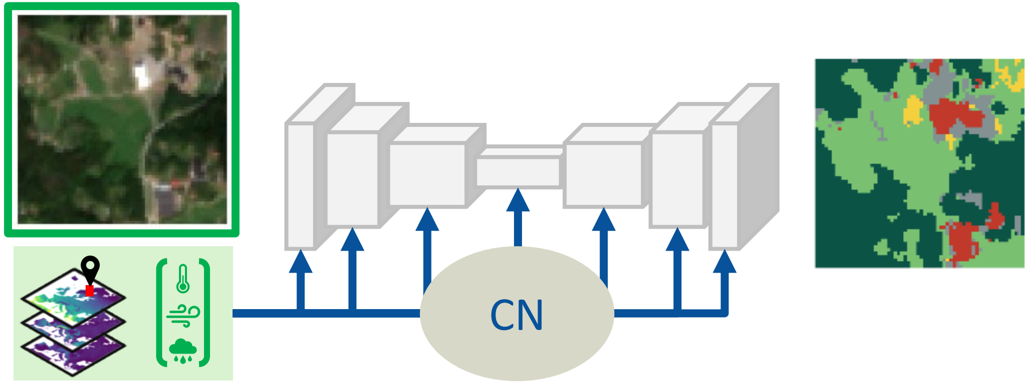


(Image: ESRI)

OTHER WORK

Climate-conditional land cover classification

- Integrate **climatic auxiliary data** into neural networks for land cover classification to tackle **geographic data shift**
- Use **climate-conditional normalization** (CN) layers
- Improves accuracy of CNN and transformer-based classifiers



OTHER WORK

Climate-conditional satellite image editing

- Use **diffusion autoencoder** model to simulate **climatic variation** in satellite imagery
- Can be used for **data augmentation**, leading to more **accurate** and **generalizable** classification when training on geographically limited datasets