ANOTHER EARTH

Synthetic Geospatial Data for Training AI-enabled Downstream Tasks

Maya Pindeus
Co-Founder, CEO

Felix Geremus
Co-Founder, CTO

Lydia Abady
Lead AI Researcher

International workshop on AI Foundation Model for EO

5-7 May 2025 | ESA ESRIN - Frascati, Italy

Motivation: Challenges in Remote Sensing and the Role of Synthetic Data

Data Scarcity and High Annotation Costs

- Labeled remote sensing data is limited and expensive.
- This restricts the amount of training data and affects model accuracy.

Domain Variability and Limited Ground Truth

- Environmental heterogeneity leads to domain shifts.
- Sparse ground truth hampers generalization across regions.

Synthetic Data Generation

- Provides scalable, labeled samples.
- Helps address data scarcity.
- Improves performance in tasks like segmentation.

Proposed Workflow: Synthetic Data for Remote Sensing Segmentation

Generative Model

Train a generative model (e.g., GAN, diffusion model) to learn spatial and spectral patterns.

Output: Synthetic images with diverse and realistic characteristics.

Synthetic Labeled Data

Simulate various land cover types, environmental conditions, and domain shifts.

Downstream Task

Train segmentation models: use the synthetic dataset to augment scarce labeled datasets to train deep segmentation models (e.g., U-Net, DeepLab)

Datasets

ESA WorldCover 2.0

- **Type:** Global Land Cover classification map
- **Resolution:** 10m GSD
- **Year:** 2021 release
- **Classes:** 11 land cover classes
- **Subset Used:**
  - Extracted **New York region**
- **Tiled to:** 512×512 pixels
- **Usage**
  - Train generative model to generate land covers
  - Train conditional generative models to generate optical images conditioned on land covers

NAIP (New York subset)

- **Type:** Aerial imagery
- **Resolution:** 1 meter GSD
- **Used Bands:** RGB
- **Acquisition:** 2021–2022
- **Subset Used:**
  - New York State region
- **Number of images used:** 126,000 tiles
- **Tile to:** 512×512 pixels
- **Usage**
  - Train conditional generative models to generate optical images conditioned on land covers

Generative Architectures

Two types of generation pipelines

Unconditional Generation – Land Cover Maps

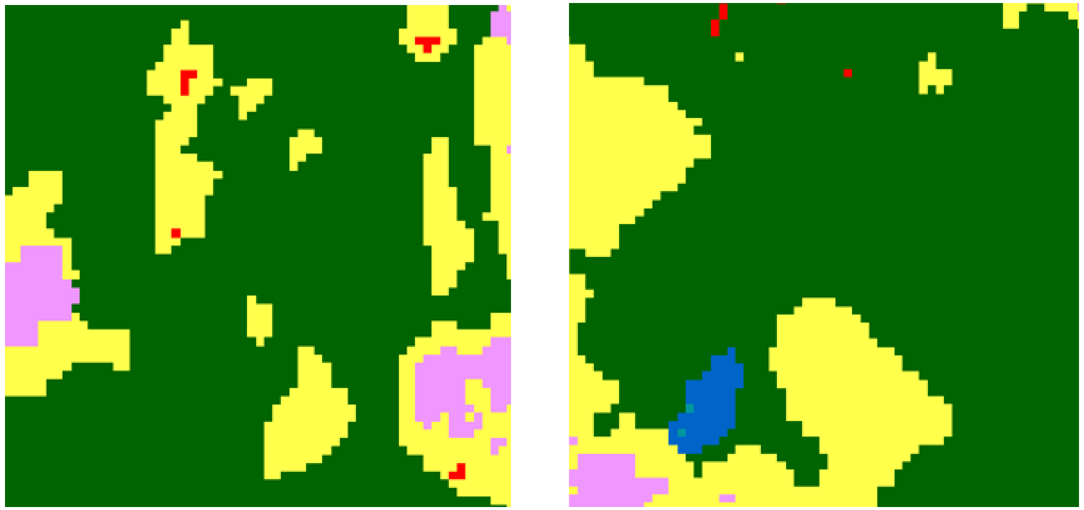
- **Model:** Latent Diffusion
- **Objective:** Generate semantic land cover maps
- **Loss Function:** Cross-Entropy

Conditional Generation – Guided Image Synthesis

- **Architectures**:
  - SPADE
  - ControlNet
- **Objectives**: Generate realistic, high resolution optical images conditioned on landcover

Unconditional Image Generation – Land Cover

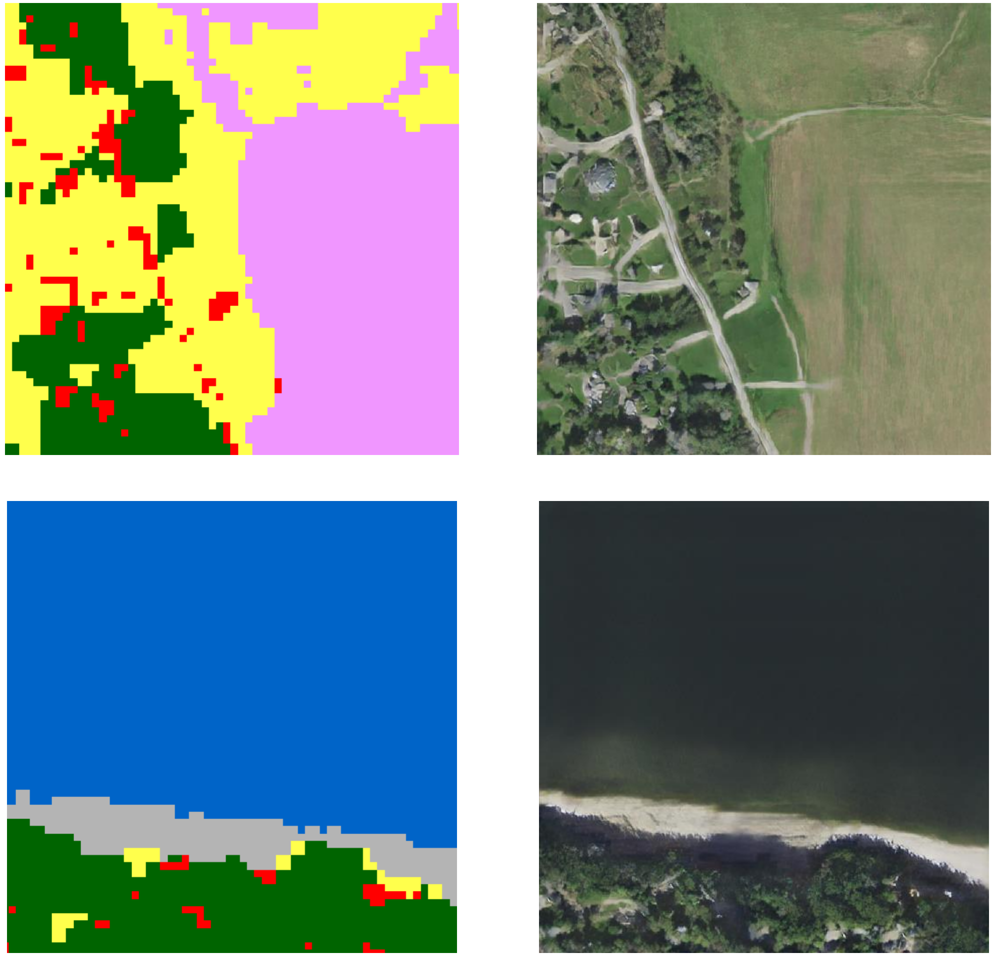
Real Land Cover



Synthetic Land Cover

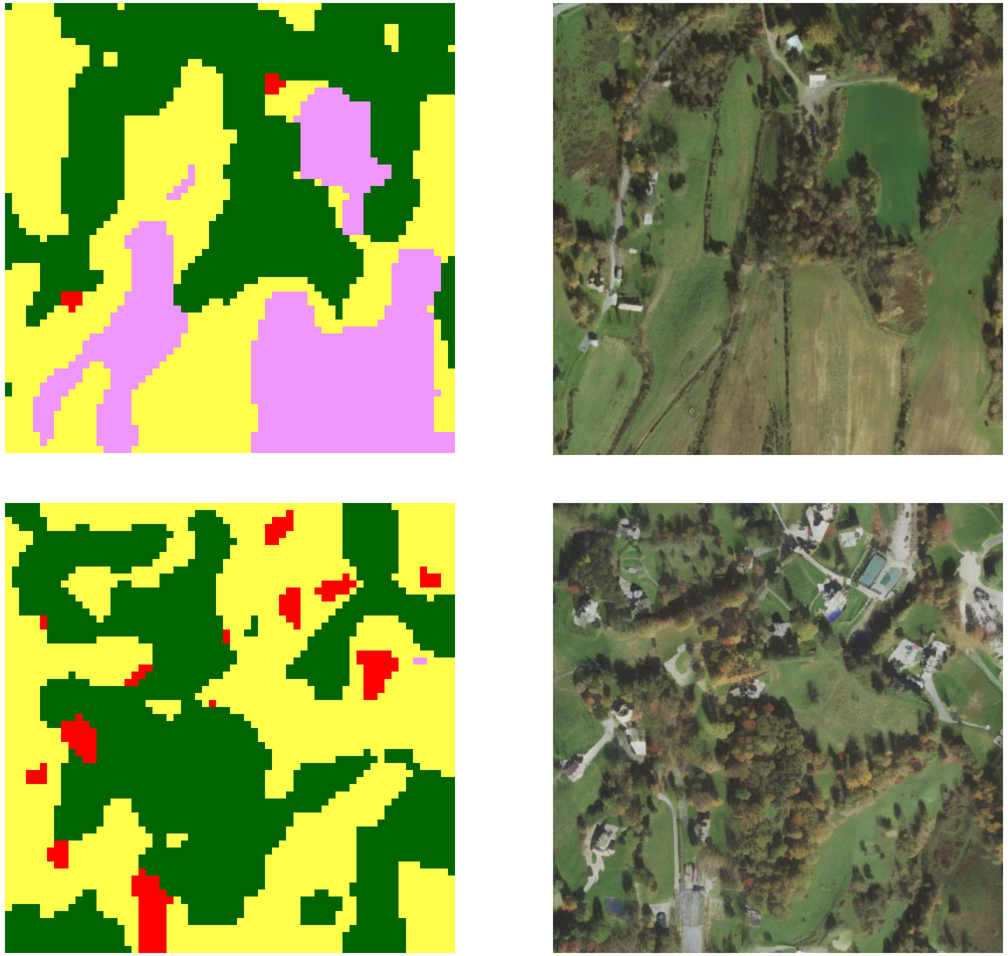
Conditional Image Generation – Spade & ControlNet

SPADE Conditional Generation



Input Output

ControlNet Conditional Generation



Input Output

Segmentation Architectures

- **Architectures**: UNet, UNet++, and DeepLabV3+
- **Dataset**
  - 4 sets were used
    - **Realrgbreallc**: Real optical images with real land covers
    - **Fakergbreallc**: conditionally generated optical images with real land cover used as conditional input to the trained controlNet
    - **Fakergbreallcspade**: conditionally generated optical images with real land cover used as input to the trained SPADE network
    - **Fakergbfakelc**: conditionally generated optical images with synthetically generated land covers used as input to the trained controlNet
  - 8,000 images for training
  - 2,000 images for testing
- **Evaluation Metric**: Intersection over Union (IoU)
- All trained segmentation models were cross tested with each other

Segmentation Results

		Test Dataset			
Training datasets	**Segmentation Architecture**	**realrgbreallc**	**fakergbreallc**	**fakergbreallcspade**	**fakergbfakelc**
realrgbreallc	**Unet**	0.76	0.73	0.73	0.73
	Unet++	0.75	0.73	0.73	0.73
	Deeplabv3+	0.74	0.73	0.74	0.73
fakergbreallc	**Unet**	0.7	0.79	0.7	0.69
	Unet++	0.7	0.79	0.7	0.67
	Deeplabv3+	0.69	0.79	0.67	0.67
fakergbreallcspade	**Unet**	0.46	0.43	0.94	0.3
	Unet++	0.47	0.41	0.95	0.27
	Deeplabv3+	0.46	0.4	0.92	0.27
fakergbfakelc	**Unet**	0.72	0.7	0.73	0.76
	Unet++	0.71	0.68	0.73	0.77
	Deeplabv3+	0.71	0.68	0.73	0.77

Key Insights

- **Real-to-Fake Generalization:** Models trained on real data generalize well to synthetic data

- **Fake-to-Real Generalization:** training on the synthetic optical images conditioned on either real of synthetic land covers generalize well on real data which is useful in the data scarce domains

- **Overfitting:** segmentation models trained on data generated using the SPADE architecture overfitted to the data generated by SPADE and did not generalize well to the other datasets

Conclusion & Future Direction

Conclusion: What Did We Aim to Learn?

- **Objective**: Investigate whether **synthetic data** can effectively support training of **downstream tasks** like segmentation.
- **Key Insight**: Synthetic data can **generalize well** and in some cases **replace or supplement real data**.

Future Directions

- **Data Mixing Experiments**: Explore **combinations of real and synthetic data** during training to study performance trade-offs and benefits.
- **Fine-tuning Scenarios**: fine-tuning of the generative model for **target-specific domains**.
- **Scaling Up Generation**: Train models on broader geographies and modalities to **increase realism and diversity** of generated datasets.

THANK YOU!

Q&A

Lydia Abady | Another Earth | lydia@anotherearth.ai