Introduction to the GeoAI Foundation Model and Its Application

Yanhong Huang
Salar Jarhan
Jiyoung Lee
Claire Simpson
Yalin Yang

Summer School 2023: Convergence Science in Action

UCAR Campus in Boulder, Colorado • August 7-11, 2023

Apply Summer School 2024: https://iguide.illinois.edu/summer-school/summer-school-2024/

Team 6

TeamAwesome

Dr. Eric Shook

University of Minnesota
Twin Cities

Dr. Diana Sinton

UCGIS & Cornell University

Yanhong Huang

University of New Mexico

Salar Jarhan

University of Wyoming

Jiyoung Lee

Louisiana State University

Claire Simpson

University of Colorado, Boulder

Yalin Yang

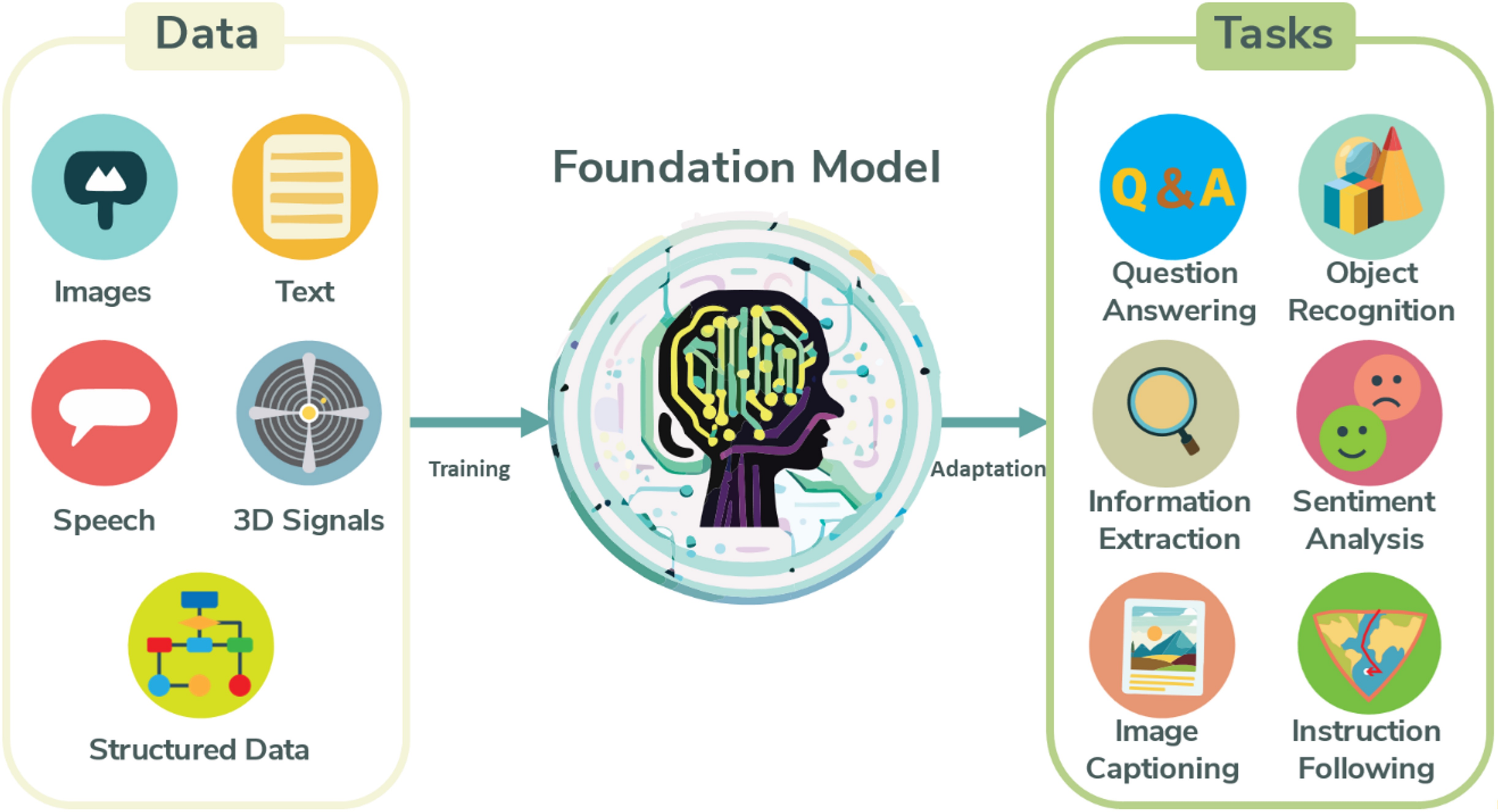
University of Texas at Dallas

Geospatial Artificial Intelligence (GeoAI)

- Utilization of AI in conjunction with **geographical data, science, and technology**
- Accelerate the comprehension of real-world business prospects, environmental effects, and operational threats.

(https://www.esri.com/en-us/capabilities/geoai/overview#:~:text=Geospatial%20artificial%20intelligence%20(GeoAI)%20is,environmental%20impacts%2C%20and%20operational%20risks.)

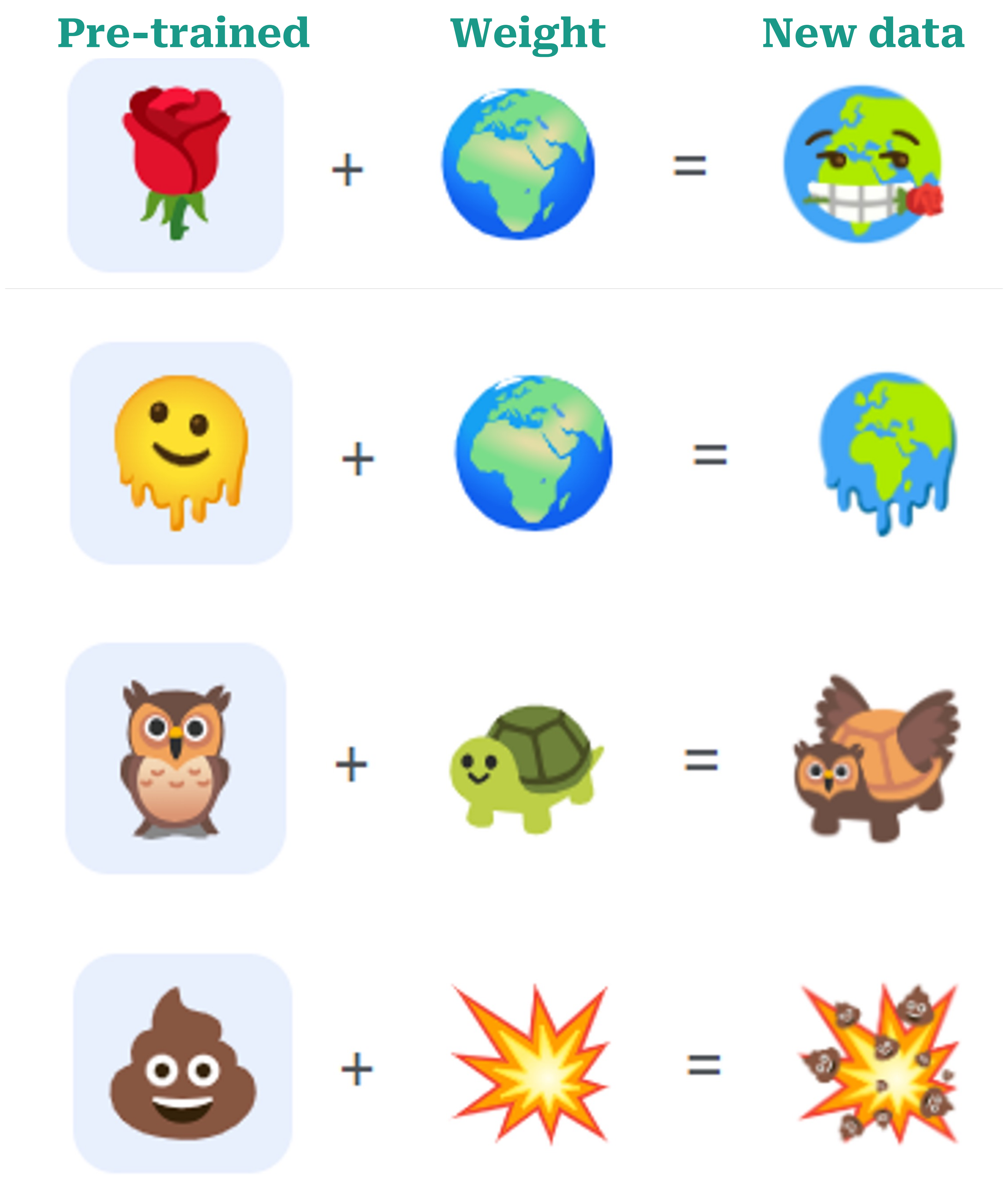
Foundation Model



(Adpated from https://blogs.nvidia.com/blog/2023/03/13/what-are-foundation-models/)

Fine-tuning

- An approach to transfer learning
- The weights of a pre-trained model are trained on new data

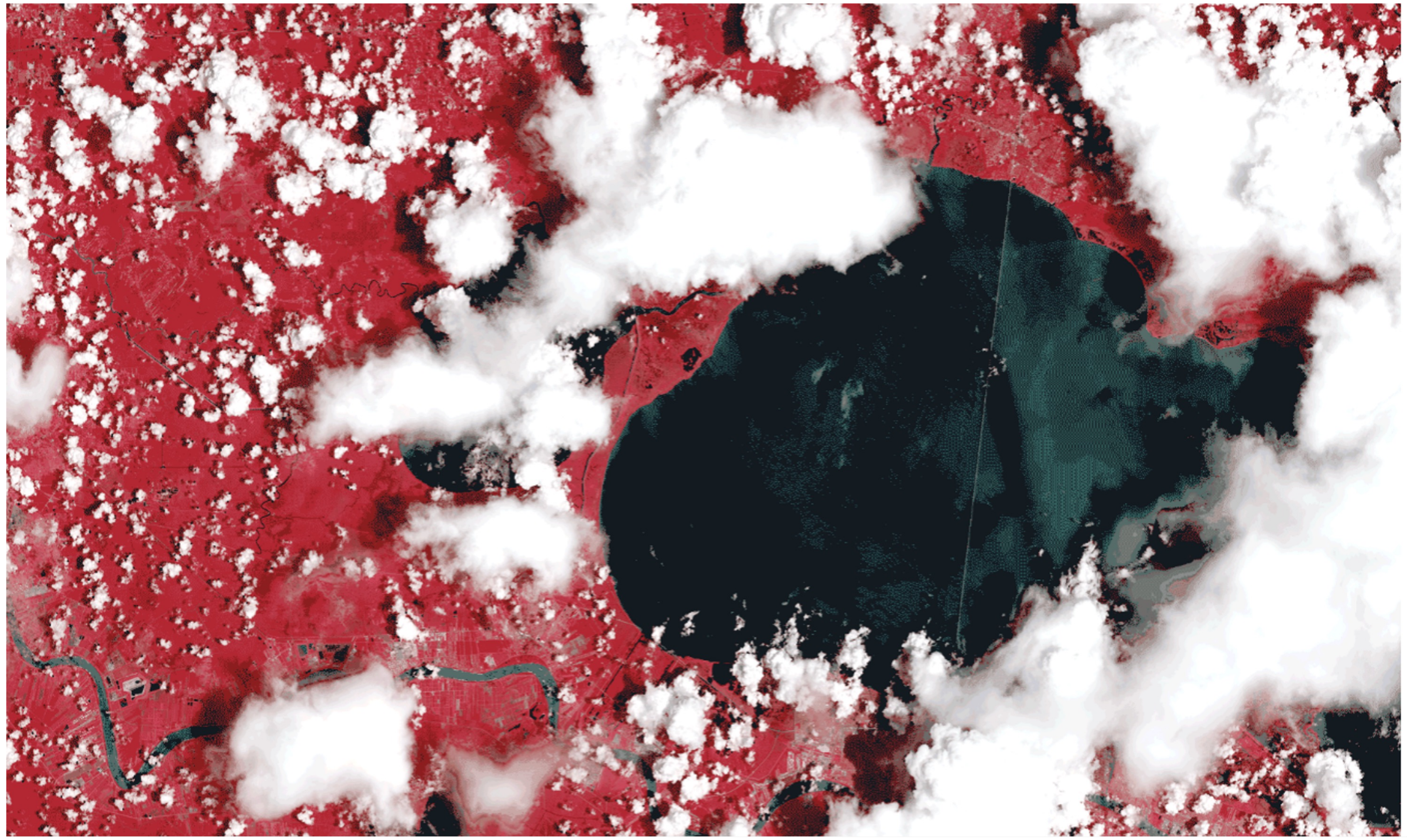


Emoji Kitchen: Cook Moji

Background

Gain **practical experience** in applying **deep learning techniques** to real-world spatial problems, at the **cutting edge of GeoAI**

"NASA's first open-source geospatial artificial intelligence (AI) foundation model for Earth observation data...is a ***milestone in the application of AI for Earth science.****"*

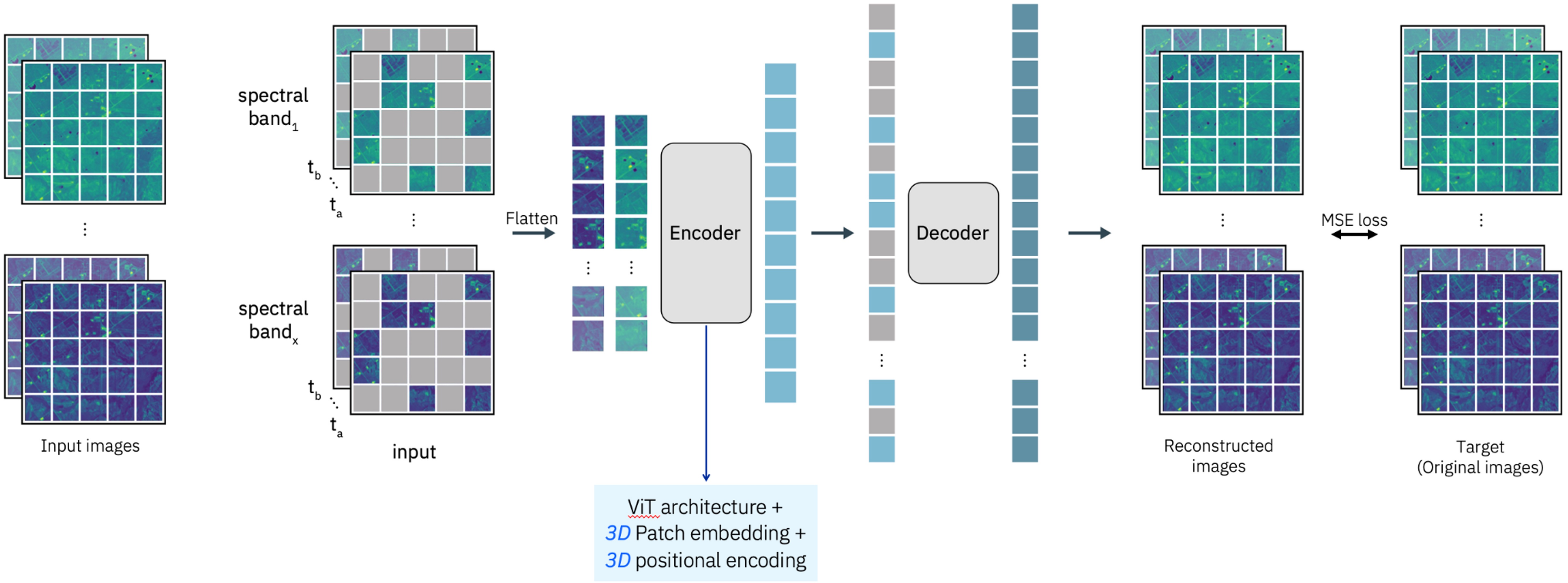


(https://www.earthdata.nasa.gov/news/impact-ibm-hls-foundation-model)

Prithvi-100M Model

- **Type:** temporal Vision transformer
- **Pre-Training:** pre-trained by the IBM and NASA team on contiguous US Harmonised Landsat Sentinel 2 (HLS) data.
- **Architecture:**
  - Self-supervised encoder developed with a ViT architecture and Masked AutoEncoder (MAE) learning strategy, MSE loss function.
  - spatial attention across multiple patches temporal attention for each patch.

Prithvi-100M Model



(Source: https://huggingface.co/ibm-nasa-geospatial/Prithvi-100M/blob/main/README.md)

Data

- Harmonized Landsat-Sentinel-2 (HLS)

<table>
<tr><th>Band Name</th><th>Data type</th><th>Scale</th><th>Unit</th><th>Saturation Flag</th><th>Spatial Resolution</th><th>Nadir Adjustment</th><th>Description</th></tr>
<tr><td>Coastal aerosol</td><td>int16</td><td>0.0001</td><td>unitless</td><td>12,000</td><td rowspan="9">30</td><td rowspan="7">yes</td><td>Nadir-adjusted</td></tr>
<tr><td>Blue</td><td>int16</td><td>0.0001</td><td>unitless</td><td>12,000</td><td>Surface Reflectance</td></tr>
<tr><td>Green</td><td>int16</td><td>0.0001</td><td>unitless</td><td>12,000</td><td></td></tr>
<tr><td>Red</td><td>int16</td><td>0.0001</td><td>unitless</td><td>12,000</td><td></td></tr>
<tr><td>NIR</td><td>int16</td><td>0.0001</td><td>unitless</td><td>12,000</td><td></td></tr>
<tr><td>SWIR 1</td><td>int16</td><td>0.0001</td><td>unitless</td><td>12,000</td><td></td></tr>
<tr><td>SWIR 2</td><td>int16</td><td>0.0001</td><td>unitless</td><td>12,000</td><td></td></tr>
<tr><td>TIRS B10</td><td>int16</td><td>0.01</td><td>Celsius</td><td></td><td></td><td>TOA BT</td></tr>
<tr><td>TIRS B11</td><td>int16</td><td>0.01</td><td>Celsius</td><td></td><td></td><td>TOA BT</td></tr>
<tr><td>QA</td><td>uint8</td><td colspan="6">refer to the QA table in the User's Guide for bit information</td></tr>
</table>

Research Objectives

1. Deploy foundation model to the **I-GUIDE Platform**
2. Replicate model fine-tuning for **flood, burn, and crop classification**
3. Create our own fine-tuned model to detect **built areas**




Fine-tuning: Flooding

Purpose: Predict flooded vs dry areas

Dataset: Sen1Floods11

- a georeferenced dataset to train and test deep learning flood algorithms
- 446 labeled 512x512 chips
- 14 biomes, 357 ecoregions, and 6 continents of the world 11 flood events.
- bands for flood mapping: Blue, Green, Red, Narrow NIR, SWIR 1, SWIR 2

Fine-tuning: Flooding

Model Overview:

Pretrained Model: Prithvi-100m.

Fine-tuning Purpose: Segmenting flood extent on Sentinel-2 images.

Label Categories:

Class 0: No water.

Class 1: Water/Flood.

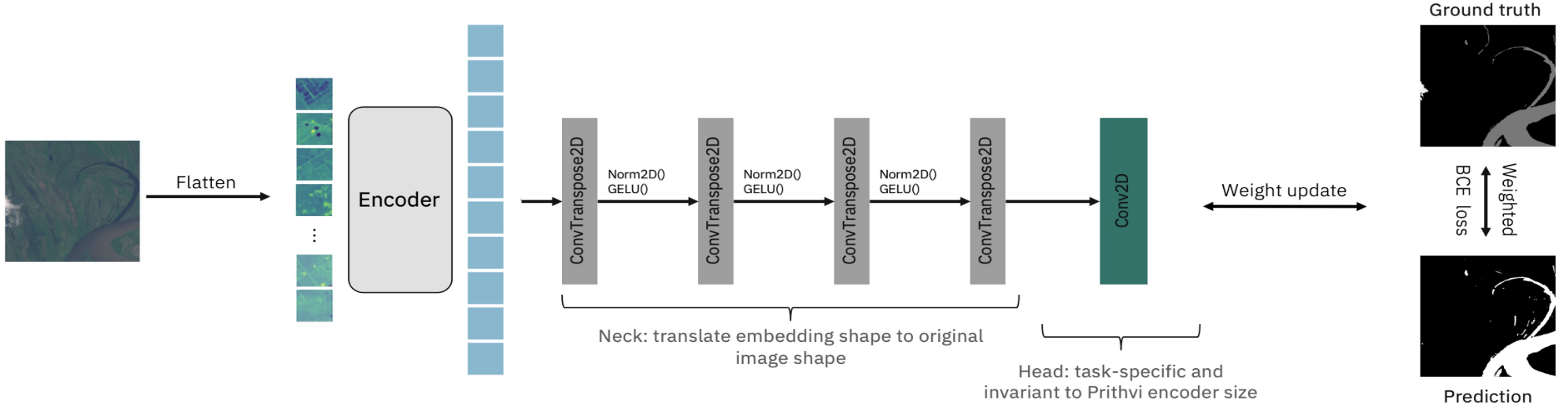
Class 2: No data/Clouds.

Model Training Details:

Initial Pretraining: 3 timesteps sequence length.

Focus: Single-timestamp segmentation.

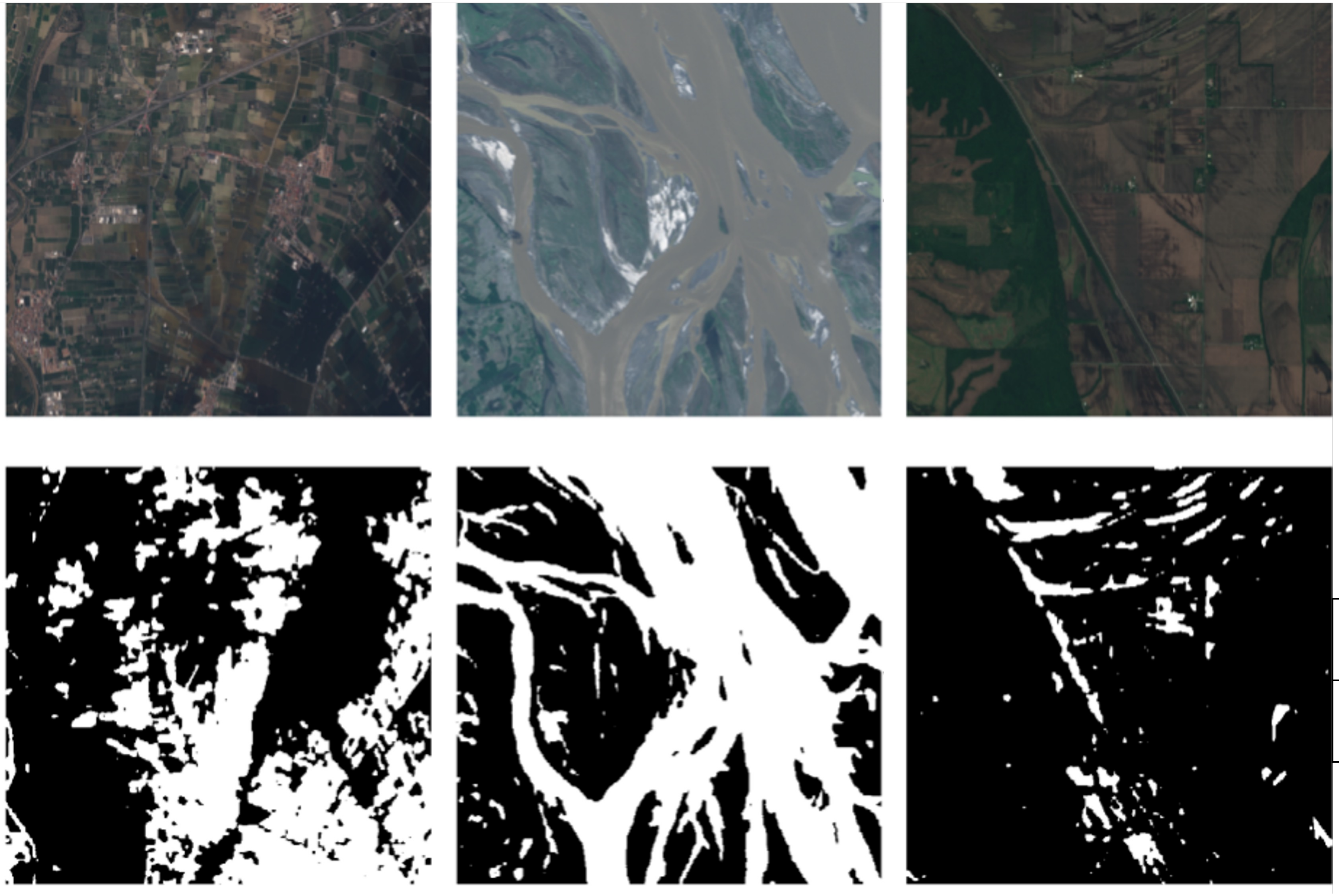
Capability: Adaptable to an arbitrary number of timestamps during finetuning.



https://huggingface.co/ibm-nasa-geospatial/Prithvi-100M-sen1floods11

Fine-tuning: Flooding

Inputs (Raw Imagery)



No water
Water / Flood

Outputs (Binary Classification)

Results

Finetuning the geospatial foundation model for 100 epochs leads to the following performance on the test dataset:

Classes	IoU	Acc
No water	96.90%	98.11%
Water/Flood	80.46%	90.54%

aAcc	mIoU	mAcc
97.25%	88.68%	94.37%

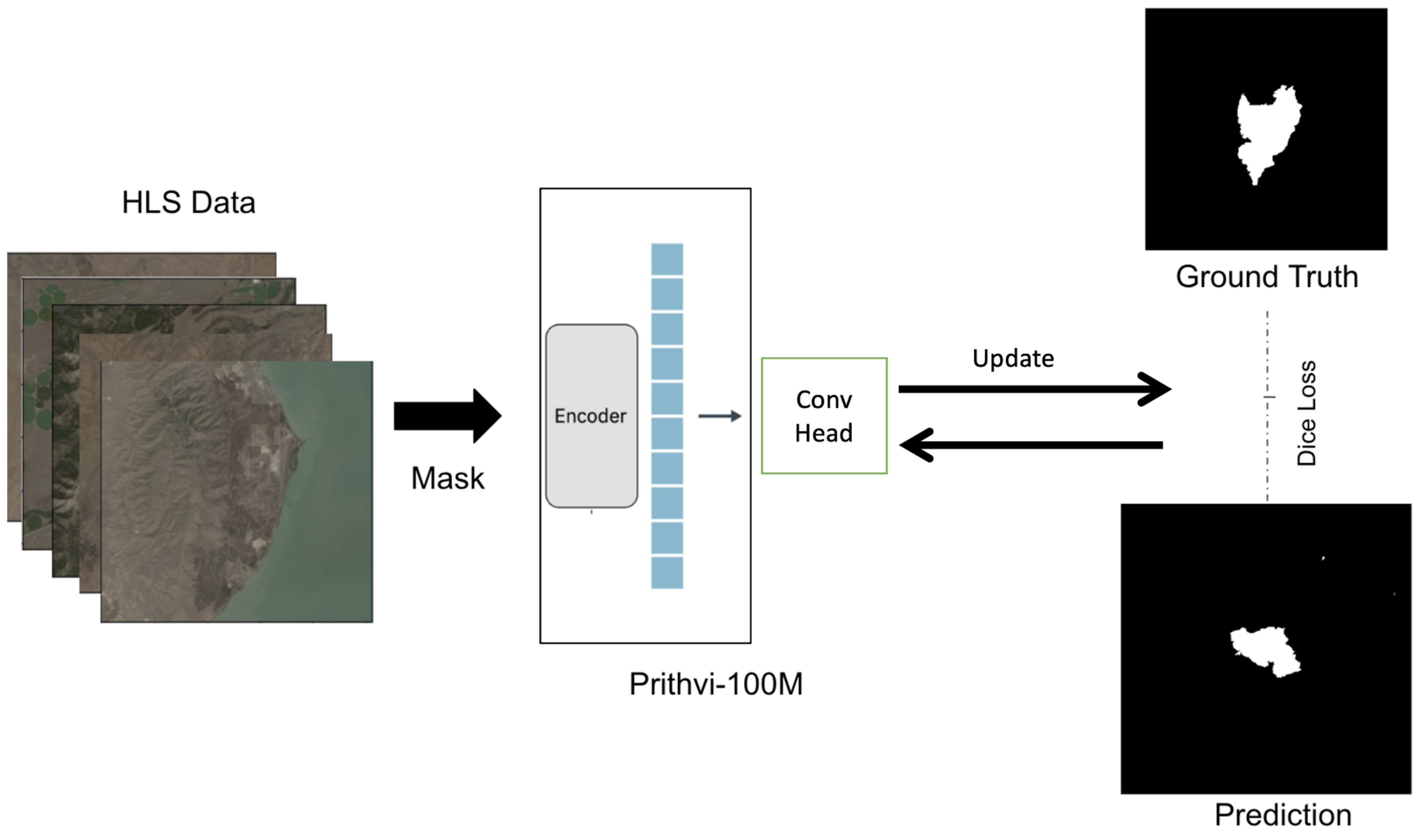
Fine-tuning: Burn Scar

Purpose: Predict burned vs unburned areas after a fire

Data: 1-timestep HLS satellite image GeoTIFFs

- 512 x 512 pixels
- Each tif file contains 6 spectral bands
- Bands: Blue, Green, Red, Narrow NIR, SWIR 1, SWIR 2
- Each GeoTIFF file for the mask contains one band, where each pixel represents the target binary class

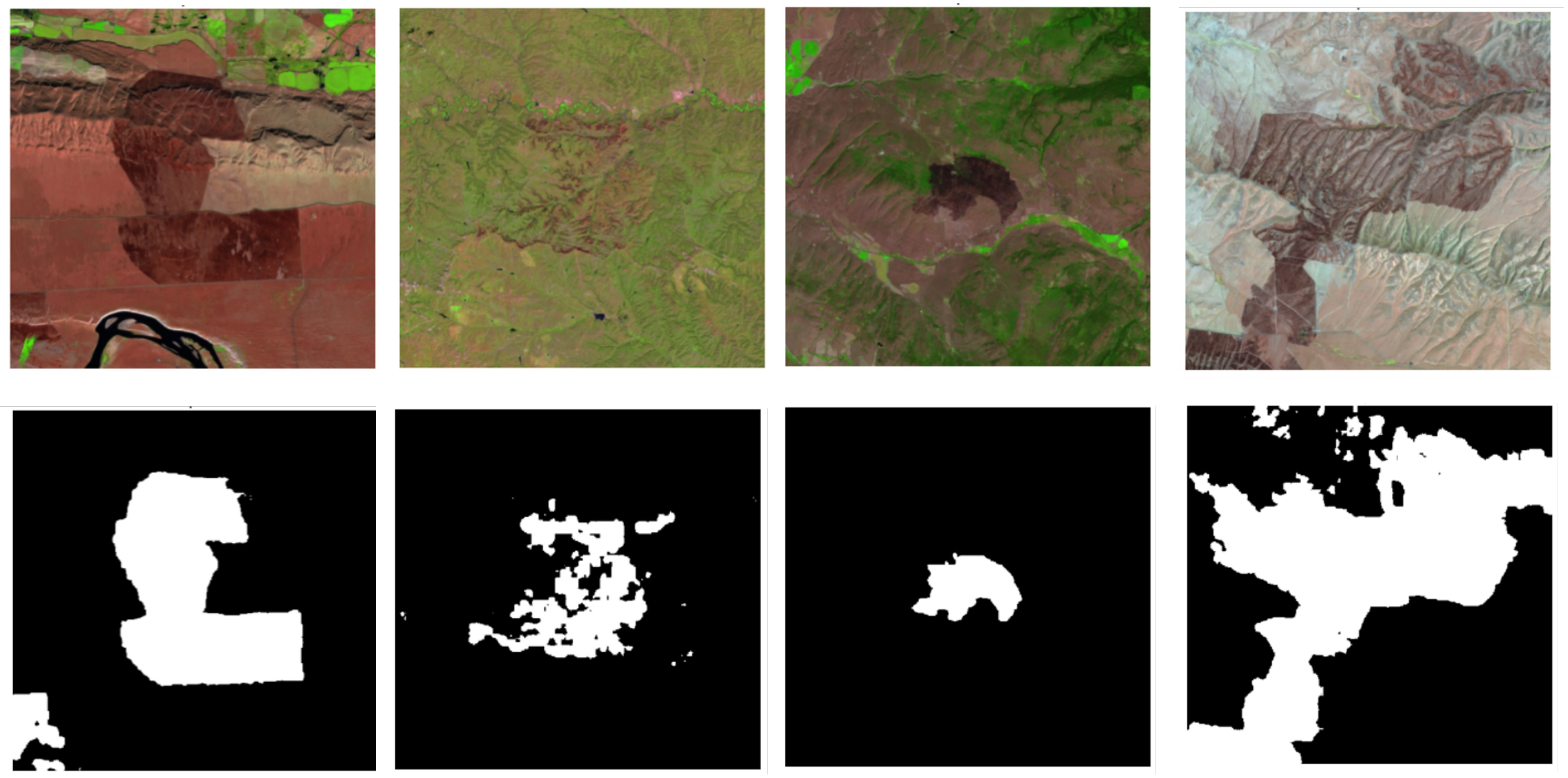
Fine-tuning: Burn Scar



- **Model Overview:**
  - Pretrained Model: Prithvi-100m.
  - Fine-tuning Purpose: Segmenting burn scars.
- **Label Categories:**
  - Class 0: Unburnt land.
  - Class 1: Burn scar.
  - Class 2: No data/Clouds.

Fine-tuning: Burn Scar

Inputs(Raw Imagery)



Outputs (Binary Classification)

Unburnt land

Burn scar

Results

Finetuning the geospatial foundation model for 50 epochs leads to the following performance on the test dataset:

aAcc	mIoU	mAcc
96%	76%	94.37%

Fine-tuning: Crop Classification

Purpose: Predict 12 crop/landcover types.

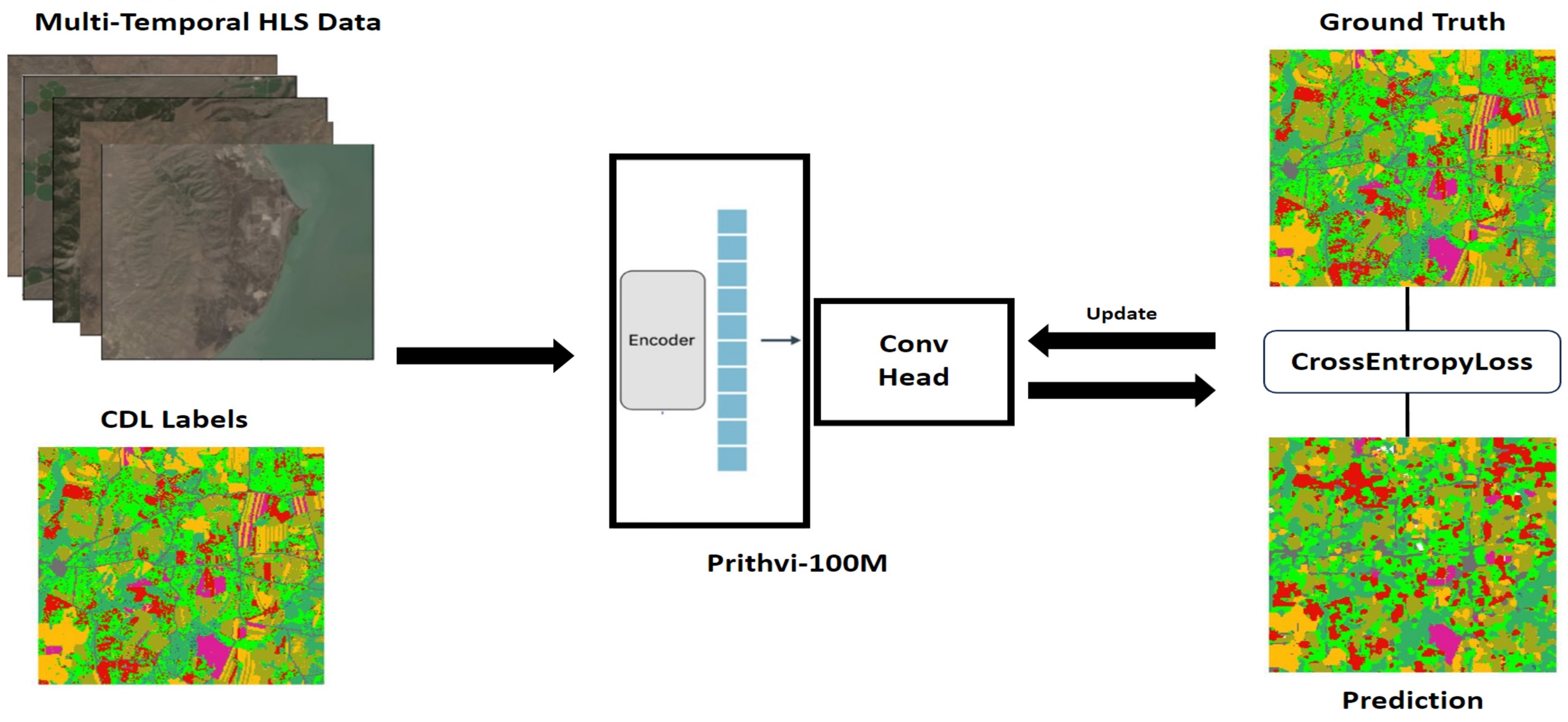
- This model differs from other fine-tuning tasks in that input data is multi-temporal (inputs contain data from three different timesteps).

Data: multi_temporal_crop_classification dataset

- 224 x 224 pixels
- Each tif file contains 18 bands: 6 spectral bands for 3 time steps stacked together
  - i.e., each GeoTIFF file represents three images acquired over the same spatial location at 3 different times
- Each GeoTIFF file for the mask contains one band, where each pixel represents the target categorical class

Fine-tuning: Crop Classification

1. **Model Overview:**
    - Pretrained Model: Prithvi-100m.
    - Fine-tuning Purpose: crop classification.
2. **In each input GeoTIFF the following bands are repeated three times for three observations throughout the growing season: Channel, Name, HLS S30 Band number:**
    - 1, Blue, B02
    - 2, Green, B03
    - 3, Red, B04
    - 4, NIR, B8A
    - 5, SW 1, B11
    - 6, SW 2, B12
3. **Label Categories:**
    - Class 0: Natural Vegetation.
    - Class 1: Forest.
    - Class 2: Corn.
    - Class 3: Soybeans.
    - Class 4: Wetlands.
    - Class 5: Developed/Barren.
    - Class 6: Open Water.
    - Class 7: Winter Wheat.
    - Class 8: Alfalfa.
    - Class 9: Fallow/Idle Cropland.
    - Class 10: Cotton.
    - Class 11: Sorghum.
    - Class 12: Other.
    - Class 13: No data/ Clouds.

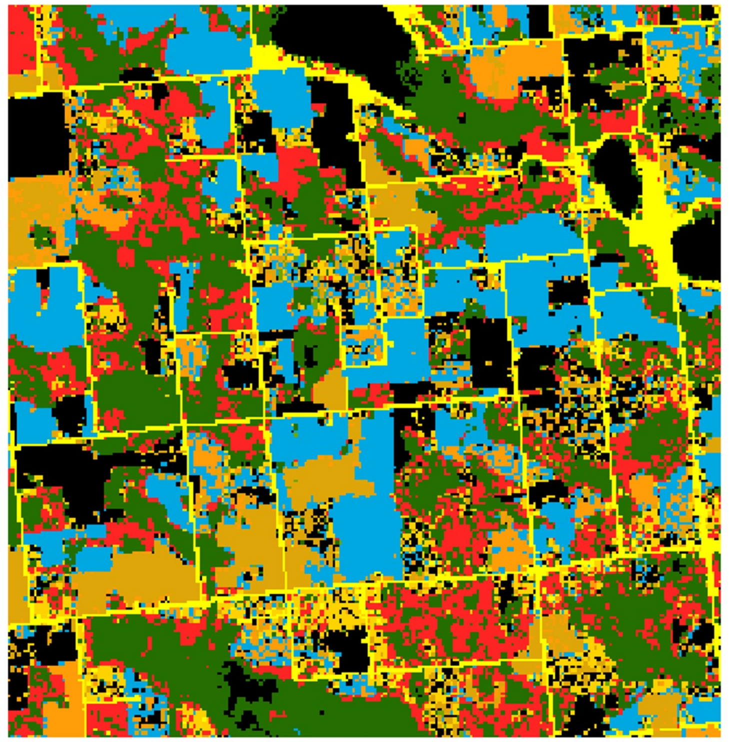


Fine-tuning: Crop Classification

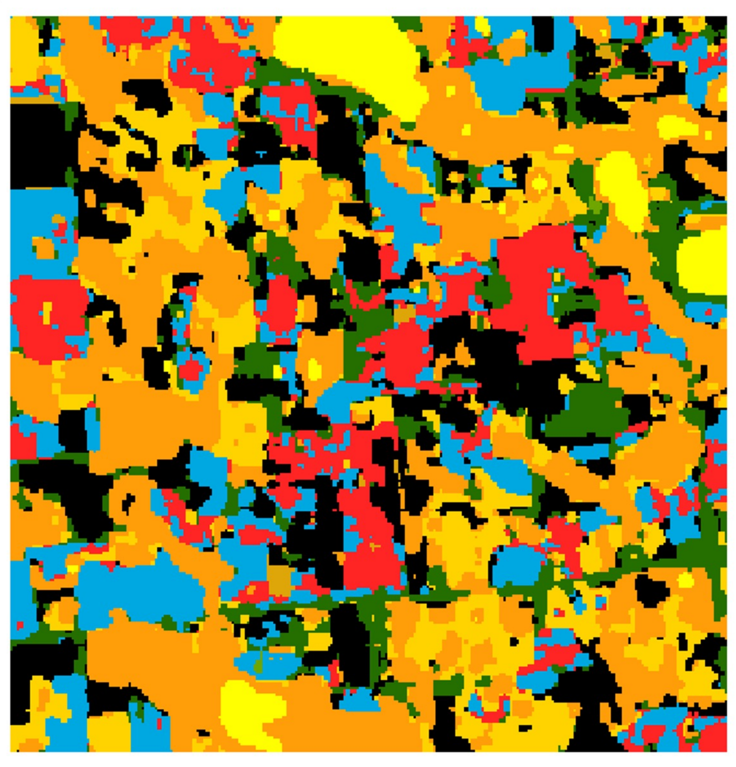
Classes	IoU	Acc
Natural Vegetation	0.4038	46.89%
Forest	0.4747	66.38%
Corn	0.5491	65.47%
Soybeans	0.5297	67.46%
Wetlands	0.402	58.91%
Developed/Barren	0.3611	56.49%
Open Water	0.6804	90.37%
Winter Wheat	0.4967	67.16%
Alfalfa	0.3084	66.75%
Fallow/Idle Cropland	0.3493	59.23%
Cotton	0.3237	66.94%
Sorghum	0.3283	73.56%
Other	0.3427	47.12%

aAcc	mIoU	mAcc
60.64%	0.4269	64.06%

Input (Multi-Temporal Imagery)



Multi-classification



Model Prediction

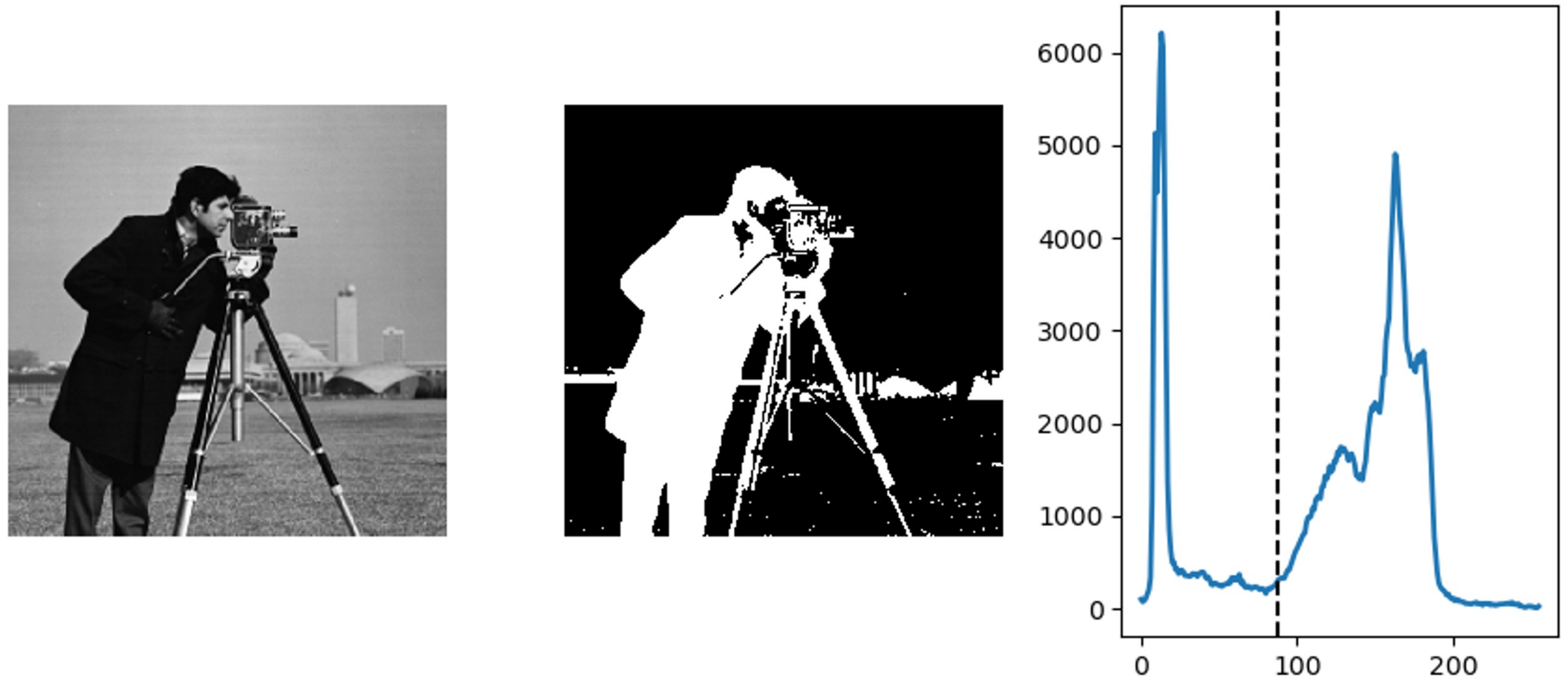
Fine-tuning Model for Our Novel Application

Normalized Difference Built-Up Index (**NDBI) for Urban Mapping** (Zha et al., 2003)

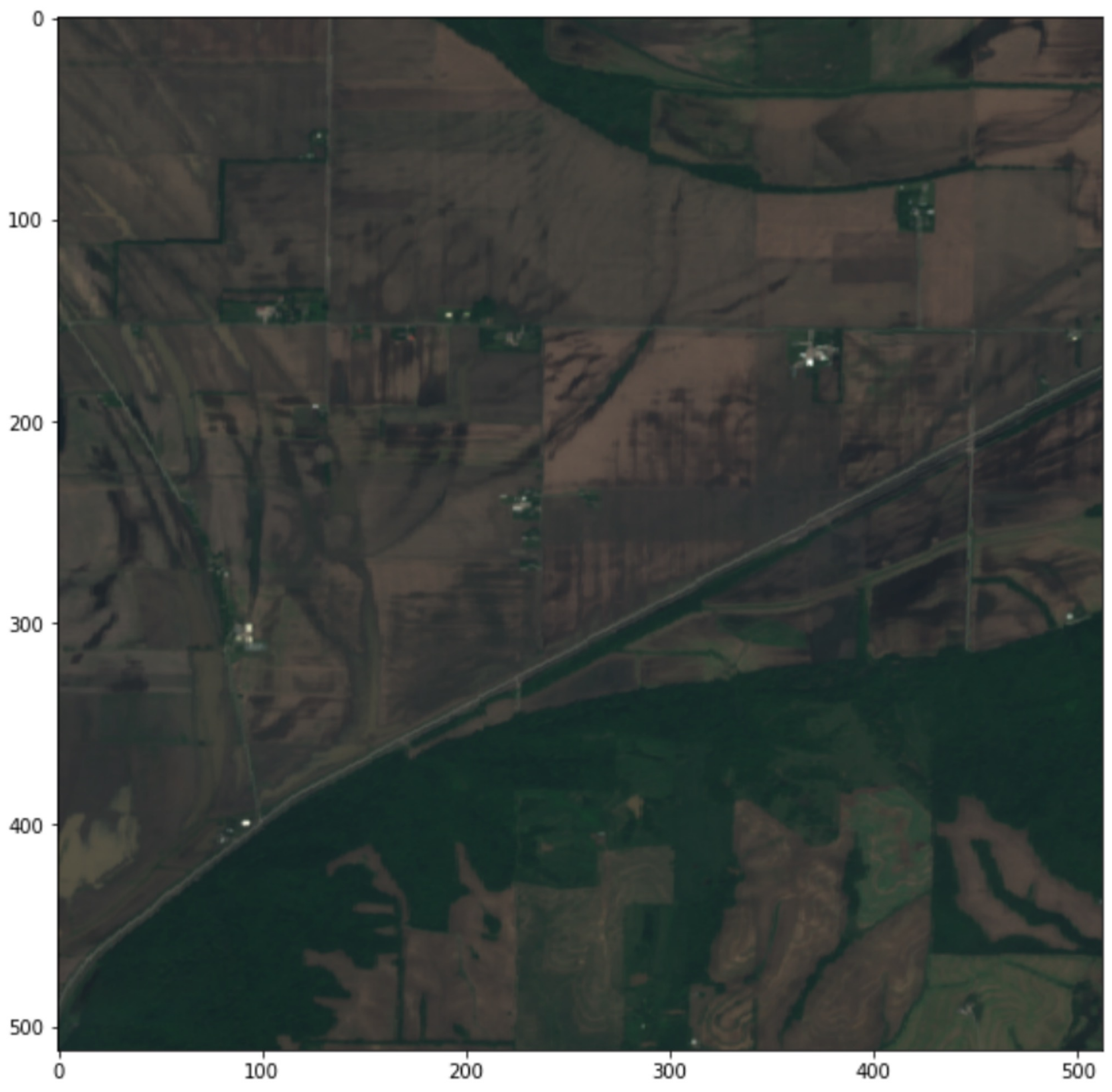
$$NDBI = \frac{SWIR - NIR}{SWIR + NIR}$$

where SWIR is the mid-infrared and NIR is the near-infrared band

Otsu's thresholding method



Fine-tuning Model for Our Novel Application



No water

Water / Flood

Non built up area

Built up area

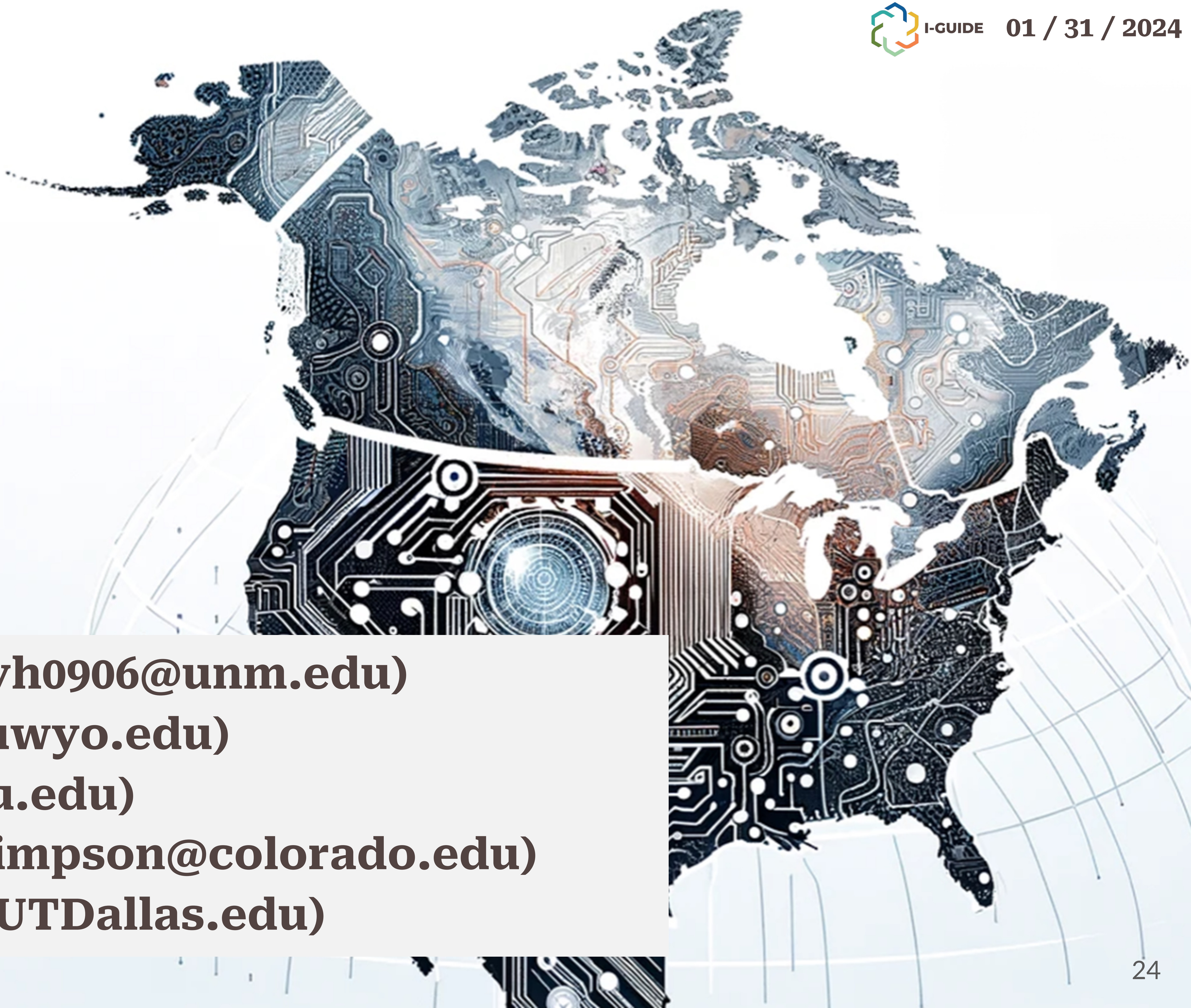
Contributions

Data Card

I-GUIDE

ABRIDGED VERSION PRODUCED FOR I-GUIDE SUMMER SCHOOL, August 2023

Team 6 dataset	Write a short summary describing your dataset (limit 200 words). Include information about the content and topic of the data, sources and motivations for the dataset, benefits, and the problems or use cases it is suitable for. The dataset utilized in this project contains temporal Harmonized Landsat-Sentinel imagery of diverse land cover and crop type classes across the Contiguous United States for the year 2022. The data card is right here https://huggingface.co/datasets/ibm-nasa-geospatial/multi-temporal-crop-classification. The primary motivation behind this dataset was to use a model capable of generating masked datasets for specific classification, leveraging training datasets for model training and validation. The resultant dataset showcases prediction results that could be vital for urban planning, environmental studies, disaster management, and other relevant fields. The process's benefits include the ability to monitor land use changes, understand urban expansion, and contribute to sustainable development practices. This dataset is particularly suitable for applications that require a detailed understanding of land use patterns, where quality spatial information about built-up areas is essential. Its utilization promises to enhance the understanding and management of urban landscapes, catering to both scientific research and practical applications in various domains.
DATASET LINK	DATA CARD AUTHOR(S)
Provide a link to the dataset:	*Select one role per Data Card Author:* *(Usage Note: Select the most appropriate choice to describe the author's role in creating the Data Card.)*
Dataset Link https://huggingface.co/datasets/ibm-nasa-geospatial/multi-temporal-crop-classification/tree/main	**Claire Simpson, Team 6:** (Contributor) **Salar Jarhan, Team 6:** (Contributor) **Yalin Yang, Team 6:** (Contributor)



Thank You
Q & A

Yanhong Huang (huangyh0906@unm.edu)
Salar Jarhan (Sjarhan@uwyo.edu)
Jiyoung Lee (jlee257@lsu.edu)
Claire Simpson (Claire.simpson@colorado.edu)
Yalin Yang(Yalin.Yang@UTDallas.edu)