NC STATE UNIVERSITY

Rapid Disaster Response

Rishabh Jain[1]
[1]North Carolina State University

CSC722_ATML_Spring_2026

The Mission

The Problem

- When a natural disaster strikes, we have satellite imagery immediately, but we lack **labels**.
- Waiting for humans to map damage takes days; emergency response needs it quickly.

The Vision

- We want to build an AI that **doesn't need to be re-trained** for every new disaster.

The Goal

- Develop a **Label Efficient** system that can **Spot the Difference** between pre-and-post-disaster images with surgical precision.

Before Disaster: Model Recognizes Normal Topographical area

After Disaster: Model Recognizes Anomalous Topographical area

[1] Before and after satellite photos show Hurricane Helene's destruction of the Florida coast, Business Insider

The Process: Phase 1 (Learn Earth)

Current Limitation

- Traditional Deep Learning methods, CNNs requires thousands of hand-labeled **Damaged vs Undamaged** examples. This fails in real-time crises.

Solution

- **Self-Supervised Masked Image Modeling (MIM)** to learn to reconstruct missing parts of satellite images.
- MIM learns the Grammar of Topography - what a roof, a road, or a tree should look like.

Intuition

- Once the AI has this "common sense" (a Foundation Model), it only needs a tiny handful of examples to become an expert at damage detection.

Dataset:

- **Prithvi** with weights (pre-trained on terabytes of Sentinel-2 data).

The Process: Phase 2 (Spot the Difference)

Model Backbone

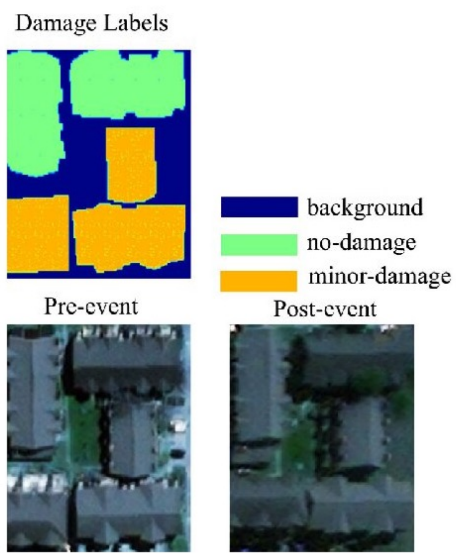
- **Vision Transformers (ViTs) -** specifically leveraging pre-trained weights from **DINOv2** or **Prithvi**.

The Strategy

- We feed in the **Pre-Event** image, so the AI knows the baseline.
- We feed in the **Post-Event** image, as the test.
- Instead of just subtracting pixels, we compare the **Latent Representations** to identify structural failures.
- Because it learned normal shapes in Phase 1, the model can instantly see if a building is **damaged** vs. just shadowed.

The Dataset: xBD Dataset (Gold standard for this task).

- 19 different disaster types.
- 0.5 meter High-resolution imagery with 850,000+ labeled buildings.



The Goals and Publication

Key Questions

- Can we achieve high accuracy (above 90%) using a small portion of labeled data (10% or less)?
- Does the model trained on floods in Americas perform well on floods in Asia?
- What is the tipping point for labels? How many disaster images do we actually need to fine-tune the Foundation model before it matches a fully supervised one? (reference to few-shot learning)
- Can a model trained on one disaster (such as California wildfires), effectively detect damage from another disaster (such as Tsunami in Japan)? (reference to Domain Shift for robust AI)
- Which works better for Earth Observation: Reconstructing pixels (MAE/RMSE) or comparing images as pairs (Contrastive Learning)?
- Where does the model/pipeline fail? What are the limitations of this project? (only be answered if a team works on it, in mid-April).

Publication target:
CVPR EarthVision 2026 (March 2, 2026) (IF POSSIBLE)
Open for Discussion

Thank you for lending me your time.
Thank you so much.

Contact for any questions:
Rishabh Jain (rjain29@ncsu.edu)
Ranga Raju Vatsavai (rrvatsav@ncsu.edu)