

Post-Disaster Building Damage Assessment Using Deep Learning on Satellite Imagery

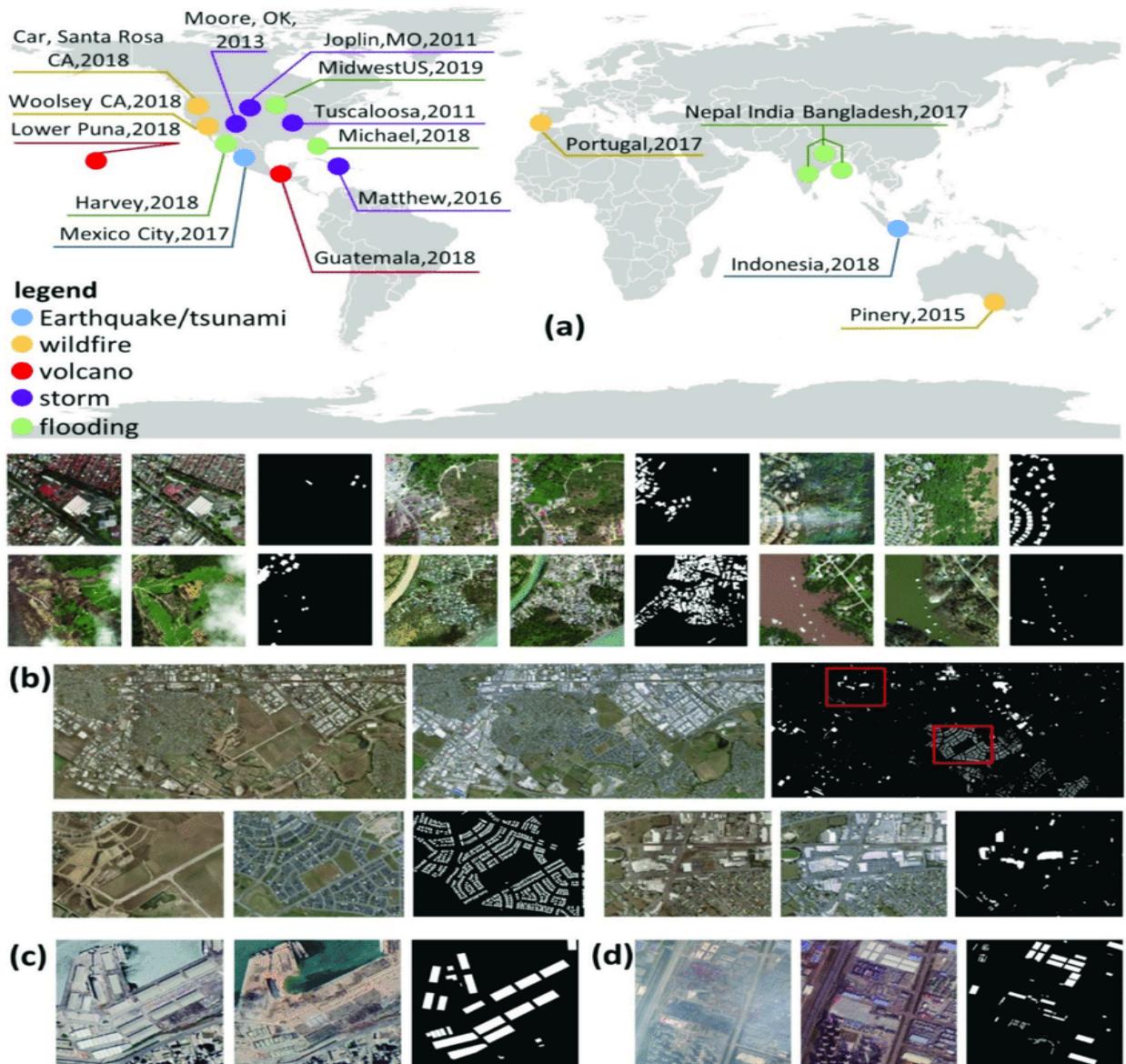
Introduction:

Disasters, both natural and man-made, pose critical challenges to emergency response systems and resource allocation strategies. As the frequency and severity of natural disasters continue to rise, developing resilient systems capable of anticipating, absorbing, and recovering from these disruptions has become a national priority in the United States. Resilience analytics—an emerging field that combines predictive modeling with geospatial analysis—provides a powerful framework for enhancing disaster preparedness and response. Building on advances in big data and artificial intelligence (AI), recent research has shown that data-driven approaches can significantly improve situational awareness and decision-making in disaster risk management (Sarker et al., 2020; Ghaffarian & Taghikhah, 2023). This project investigates the application of deep learning and satellite imagery for rapid and accurate assessment of building damage following disasters, utilizing the xView2 dataset which includes events like windstorms, floods, and tsunamis. Initially, our research questions focused on integrating heterogeneous data sources—including social media, IoT sensor data, and emergency service logs—to support real-time response and public sentiment analysis. However, due to the time-intensive nature of multi-source integration and the complexity of image classification, we refined our scope to focus exclusively on geospatial data. Our research concentrates on the direct analysis of pre- and post-disaster satellite imagery. To this end, we developed and evaluated a system based on a Siamese UNet architecture. This model is specifically designed to analyze paired pre- and post-disaster imagery to identify and classify structural damage at the building level through semantic segmentation. This shift allowed us to develop a more focused and technically rigorous approach to structural damage assessment while still pursuing the broader goal of enhancing disaster preparedness through real-time predictive analytics. Our research questions center on the efficacy of this Siamese UNet model for precise damage prediction, the model's ability to leverage temporal changes captured in satellite imagery, and the broader implications of employing such geospatial-AI integration for enhancing emergency response planning and resource allocation. By focusing on a deep learning approach tailored for multi-temporal satellite image analysis, this project aims to contribute to more efficient and informed disaster response efforts.

Description of the dataset:

The dataset employed in this study is the xView2 dataset, a comprehensive collection of high-resolution satellite imagery designed for building damage assessment in post-disaster scenarios. Developed by the Defense Innovation Unit and Carnegie Mellon University Software Engineering Institute, xView2 comprises over 44,000 images from 17 natural disaster events across six disaster types, including earthquakes, tsunamis, floods, volcanic eruptions, wildfires, and windstorms . Each image is paired with a corresponding post-disaster image, providing a total of 9,168 pre-disaster and post-disaster image pairs, each with a resolution of 1024×1024 pixels. The dataset's primary focus is on building damage assessment, with each building annotated according to a four-level damage scale: no damage (0), minor damage (1), major damage (2), and destroyed (3). This classification enables fine-grained analysis of structural impacts resulting from various disaster events. Notably, the dataset exhibits a class imbalance, with the majority of buildings labeled as "no damage" and fewer instances of "major damage" and "destroyed" categories.

For visualization, the following figure illustrates a sample of the xView2 dataset, highlighting the spatial distribution of building damage across different disaster events:



In this study, we utilized image patches derived from the xView2 dataset to train and evaluate a Siamese UNet based model, aiming to predict post-disaster building damage with pixel-level accuracy.

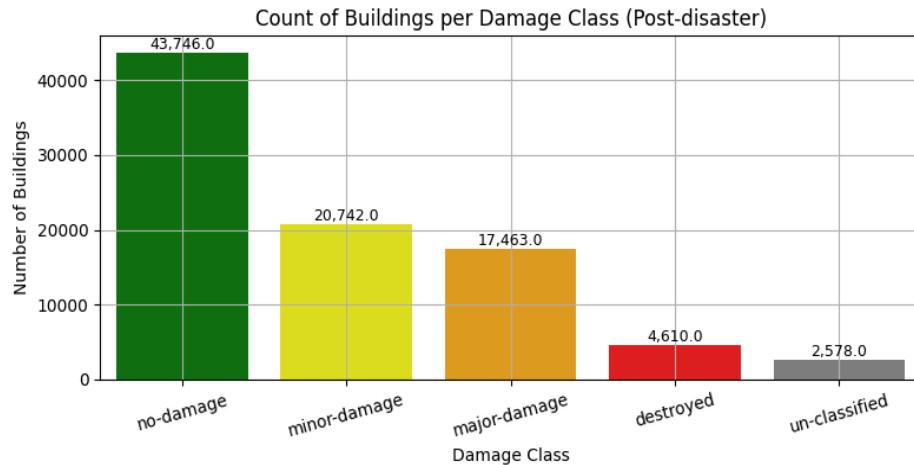
Data Cleaning and Preparation:

To prepare the xView2 dataset for model training, we performed a series of data cleaning and preprocessing steps to ensure data consistency, relevance, and model-readiness. First, we filtered the dataset to include only disaster events related to windstorms, as this was enough information to train and get an accurate performing model but did not overwhelm it with a lot of data. We then reviewed the dataset for missing values, particularly in the damage annotation labels and geospatial metadata, and excluded any image pairs with incomplete or corrupted data. Following cleaning, the satellite imagery was divided into smaller, uniformly sized patches to facilitate

efficient training with our Siamese UNet-based model. Each patch preserved the semantic information of its corresponding pre- and post-disaster image pair, with aligned damage annotations carried over at the pixel level. Additionally, the dataset was normalized and formatted into a consistent input structure, ensuring compatibility with our model pipeline and reducing the risk of training artifacts. These preparation steps were critical in enhancing data quality, balancing the classes across disaster types, and ensuring reliable model performance in subsequent analysis.

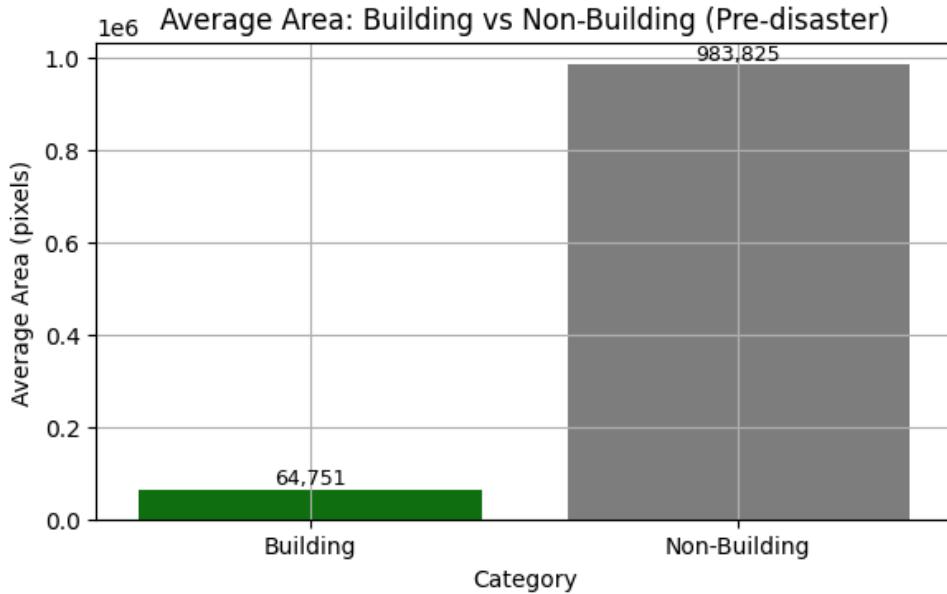
Exploratory Data Analysis

To better understand the characteristics and patterns within the xView2 dataset, we conducted an exploratory data analysis (EDA) focused on buildings affected by windstorms. This involved visualizing building damage levels, analyzing geometric properties, and mapping spatial distributions. First, the distribution of damage grades across all annotated buildings was analyzed. The bar chart revealed a heavy skew toward the "no damage" category (43,746 buildings), with fewer instances labeled as "minor damage" (20,742 buildings), "major damage" (17,463 buildings), and "destroyed" (4,610 buildings), and a subset marked as "unknown" (2,578 buildings) as a result of annotation gaps or unclear imagery. This class imbalance is a notable characteristic of the dataset and has direct implications for model training, requiring appropriate strategies, such as class weighting or oversampling, to avoid bias in predictions.



Next, we examined the physical characteristics of buildings, specifically focusing on their area. After ensuring all geometries were valid and transforming coordinate reference systems to calculate real-world areas in square meters, we visualized the distribution of building sizes. The initial histogram of raw areas showed a skewed distribution, many small values and a few very large ones because real-world building areas vary greatly making most data crammed into the first few bins. To address this skew, we applied a log transformation, which revealed a more normalized distribution, helpful for model interpretability visualizing the mid-sized and small buildings that were hidden before and feature scaling. Finally, we analyzed the average area of building versus non-building on the pre-disaster dataset by pixels. This revealed that across the windstorms data, on average 64,751 pixels displayed the image and the remaining 983,825 pixels of the image were non-building. This is an important finding that was implemented into the

training of the model to extract meaningful insights for training. Through this EDA process, we uncovered several key insights: (1) the dataset exhibits class imbalance with a dominance of undamaged buildings, (2) building sizes vary significantly, with most being small to moderately sized, and (3) roughly 6% of the image's pixels across the dataset represented the building that can inform downstream modeling. These findings guided decisions in preprocessing, model design, and evaluation strategies.

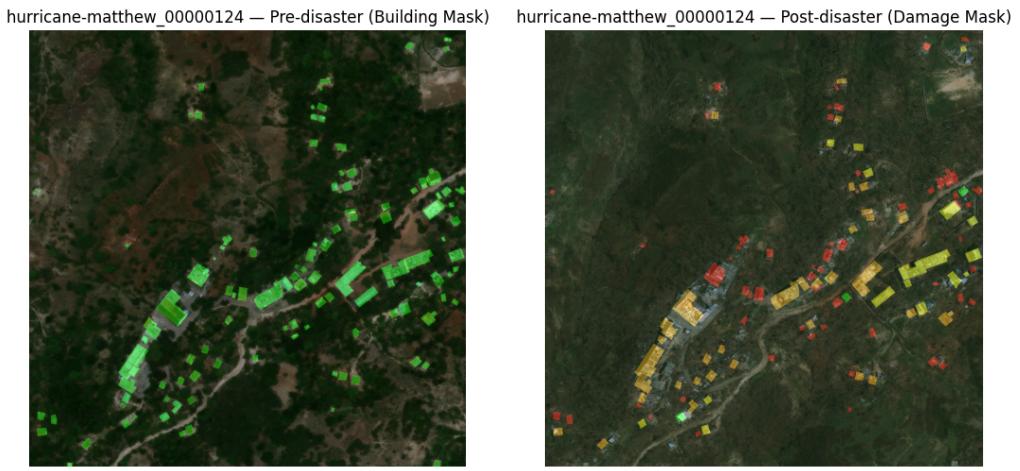


Modeling and Results:

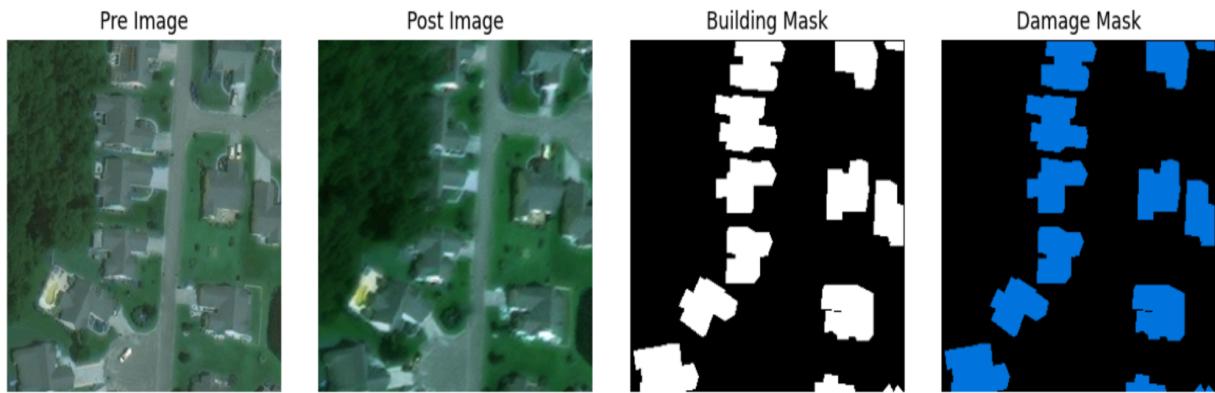
This project applies a deep learning approach to analyze satellite imagery captured before and after natural disasters, aiming to automatically detect buildings and classify the level of damage sustained. The system comprises several stages: preparing the input data, generating pixel-wise labels, dividing the data into manageable units, and designing a neural network capable of performing both segmentation and classification.

We process satellite imagery from .tif format by extracting the RGB channels to create standard visual images. To handle variations in lighting conditions across different satellite captures, we apply normalization techniques that adjust pixel values to a consistent range. This normalization step is crucial for the neural network to learn effectively without being influenced by brightness inconsistencies.

The project uses structured JSON files containing building outlines and damage classifications to create training labels. We convert these vector-based annotations into pixel-level masks where each building pixel receives a numerical code representing its damage category (ranging from undamaged to completely destroyed). This transformation creates ground truth data that our model can learn from.



Since satellite images are large and complex, we divide each scene into smaller square sections or "tiles" measuring 256 by 256 pixels. This step simplifies the learning process and allows the model to focus on localized features. Only patches that contain labeled buildings are kept to ensure training data remains relevant. Each patch contains four components such as the pre-disaster image, the post-disaster image, the building mask and the damage mask. These are saved in compressed files to speed up loading during training.



The deep learning model used is a variant of the U-Net architecture, known as **Siamese U-Net**. It has two identical input paths, one for the pre-disaster image and one for the post-disaster image. Each path extracts features independently but using shared weights. These features are then combined to highlight changes in the image pair.

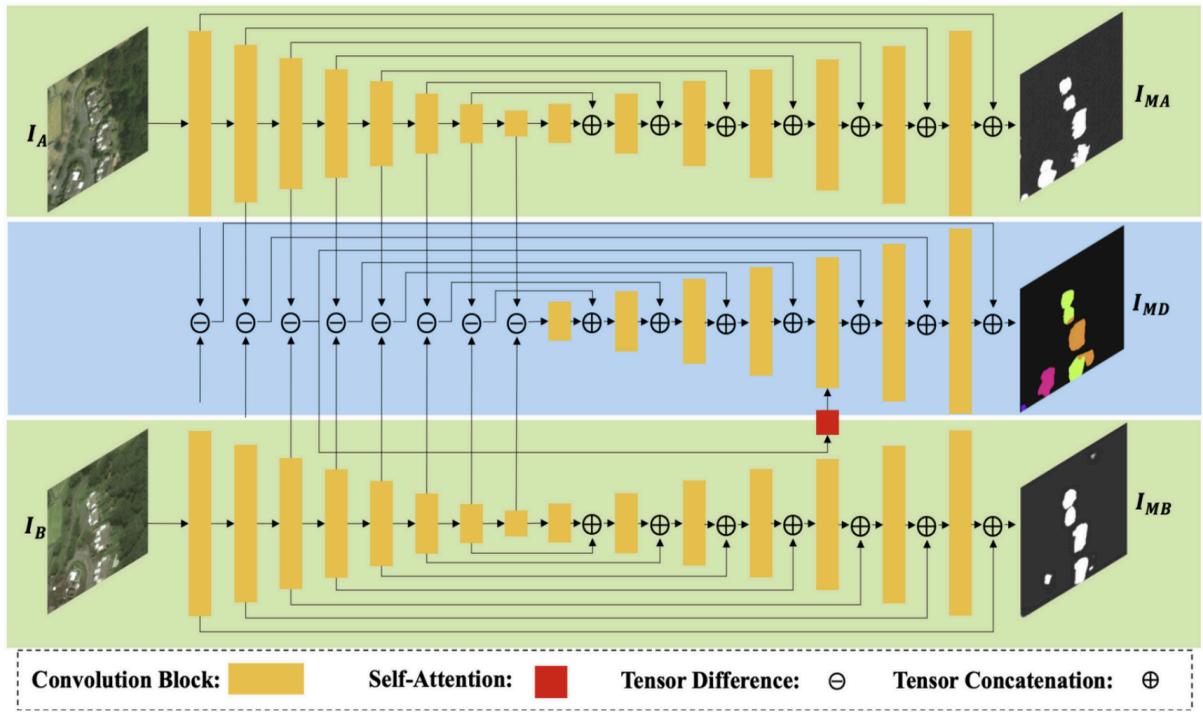


Figure: Siam-U-Net model where I_A and I_B are the pre-disaster and postdisaster input images. I_{MA} and I_{MB} are the corresponding output building segmentation masks. I_{MD} is the output damage scale classification mask.

By combining the strengths of encoder-decoder networks with the ability to compare two inputs, this architecture is well-suited for change detection and damage assessment tasks. The model outputs **binary segmentation mask** from pre and post disaster images along with **multi-class damage map** that assigns each pixel a damage category.

The model is trained with two objectives in mind: identifying where buildings are located, and determining the extent of damage they have sustained. Two loss functions are used, one for segmentation and another for classification, and the total training loss is a weighted sum of both. This combined objective helps the model learn spatial structure and semantic meaning simultaneously.

Our training configuration includes the following carefully selected parameters:

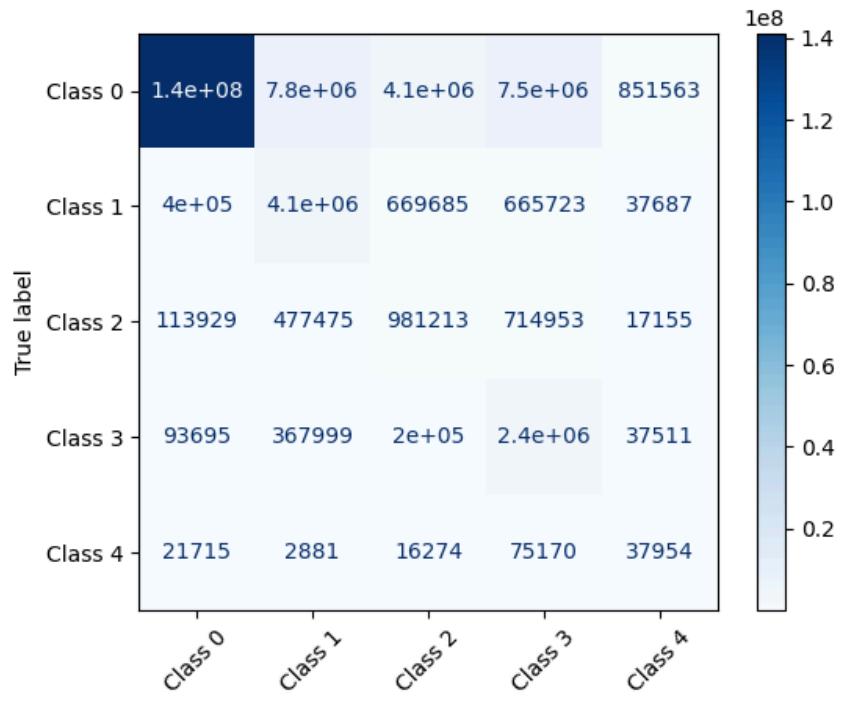
- Learning rate: 0.001
- Loss function: Cross Entropy Loss
- Optimizer: Adam
- Batch size: 32
- Epochs: 30



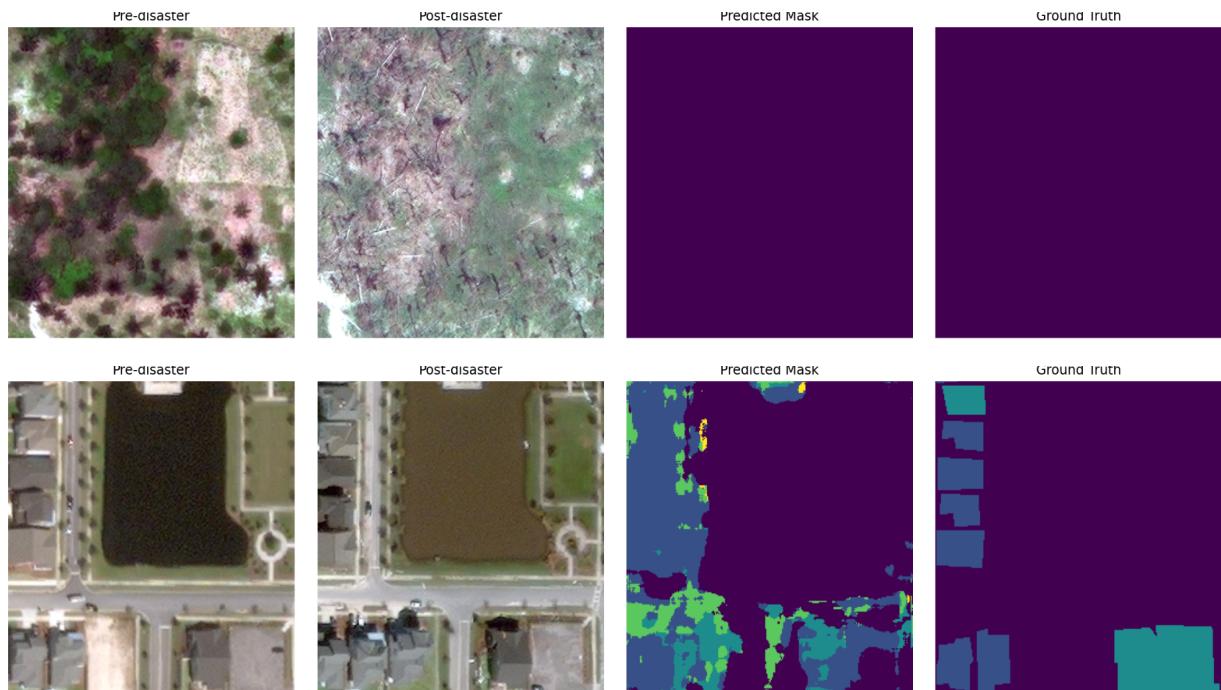
Figure: Loss Convergence

The model undergoes training with these parameters while monitoring validation performance. We implement early stopping criteria based on validation metrics to select the optimal model version that generalizes best to unseen disaster events.

Evaluation on the xView2 dataset showed an average pixel-level accuracy of 85.93%. Class-wise analysis revealed strong precision and recall for the “destroyed” category, while performance was lowest for “minor damage,” likely due to its more subtle visual patterns. Visual inspections confirmed that the model produced accurate segmentation masks for clearly defined buildings, particularly in urban regions with consistent lighting. However, model performance was less stable in cluttered or low-contrast scenes.



Overall, the results demonstrate the effectiveness of Siamese UNet architectures in identifying significant structural damage from satellite imagery, with robust performance for higher-severity damage categories. These findings establish a solid foundation for building real-time, scalable disaster response tools based on geospatial AI.



Discussion and future work:

Our findings demonstrate the potential of Siamese UNet architectures for rapidly assessing post-disaster building damage using satellite imagery. This approach effectively captures temporal changes between pre- and post-disaster scenes, offering a scalable alternative to manual assessments. However, our analysis is not without limitations. One key constraint was the semantic segmentation resolution, which lacked the precision needed to distinguish subtle damage patterns in dense urban areas. Enhancing pixel-level classification could address this by providing more fine-grained differentiation between damaged and undamaged structures. Additionally, background elements such as roads, vegetation, and water bodies sometimes confused the model, highlighting the need for improved filtering techniques. Future research could explore the integration of pre-segmentation methods or building footprint masks to ensure the model focuses on relevant areas. Another limitation involved inconsistencies in satellite image quality due to varying atmospheric conditions and sensor settings. We learned that preprocessing plays a crucial role in model reliability; with additional time, we would experiment with techniques such as brightness normalization, histogram equalization, or CLAHE to improve consistency across images. Finally, although we focused on windstorm-related damage in this project, the broader applicability of our method to other disasters is a strong direction for future research. Through this project, we gained valuable experience in balancing technical ambition with practical constraints. One important lesson was the need to narrow scope early when integrating multiple data sources becomes infeasible. Given more time, we would revisit our initial goal of multi-modal data fusion (combining satellite imagery with social media or sensor data) to enrich situational awareness and build more holistic disaster response tools.

Python

References:

- Arena, F. (2023, July). *Artificial intelligence in emergency response systems* [PDF]. ResearchGate. https://www.researchgate.net/publication/387829449_Artificial_Intelligence_in_Emergency_Response_Systems
- Ghaffarian, S., & Taghikhah, F. (2023). *Explainable artificial intelligence in disaster risk management: Achievements and prospective futures* [Preprint]. SSRN. <https://doi.org/10.2139/ssrn.4421109>
- Gupta, R., Hosfelt, R., Sajeev, S., Patel, N., Goodman, B., Doshi, J., ... & Heim, E. (2019). *xView2: A dataset for building damage assessment using satellite imagery* [Preprint]. arXiv. <https://arxiv.org/abs/1911.09296>
- Hao, H., Baireddy, S., Bartusiak, E. R., Konz, L., LaTourette, K., Gribbons, M., Chan, M., Delp, E. J., & Comer, M. L. (2021). An attention-based system for damage assessment using satellite imagery. 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, 4396–4399. <https://doi.org/10.1109/igarss47720.2021.9554054>

IBM. (n.d.). *The xView2 AI challenge: Building resilience from above*. IBM Insights. <https://www.ibm.com/think/insights/the-xview2-ai-challenge>

Numan, A. (2024, May). *Artificial intelligence in disaster management: Effectiveness and limitations* (Master's thesis). St. Cloud State University. https://repository.stcloudstate.edu/cgi/viewcontent.cgi?article=1032&context=cjs_etds

OpenStreetMap contributors. (n.d.). *OpenStreetMap*. <https://www.openstreetmap.org>

Sarker, M. N., Peng, Y., Yiran, C., & Shouse, R. C. (2020). Disaster resilience through big data: Way to environmental sustainability. *International Journal of Disaster Risk Reduction*, 51, 101769. <https://doi.org/10.1016/j.ijdrr.2020.101769>

Uddin, M. R., Islam, M. R., & Islam, M. S. (n.d.). *Potential use of artificial intelligence (AI) in disaster risk and crisis management*. National Center for Biotechnology Information (NCBI). <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10712270/>