Enhancing Object Geolocations in Imagery to Improve Disaster Damage Mapping and Assessment

Jooho Kim[1], Sayok Bose[2], Sarah Brasseaux[2] and Junho Kim[3]

[1] *Institute for a Disaster Resilient Texas at Texas A&M University*
[2] *Computer Science at Texas A&M University*
[3] *Computer Science at Rice University*

Introduction

Community-driven or field survey imagery is vital for conveying disaster impacts by providing data on damage levels, locations, and extent. **This study develops an algorithm to enhance geolocation calibration in such imagery, improving the accuracy of disaster impact assessments and related geodatabases.** We leverage models like Contrastive Language-Image Pre-Training (CLIP) and Vision Transformers (ViT) to address current geocoding limitations and produce more precise damage maps across a range of disaster events.

For this case study, we analyzed data from the 2023 tornado in Norman, Oklahoma, using the **NOAA database**[1] to evaluate our framework. We compared the NOAA dataset with Google Street View images to align coordinates for each image. Our framework leveraged **open-source models like CLIP and ViT to compute similarity metrics**. These models were applied to both original images with background context and cropped images focusing solely on buildings. Attention maps provided further insights into the focal points of the models, enhancing our analysis.

Data Description

Data sources: NOAA Damage Assessment Toolkit (2023 Norman, OK tornado) and Google Street View (GSV).

Norman, OK Tornado Building Types Count:

Apartments, Condos, Townhouses [3 stories or less] (ACT)	Low-Rise Building [1-4 Stories] (LRB)	Manufactured Home - Single Wide (MHSW)	Metal Building System (MBS)	One- or Two-Family Residences (FR12)	Small Retail Building [Fast Food Restaurants] (SRB)	Small Professional Building [Doctor's Office, Branch Banks] (SPB)
7	1	2	4	182	2	2

Sample images from the NOAA dataset


35.1937, -97.3987

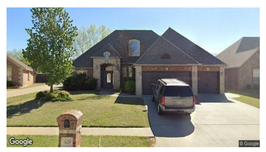

35.1733, -97.4333


35.1936, -97.3988

Sample images from the GSV dataset

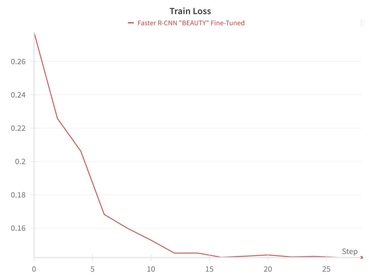

35.32178, -97.2123

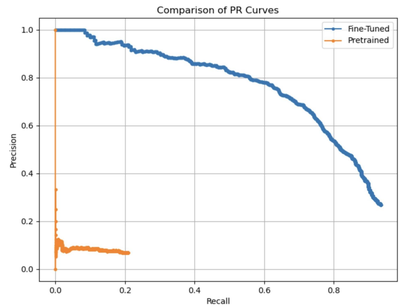

35.24687, -97.32171


35.212072, -97.392560

Building Detection and Image Cropping

- **Fine-tuned Faster R-CNN with Building Dataset**[2]





- **Bounding Boxes for Damaged Buildings in NOAA Dataset**

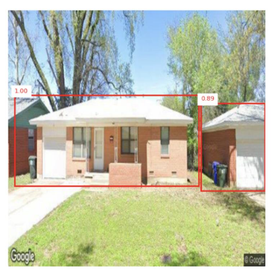






- **Bounding Boxes for Buildings in GSV Dataset**

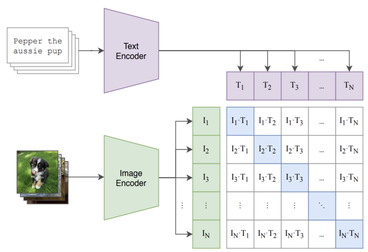




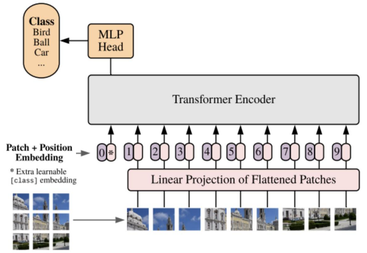


Model as Image Similarity Metric

CLIP

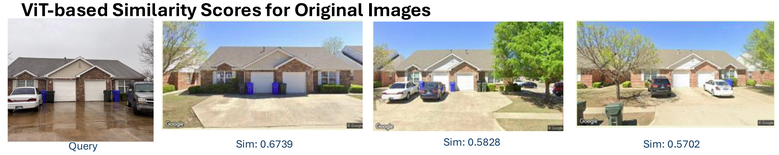


ViT



Results

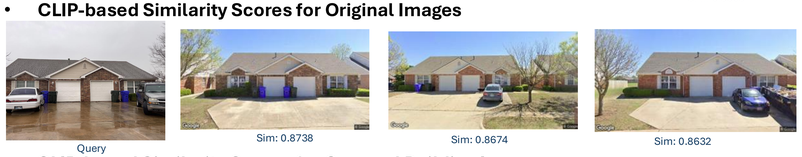
- **ViT-based Similarity Scores for Original Images**



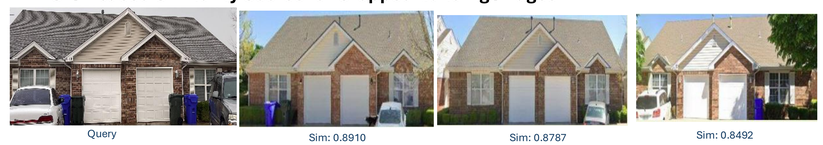
- **ViT-based Similarity Scores for Cropping Building Images**



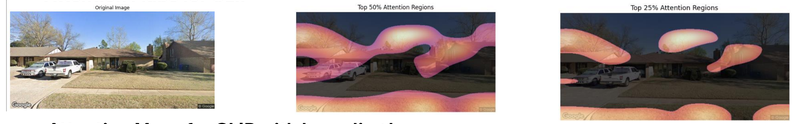
- **CLIP-based Similarity Scores for Original Images**



- **CLIP-based Similarity Scores for Cropped Building Images**



- **Attention Maps for CLIP**



- **Attention Maps for CLIP with bounding boxes**



Limitations and Future Work

Ground Truth Dataset: We currently lack a ground truth dataset to validate image correspondences as well as for **deep model analysis**.

Generalizability: Our work is restricted by geographic scope, building damage levels, and background artifacts. Future study can extend and evaluate our framework across diverse regions and more severely damaged structures.

References

[1] National Oceanic and Atmospheric Administration. NOAA Damage Assessment Toolkit. https://apps.dat.noaa.gov/stormdamage/damageviewer/

[2] Zhao, Kun, et al. "Bounding boxes are all we need: street view image classification via context encoding of detected buildings." IEEE Transactions on Geoscience and Remote Sensing 60 (2021): 1-17.