Evaluating the Efficacy of Vision-Language Models for Automated Wildfire Damage Classification

Ella Wilson[1], Ryan Chen[2], Barbaros Cetiner, Ph.D.[3] Joseph Wartman, Ph.D.[4] Laura Lowes, Ph.D. [4] Jeffrey Berman, Ph.D.[4]

[1]McGill University, [2] The Cooper Union School for the Advancement of Science and Art, [3]University of California, Berkeley, [4]University of Washington

Purpose

This research project automates damage classification by evaluating the efficacy of artificial intelligence models in their ability to detect and assess different levels of infrastructure damage based off imagery.

Background

- Wildland Urban Interface expansion [1]
- Risk in Manual Assessment [2]
- Limitations of past models:
  - Binary [3]
  - Pre-hazard imagery [4]

Objectives

- Multi-class damage index
- Damage map of burned area

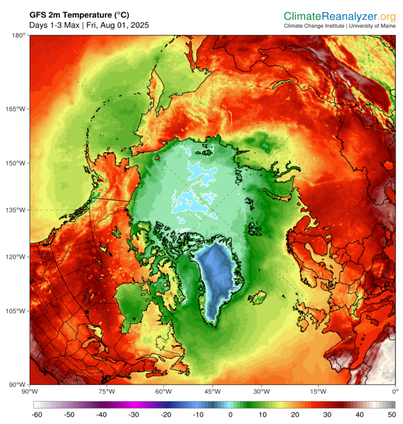

Fig. 1. Max Potential Temperatures in the Arctic by Climate Reanalyzer

Instrumentation/Software

Fig. 2. Mosaic X Camera

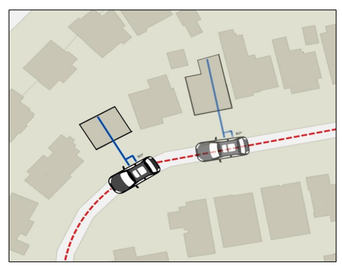

Fig. 3. Diagram of Rig Position

Table. 1. Software and Version used in this Project

Software Name	Version
MapBox GL JS Outdoor	2.0
Photo GPS Extractor	9.0
BRAILS	3.1.3
GEMINI	2.5 Pro
ArcGIS Pro	3.5

Datasets

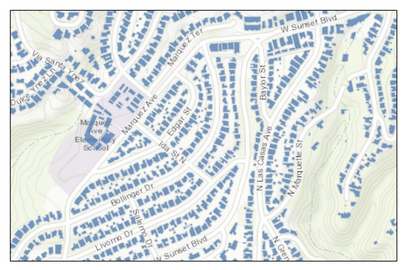

Fig. 6. Los Angeles 3D Building Footprints by Los Angeles GeoHub

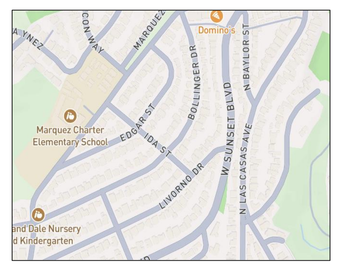

Fig. 7. Unique identifiers from MapBox GL JS OpenStreet Maps

Fig. 4. Raw panorama from PRJ-5815 Palisades Fire Burn Area Dataset

Fig. 5. Sample Final Image for Testing

Combustion (CHS) Index

Street-view definitions:

CHS Level 0: Unaffected
- Exterior walls/windows/doors are intact
- No signs of soot/charring

CHS Level 1: Radiant Heat Damage
- Scorched paint/soot stains on walls
- Melted vinyl siding, fascia, or gutters

CHS Level 2: Partial Structural Combustion
- Heavy charring of load-bearing elements
- Visible breach in an exterior wall

CHS Level 3: Major Structural Combustion
- Only skeletal frame remains standing
- Collapsed roof

CHS Level 4: Complete Combustion
- No walls are left standing
- Structure is a pile of ash and rubble
- Collapsed structure in basement frame
- Non-combustible items are visible

CHS Level 'No Data': Damage is unconfirmed
- Shrubbery covering large areas
- Construction/demolition
- Presence of Construction Vehicles on site

CHS Level 0

CHS Level 1

CHS Level 2

CHS Level 3

CHS Level 4

CHS Level No Data

Fig. 8. Example of CHS levels on residential buildings (RAPID)

Methods

- ***Prompt engineering*** – includes damage assessment instructions, wildfire context, and CHS index (entire prompt extends two pages)
- ***Prompt iteration*** – repeated prompt revisions with continual testing of such prompts with Gemini Pro 2.5 for improved results

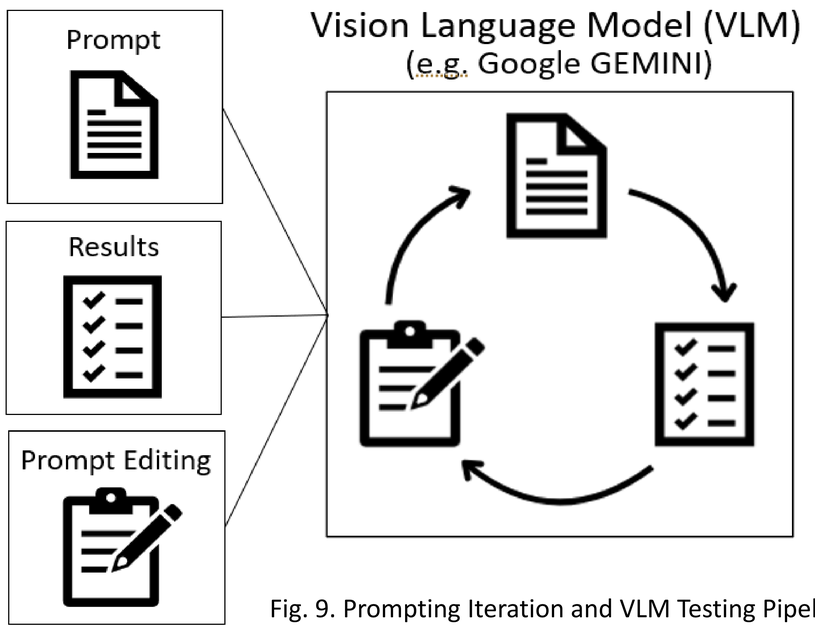

Fig. 9. Prompting Iteration and VLM Testing Pipeline

Results

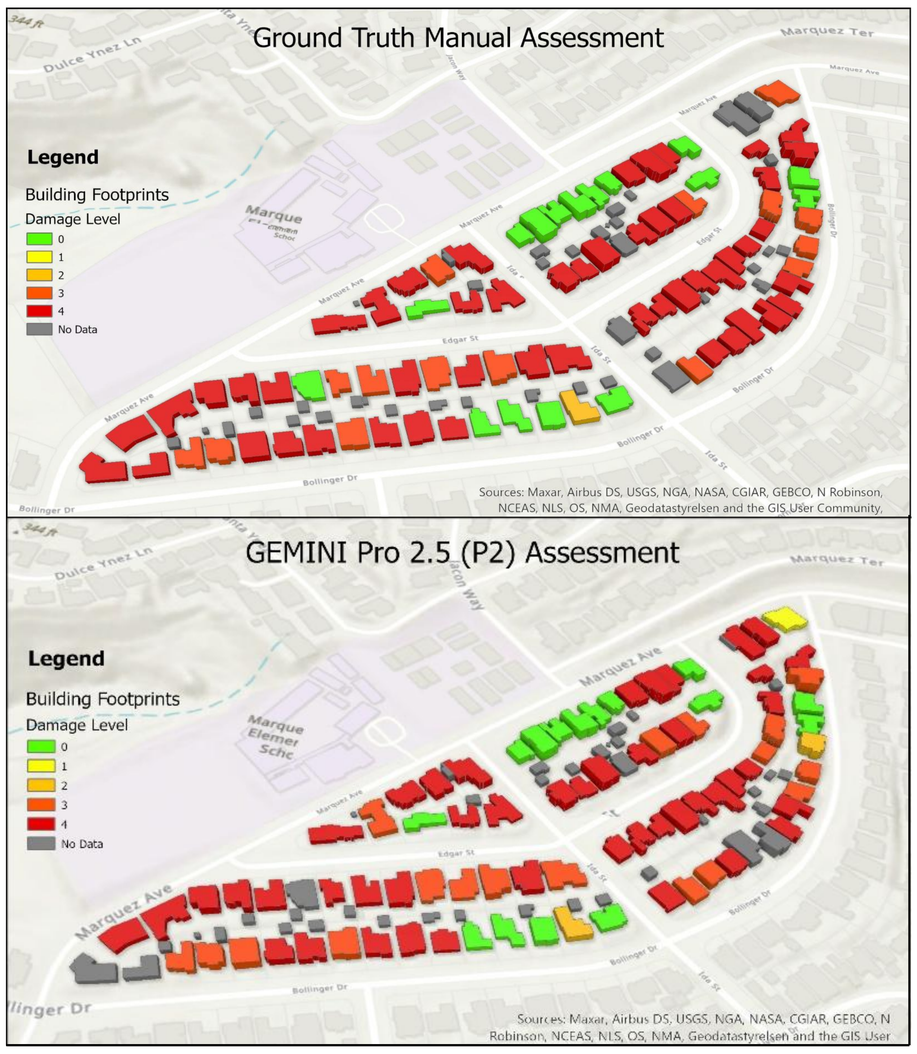

Fig. 10. ArcGIS 3D Damage Map Comparing Ground Truth and VLM Assessments

Discussion

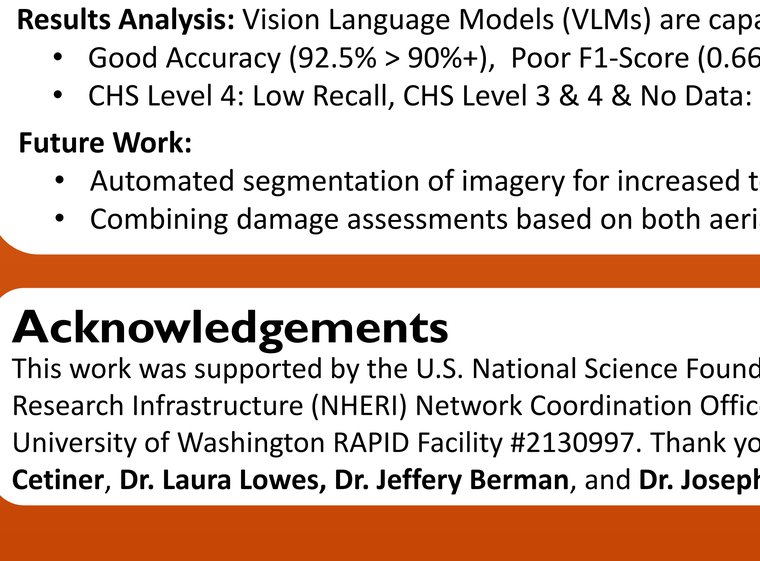

Fig. 11. Confusion Matrix Analyzing Assessment Comparisons

Table 2. Performance Analysis Metrics between Initial and Final Prompts including all CHS Levels

	Accuracy	F1-Score
Initial Prompt	91.2%	0.738
Final Prompt	92.5%	0.664
Difference	+1.3%	-0.074

Table 3. Performance Analysis Metrics between Initial and Final Prompts including CHS Levels 0-4

	Accuracy (0-4)	F1-Score (0-4)
Initial Prompt	92.1%	0.752
Final Prompt	93.8%	0.800
Difference	+1.7%	+0.048

Results Analysis: Vision Language Models (VLMs) are capable of damage assessment, but need improvement
- Good Accuracy (92.5% > 90%+), Poor F1-Score (0.664 < 0.8+) → issues primarily regarding 'No Data'
- CHS Level 4: Low Recall, CHS Level 3 & 4 & No Data: Low Precision

Future Work:
- Automated segmentation of imagery for increased testing efficiency and widespread application
- Combining damage assessments based on both aerial and street-level imagery

Acknowledgements

This work was supported by the U.S. National Science Foundation Natural Hazards Engineering Research Infrastructure (NHERI) Network Coordination Office Award #2129782 and The University of Washington RAPID Facility #2130997. Thank you to my mentors **Dr. Barbaros Cetiner**, **Dr. Laura Lowes, Dr. Jeffery Berman**, and **Dr. Joseph Wartman.**