Gen-AI enabled Disaster Scenes Analysis the Aftermath of 2025 Mandalay, Myanmar Earthquake

ZhiQiang Chen

TC33: Advancing Artificial Intelligence Education in Civil Engineering (ASCE). Email: chenzhiq@umkc.edu

April 2, 2025

1 Motivation

Disaster situational awareness is crucial for initiating rapid and sustainable community recovery, a key pillar of long-term resilience. The faster and more accurately we assess post-disaster impact, the more effectively resources can be allocated.

In this context, leveraging real-time or near-real-time data from all available sources becomes highly valuable. Among these, public sources such as online news media and citizen-generated content (e.g., social media) serve as rich, albeit noisy, indicators. If extracted and processed properly, their coverage can reflect the severity of damage and progress in rescue and recovery, often with embedded geospatial locations. Collectively, more objective disaster response monitoring and planning can be expected.

However, these data sources present *deep uncertainties*. Unlike physical sensor data, they lack mechanistic foundations and are influenced by journalism norms, public bias, and inconsistencies.

Generative AI (GenAI) presents a promising opportunity to mine and interpret such data. Through multi-modal large language models (LLMs and vision models), we can perform high-level reasoning over textual and visual content, extracting structured, relevant information for disaster intelligence.

2 Methods

In this rapid practice, the methods adopted consist of three main components:

2.1 News and Image Data Extraction

We use keyword-based search ("Mandalay", "Myanmar", "earthquake") to retrieve news articles using the NewsAPI. From these articles, we extract:

- The article title, URL, and publish date
- The full news text
- Associated images, if available

All data are stored in a structured format, and geolocations (if mentioned) are extracted using rule-based named-entity parsing.

2.2 AI-Based Multi-Modal Analysis

For each article-entry pair, we apply the following generative AI models:

1. **Text summarization using GPT-3o**: We prompt the model to extract concise, disaster-focused summaries, with emphasis on building damage, human impact, and rescue operations.
2. **Image captioning and tagging using GPT-4o**: Images are fed through a vision-enabled model that generates natural captions and high-level tags (e.g., `Damaged Building`, `Rescue`, `Debris`, `People`).
3. **Sentiment analysis**: The summary is passed again to GPT to infer a sentiment level (e.g., "Concerned", "Urgent", "Neutral").
4. **Discrepancy scoring**: To evaluate the alignment between visual and textual content, we computed a discrepancy score D for each image-summary pair using cosine distance between their semantic embeddings:

   $$D = 1 - \cos(\theta) = 1 - \frac{\langle v_{\text{caption}}, v_{\text{summary}} \rangle}{\|v_{\text{caption}}\| \cdot \|v_{\text{summary}}\|}$$

   where v_{caption} and v_{summary} are embedding vectors produced from the AI-generated caption and summary, respectively.
   A lower D value indicates a closer semantic match between the visual content and the news narrative, while a higher D suggests a divergence between what is shown in the image and what is emphasized in the article. This measure enables us to identify potential mismatches or over-/under-emphasis in disaster reporting.
5. **Prompting as Inference**: Each prompt is constructed as a structured inference request. For example, for image captioning:

   > Describe this image in a caption-style sentence focused on earthquake effects... Use a short confidence label... Do not include tags.

2.3 LLM-enabled Loss and Resilience Statistics

Using a prompt-engineered GPT-4o model, we classify each scene into a loss and resilience level. The prompt combines both the image caption and text summary and returns two integers:

- **Loss Level (1 to 3)**:
  - 3 = Mentions of injured/dead people or fully collapsed buildings with visible debris
  - 2 = Partially collapsed buildings and other damage
  - 1 = No visible or described damage
- **Resilience Level (1 to 3)**:
  - 3 = Many rescue or medical staff involved
  - 2 = Small-scale search or rescue effort
  - 1 = No visible rescue or recovery presence

3 Results

3.1 News Article Sources

We collected a total of 199 unique news articles from 12 distinct international news media agencies. These articles formed the foundation of our analysis pipeline, offering both textual and visual data related to the 2025 Mandalay earthquake. The dataset includes a diverse array of sources spanning regional and global outlets.

The top 10 most frequently contributing agencies are listed in Table 1. Notably, Bangkok Post, BBC, and IBTimes each contributed over a dozen articles, reinforcing the global reach of disaster coverage.

Table 1: Top 10 Contributing News Agencies by Article Count

Rank	News Agency	Article Count
1	bangkokpost.com	19
2	bbc.com	17
3	ibtimes.com	17
4	channelnewsasia.com	12
5	aljazeera.com	11
6	abc.net.au	11
7	abcnews.go.com	8
8	dw.com	7
9	newsweek.com	6
10	punchng.com	5

3.2 Interactive Gallery Product

The output of the above AI reasoning steps is saved as structured JSON files. We first generate an interactive HTML-based gallery that displays:

- A grid of disaster-related images
- Clickable modals showing summaries, captions, tags, sentiment, and discrepancy score

Fig. 1 shows a screenshot of this page; and all data are accessible through this Github link.

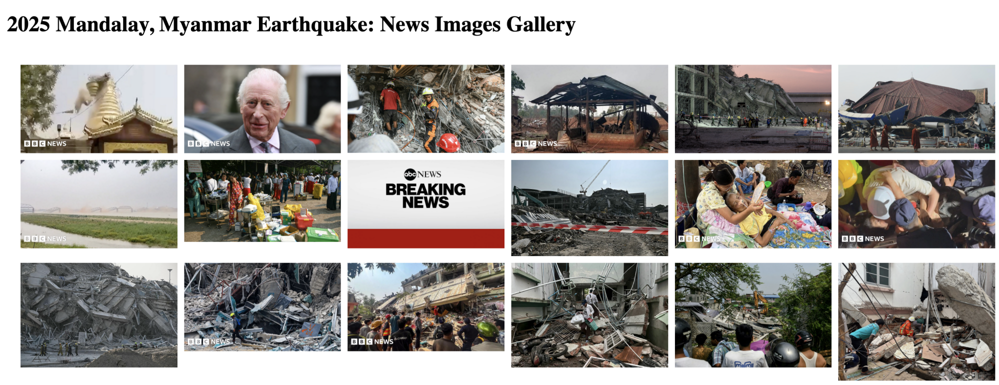

Figure 1: Interactive webpage showing Mandalay Disaster Scenes and Gen-AI generated Summaries. Please visit this link for access.

3.3 Intrinsic Language-vision Discrepancy Scores

It is interesting to evaluate the intrinsic alignment between visual and textual content. Table 2 summarizes the distribution of D across all records. Most values fall in the high range (above 0.83), indicating that in many cases, image captions and textual summaries express somewhat different focuses or content emphasis.

Table 2: Summary Statistics of Discrepancy Scores

Statistic	**Value**
Minimum	0.619
25th Percentile	0.829
Median (50%)	0.872
75th Percentile	0.906
Maximum	1.000
Mean	0.868

3.4 Statistics and Analysis

We applied the LLM-enabled classification process to more than 300 image-summary pairs related to the 2025 Mandalay earthquake. Each record was analyzed using GPT-based prompts to infer sentiment, loss, and resilience levels.

We then generated visual summaries:

- Figure 2: **Histogram of Sentiment**: Categorized into *Tragic*, *Distressing*, *Concerned*, and *Hopeful*
- Figure 3: **Histogram of Loss Level**: Ranging from 1 (None) to 3 (High)
- Figure 4: **Histogram of Resilience Level**: Ranging from 1 (None) to 3 (High)

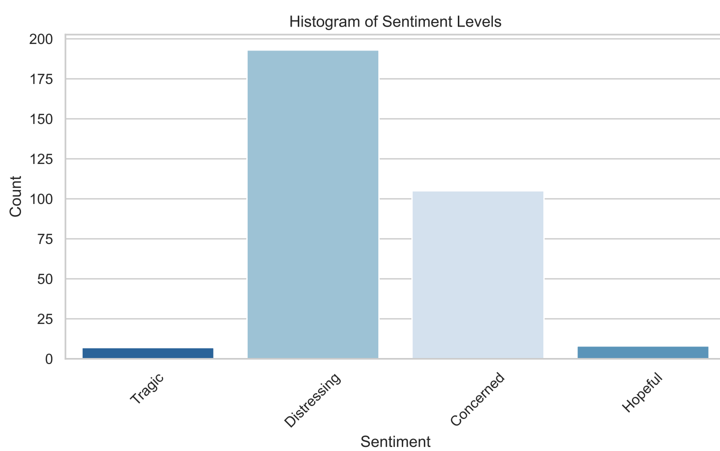

Figure 2: Histogram of sentiment in the collected news coverage.

Finally, we computed a joint distribution of loss vs. resilience, shown in Table 3.

4 Discussion

While GenAI provides powerful tools, from the perspective of Civil Engineering profession, this methodology has limitations:

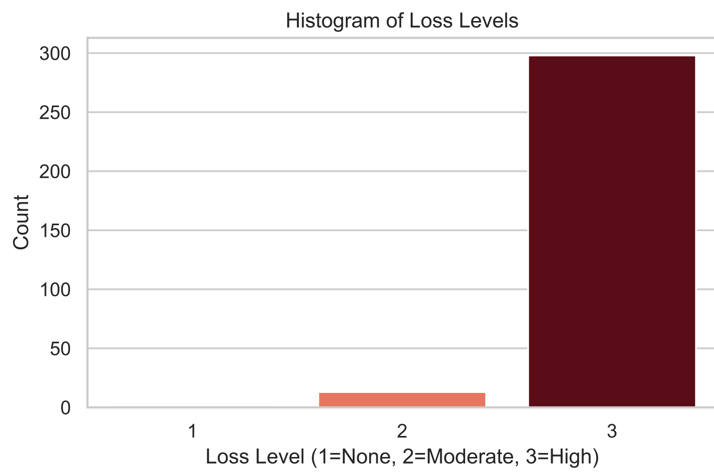

Figure 3: Histogram of loss levels in the collected news coverage.

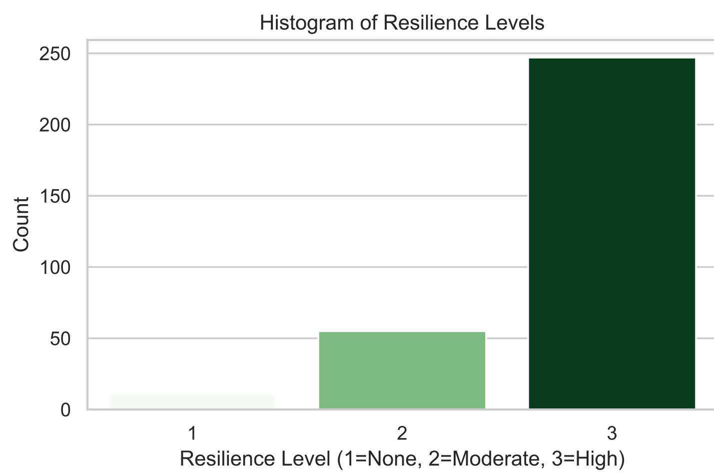

Figure 4: Histogram of resilience levels in the collected news coverage.

Table 3: Joint Matrix of Loss vs Resilience Levels

Loss Level	**Resilience Level**		
	1 (None)	2 (Moderate)	3 (High)
1 (None)	1	0	1
2 (Moderate)	2	5	6
3 (High)	8	50	240

- **No physical grounding**: News and images lack direct linkage to physical damage models
- **Bias and noise**: Both journalism and public expression are subjective and inconsistent
- **Inference instability**: Model outputs may vary with minor rewordings
- **No uncertainty quantification**: Current GPT outputs lack confidence bounds

To make such AI-derived insights actionable for planning, **human-in-the-loop supervision is essential**. Moreover, uncertainty propagation must be controlled to avoid amplification. Future directions include:

- Test-time ensembling and scaling
- Fine-tuning on disaster-specific data
- Probabilistic modeling of inference outputs

5 Conclusion

This work proposes a generative-AI-powered system that extracts, interprets, and visualizes disaster-related visual and textual signals from public news. Through large language and vision models, we produce structured insight from unstructured sources. While challenges remain, this represents a step toward more agile, multi-modal, geospatially aware disaster response intelligence.