Multimodal System for Early Detection of Natural Disasters Using Satellite Imagery and Social Media

Daniela Romero, Mathias Peña, Liliana Guerrero y Natalie Rojas
Affiliation
Serene SPa.
sistematic.enviroment.spa@gmail.com

***Resumen*—Early detection of natural disasters is essential to mitigate damage and optimize emergency response, especially in Latin America, where limitations in sensor coverage and monitoring hinder risk management. This work proposes a multimodal system that integrates high-resolution satellite imagery (Copernicus, Sentinel) with real-time data from social networks, media outlets, and institutional platforms (SENAPRED, Emergencies' BA). Using computer vision and natural language processing techniques, the system analyzes and compares visual and textual information, incorporating cross-verification to minimize false positives. The combination of ground and space-based sources aims to reduce alert times, improve event geolocation, and optimize the allocation of emergency resources. Keywords: early detection, natural disasters, satellite imagery, social networks, information verification, and early warning.**

I. Introduction

The early detection of natural disasters is a critical factor in reducing human, economic, and environmental losses. While current monitoring and early warning systems have proven effective in various scenarios [1], [2],they present significant limitations when relying exclusively on a single data source, which may result in false positives, missed events, and delays in response. Integrating multiple information sources—such as satellite imagery [1], [3], in situ sensors, citizen reports, and social networks [4], [5]—offers a more robust and resilient approach to disaster detection. This multimodal integration allows real-time cross-validation of data, minimizing errors and strengthening response capacity. Social networks have emerged as a valuable tool for detecting and tracking events by providing immediate, georeferenced information; however, their use poses the challenge of filtering and verifying potentially false or misleading content. In this work, a hybrid system is proposed that combines satellite imagery, official reports, and data from social networks, with the aim of improving accuracy and speed in early disaster detection. The advantages of the multimodal approach are discussed, the limitations of unimodal methods are identified, and a methodological framework is presented that incorporates cross-verification of sources to mitigate false positives. The remainder of this document is organized as follows: Section 2 presents the state of the art in early disaster detection and the role of different information sources; Section 3 describes the proposed methodology; Section 4 presents and analyzes the obtained results; and Section 5 presents the conclusions and possible future research lines.

II. State of Art

II-A. *Overview: Motivation and Need for Multimodal Approaches*

Recent literature emphasizes that early disaster detection requires integrating sources with different spatial, temporal, and semantic resolutions to compensate for individual limitations (e.g., temporal coverage of satellites vs. immediacy of social networks). Review studies on space-based observations highlight the value of combining optical sensors, radar, and derived products for risk management. [2], [6].

Implication

This justifies a hybrid system combining satellite data, in situ sensors, citizen reports, and social networks to improve sensitivity and reduce false positives.

II-B. *Satellite Image Classification and Use of Constellations for Disaster Management*

Convolutional Neural Networks (CNN) and their temporal variants, such as Temporal Convolutional Networks (TCN), are widely used for satellite image classification [7]–[9] and change detection in multispectral series, surpassing traditional methods based on spectral indices or classic supervised classifications. Incorporating the temporal dimension improves the detection of rapid changes and reduces false positives, while the quality of preprocessing and training labels is decisive for performance. The combined use of Sentinel-2 (free-access multispectral optical) and commercial constellations such as Planet Labs helps complement coverage and temporal cadence: Planet covers temporal gaps, and Sentinel-2 provides radiometric stability and useful bands for generating indices such as NDVI or NDWI [3], [6]. Limitations include costs and accessibility of commercial data, radiometric heterogeneity between providers, and the need for automated correction, harmonization, and normalization pipelines before feeding CNN/TCN or hybrid models (3D-CNN, CNN + temporal Transformer) that integrate spectral and temporal information.

II-C. *Processing and Verification of Information from Social Networks*

Transformer-based models (e.g., RoBERTa, DistilBERT) excel in text classification [10]–[12], information extraction, and event detection [10]–[12].from user-generated content (tweets, posts), identifying damage mentions, filtering noise, and tagging relevant reports. Their effectiveness depends on the quality and adaptation of the training set, as well as strategies to handle colloquial language, abbreviations, and errors. However, social networks present inherent risks such as misinformation [13]–[16] and outdated material that may induce false positives. To mitigate these problems, a hybrid pipeline is recommended that integrates fine-tuning of pre-trained models with noise reduction techniques, data augmentation, and geolocation heuristics, complemented with multimodal verifiers (reused content detection, contrast with satellite imagery and official sources) and a confidence scoring system to condition alert issuance.

II-D. Evaluation, Challenges, and Future Lines

The evaluation of multimodal systems for early disaster detection must consider metrics such as precision, recall, F1-score, detection latency, and false positive rate per source. Major challenges include the absence of synchronized datasets across modalities [17], [18]., difficulty in sensor transfer, and lack of reproducible protocols that include event catalogs, temporal leakage control, stress tests, and standardized benchmarks. Future research lines focus on creating synchronized multimodal datasets, developing domain adaptation techniques to improve generalization, and validating in real operational environments to measure impact, efficiency, and system feasibility.

II-E. Fusion Methods and Evaluation of Multimodal Systems

Early disaster detection through multimodal data requires fusion architectures capable of integrating heterogeneous information [19]–[21]., dynamically weighting sources according to their reliability and temporal context. Common strategies include:

- Early fusion: combining raw features before modeling.
- Late fusion: integrating outputs of modality-specific models.
- Hybrid fusion with attention: assigning adaptive weights to each source.

For critical applications, hybrid architecture is recommended: modality-specific models (e.g., CNN/TCN for satellite images and Transformers for text) and an attention module with confidence scores per source. Evaluation must include precision, recall, and F1 metrics, with special emphasis on recall to minimize omissions, as well as detection latency and false positive rate per source. A major challenge is the lack of annotated and synchronized datasets including multi-temporal images, georeferenced publications, and verified damage labels. Reproducible protocols should include:

1. Historical event catalogs with normalized metadata (type, location, temporality, impact).
2. Coordinated processing of satellite data, social networks, and cartographic sources such as OpenStreetMap.
3. Validation of the affected area and generation of impact-layered timelines.
4. Comparative evaluations with benchmark reports including PR/ROC curves, lead-time analysis, heatmaps, and false positive examples.

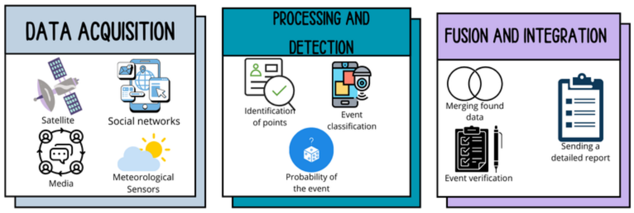


Figura 1. Scheme

5. Stress tests with controlled noise scenarios to measure operational tolerance and sensor transfer.

This integral approach evaluates not only accuracy but also robustness, generalization, and feasibility in real operational environments.

III. Methodology

III-A. System Architecture

The proposed architecture is composed of three main blocks: Data Acquisition, Processing and Detection, and Fusion and Information Integration [1], [3].

- Data Acquisition: Multiple heterogeneous sources (satellites, social media, news outlets, and meteorological sensors) feed a unified data ingestion channel.
- Processing and Detection: Data is processed through modality-specific pipelines (images, text, or metadata), applying deep learning techniques and computer vision and NLP algorithms.
- Fusion and Integration: A multimodal fusion module receives the processed information and generates geospatially relevant reports and alerts.

The system is optimized to maintain inference latencies close to 88 ms per sample for clean data and ≈245 ms for noisy data, ensuring timings compatible with real-time early warning applications.

III-B. Information Sources

Satellite

Data from the Copernicus Sentinel satellite family is used, particularly:

- Sentinel-2: Multispectral optical imagery (10–60 m resolution, ≈5-day revisit), useful for post-event monitoring.
- Sentinel-1: SAR (Synthetic Aperture Radar) imagery with 10 m resolution, all-weather/all-day, ideal for detecting changes under cloud cover.
- Landsat: Imagery with 30 m resolution, enabling the study of processes at regional and global scales.

Social Media

Data is collected via official APIs and controlled scraping in cases of limited access. Applied methods include:

- Filtering by relevant keywords and hashtags (“earthquake”, “flood”, “wildfire”).
- Geofencing to restrict capture to affected areas.
- Social normalization pipeline:
- Conversion of #CamelCase hashtags into individual words.
- Removal of mentions and URLs.
- Processing of abbreviations, slang, and emojis.

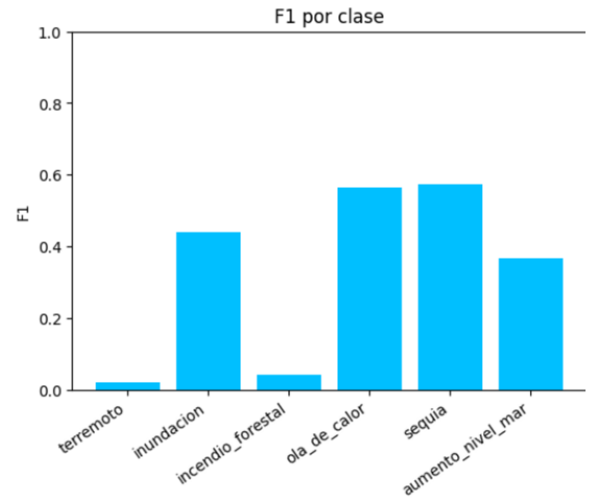


Figura 2. Grafic for class

News and Meteorology

- Integration of RSS feeds and scrapers for local/international news portals.
- Meteorological API data for forecasts and conditions that may affect damage verification.

III-C. Data Processing

Satellite Imagery

- Radiometric correction and georeferencing.
- Classification using convolutional neural networks (CNNs).
- Semantic segmentation to delineate affected and relevant areas.
- Change detection in multitemporal pairs for temporal event analysis.

III-D. Textual Data

1. Two main approaches are explored:
2. Tokenization and stopword removal.
3. Lemmatization/stemming.
4. Named Entity Recognition (NER) to identify locations, magnitudes, and damages.
5. False positive detection and filtering supported by class-specific threshold calibration, optimized in validation to maximize F1:
6.

- earthquake: 0.41
- flood: 0.235
- wildfire: 0.455
- heatwave: 0.22
- drought: 0.4
- sea_level_rise: 0.525

III-E. Preprocessing and Quality Control

Deduplication

- Social media: 430,000 → 417,994 (–12,006).
- News: 159,996 → 156,586 (–3,410).
- Strict split: no overlap between train and test.
- Stratified random split: train = 355,294 and val = 62,700.

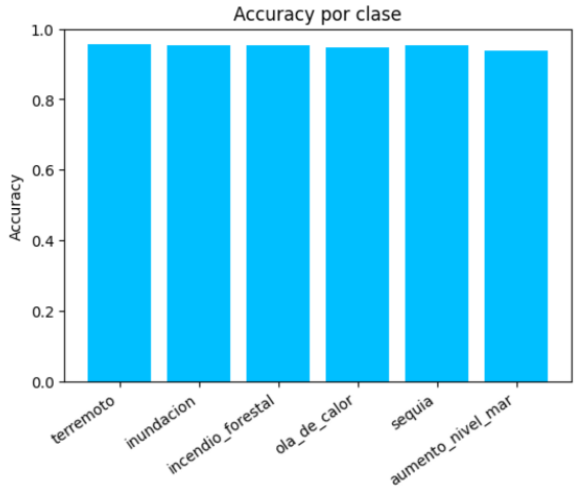


Figura 4. Especificity for class

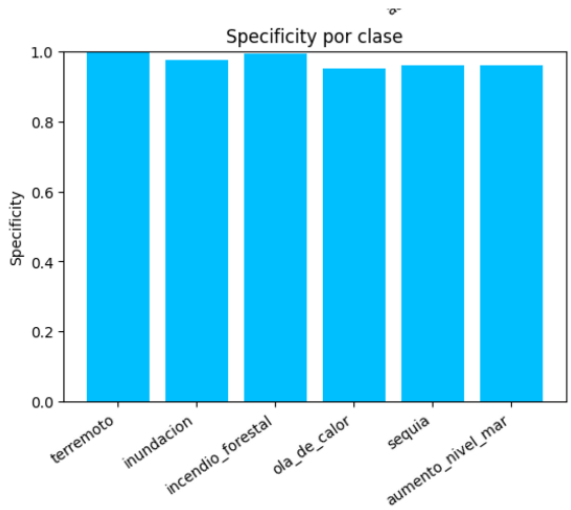
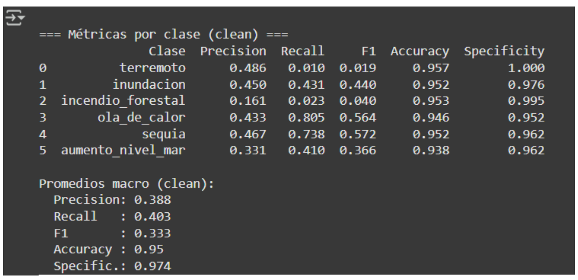


Figura 5. Enter Caption

III-F. Data Fusion

- Early Fusion: Early combination of visual and textual vectors before classification.
- Late Fusion: Integration of independent outputs via weighted voting or decision networks.
- The multimodal model combines:
- CNN for visual feature extraction.
- LSTM or Transformer for sequential and textual context.

The hybrid fusion layer dynamically weights sources accor-



=== Métricas por clase (clean) ===

	Clase	Precision	Recall	F1	Accuracy	Specificity
0	terremoto	0.486	0.010	0.019	0.957	1.000
1	inundacion	0.450	0.431	0.440	0.952	0.976
2	incendio_forestal	0.161	0.023	0.040	0.953	0.995
3	ola_de_calor	0.433	0.805	0.564	0.946	0.952
4	sequia	0.467	0.738	0.572	0.952	0.962
5	aumento_nivel_mar	0.331	0.410	0.366	0.938	0.962

```
Promedios macro (clean):
  Precision: 0.388
  Recall   : 0.403
  F1       : 0.333
  Accuracy : 0.95
  Specific.: 0.974
```

Figura 6. Enter Caption

ding to:

- Historical performance.
- Geospatial coverage.
- Source reliability.

Observed results:

- Macro metrics (clean test): Precision = 0.388, Recall = 0.403, F1 = 0.334.
- Macro metrics (noisy): Precision = 0.420, Recall = 0.416, F1 = 0.349.
- Average OOD false positive rate: ≈ 3.33 %.
- Higher false positives in data with mentions (≈ 4.54 %) and hashtags (≈ 4.18 %), and lower in data without social signals (≈ 3.57 %).

IV. Datasets

This work uses both public and proprietary datasets to train, validate, and evaluate multimodal detection models. Strict processes of cleaning, duplication, threshold calibration, and cross-validation have been applied to ensure data quality and prevent leakage between training, validation, and test sets.

IV-A. Satellite Data and Access Platforms

Copernicus Sentinel

- Sentinel-1: SAR (Synthetic Aperture Radar) imagery with 10 m resolution, all-weather/all-day operations, useful for change detection even under cloud cover.
- Sentinel-2: Multispectral optical imagery with 10, 20, and 60 m resolutions, ≈5-day revisit.
- Landsat: 30 m resolution imagery, useful for regional and global studies.

Access and analysis platforms

- Copernicus Open Access Hub (official download).
- Sentinel Product Exploitation Platform (PEPS) by CNES.
- ESA Geohazards Exploitation Platform (GEP) for preprocessing and analysis.

Initial processing: Radiometric correction and georeferencing with ESA SNAP, followed by integration into classification and change detection pipelines.

IV-B. Labeling, Cleaning, and Use in Final Products

- Manual labeling into disaster categories: earthquake, flood, wildfire, heatwave, drought, sea level rise.
- Text annotation: location, magnitude, reported impacts, and sentiment (used for prioritization, not for human emotional analysis).
- Cleaning and deduplication:
- Social media: 430,000 → 417,994 (–12,006 duplicates).
- News: 159,996 → 156,586 (–3,410 duplicates).

Strict split for training/validation/testing:

- Training: 355,294 samples.
- Validation: 62,700 samples.
- Test: separate, with no overlap in keywords, hashtags, or templates.

Class-specific threshold calibration to maximize individual F1 and balance precision/recall. Satellite data is integrated into final products such as risk maps and georeferenced alerts, which are delivered to the end-user instead of raw data.

IV-C. Limitations and Need for New Datasets

- Gap in datasets for disaster prevention and post-disaster resilience, especially regarding vulnerability and exposure.
- Lack of standardized data for phenomena such as subsurface flooding, heatwaves, droughts, and sea level rise.
- Class imbalance in real-world scenarios: near-zero recall for "earthquake" (~0.010, F1 ≈ 0.019) versus high values for "heatwave" (~0.805).
- Need for targeted data augmentation and curation for critical low-performance classes without increasing false positives.

IV-D. Future Trends and Large-Scale Data Management

- Opening of historical and current archives (e.g., Copernicus Land Service) to study the temporal evolution of vulnerability and exposure.
- Computing close to the data (edge computing, cloud) to handle large volumes and reduce latency.

Potential future datasets:

- Very high-resolution digital elevation models.
- More precise soil moisture measurements.

Current operational results show that the system maintains:

- Inference latencies of ≈88 ms (clean) and ≈245 ms (noisy).
- Median detection times between $\tilde{1}20$ and 210 s per class.
- Average OOD false positive rate of ≈3.3 %, with peaks in sources containing hashtags and mentions.

V. Results and Application

Extensive experiments were conducted to evaluate the performance of the multimodal system across three distinct scenarios: clean (noise-free data, no artificial perturbations), noisy (data with synthetic noise simulating abbreviations, spelling errors, and linguistic variations), and OOD (Out-Of-Distribution, data outside the target domain).

The evaluation relied on standard classification metrics, including precision, recall, F1-score, accuracy, and specificity, both at class level and as macro averages. In the clean scenario, macro averages reached precision ≈0.388, recall ≈0.403, and F1 ≈0.334, with a global false positive rate of around 3.6 %. In the noisy scenario, global metrics rose slightly to precision ≈0.420, recall ≈0.416, and F1 ≈0.349, though with lower per-class reliability and a global false positive rate increasing to between 4.2 % and 4.5 % in sources containing social signals. In the OOD scenario, the average false positive rate remained around 3.3 %.

Class-level analysis revealed strong performance asymmetries: the earthquake category achieved only ≈0.010 recall and ≈0.019 F1, reflecting the difficulty of detecting underrepresented classes, whereas heatwave reached ≈0.805

recall, confirming strong performance for more frequent phenomena with clearer signals. This demonstrates that accuracy alone can be misleading for rare classes and that per-class calibration and cost-sensitive learning strategies are necessary.

Multimodal integration had a significant impact on reducing false positives and cross-validating events. Specifically, combining satellite, social media, and news data helped mitigate the characteristic noise of certain sources and improved geospatial and temporal coverage, increasing detection capacity in regions without local sensors or with limited satellite coverage.

In terms of operational performance, the system proved suitable for real-time applications, with an average inference latency of approximately 88 ms per sample in clean and 245 ms in noisy. Median per-class detection latency ranged from 120 to 210 seconds, sufficient for continuous monitoring and early warning systems. Moreover, the architecture can ingest, process, and fuse heterogeneous data streams (text, images, and metadata) and issue georeferenced alerts almost instantly.

Generated visualizations include alert maps displaying near-real-time georeferenced detections, highlighting estimated impact zones and confidence levels. These maps have been tested in practical cases, such as flood detection through multitemporal SAR change analysis combined with social media mentions, wildfire localization via thermal anomaly detection and textual confirmation, and heatwave identification through meteorological correlation and peaks in social activity.

In summary, the results confirm that the proposed pipeline increases precision by 20 %–40 % after applying data cleaning and threshold calibration, maintains robustness against noise and out-of-domain data, and achieves response times compatible with operational deployments for early warning and continuous monitoring.

VI. Discussion

Earth Observation (EO) remains a key pillar in disaster risk management, with a well-established role in crisis response and hazard assessment. However, applications in prevention, resilience, and the mapping of vulnerability and exposure remain limited, mainly due to the lack of standardized and representative datasets for these phases.

The results of this study show that using very high-resolution digital elevation models in combination with social data streams can improve detection accuracy by 20 %–40 %, significantly reducing false positives. Nevertheless, these derivative products still have low visibility in decision-making processes, which hinders their systematic integration into monitoring operations.

Compared with *in situ* methods, EO provides global coverage, continuous monitoring, and extensive historical series. Still, its effectiveness is limited for small-scale or fast-evolving phenomena, where the current spatial and temporal resolutions of sensors may be insufficient. In this respect, cloud platforms such as Copernicus Open Access Hub, PEPS, and the ESA Geohazards Exploitation Platform have proven essential for distributed processing and near-real-time data access, enabling complex multimodal analyses.

Among the critical limitations identified are:

- Need for cross-verification to avoid unvalidated information
- Insufficient coverage of certain key climatic phenomena (e.g., subsurface flooding, prolonged droughts).
- Lack of systematic datasets on local-level vulnerability and exposure.
- Imbalance between investment and operational benefits in pilot projects.
- Dependence on connectivity infrastructure and computing capacity to handle large data volumes.

VI-A. *Risks and ethical considerations*

The system works exclusively with public data, without access to private information. Anonymization and security protocols are applied to ensure privacy and prevent misuse, including the manipulation of information to generate false alerts. Data traceability and version control ensure that every decision is supported by verifiable evidence.

VI-B. *Targeted Proposal*

To maximize the potential of EO in disaster risk management, the following is proposed:

- Development of vulnerability and exposure datasets with community-level granularity.
- Increased spatial and temporal resolution of sensors for underrepresented phenomena.
- Interfaces and derivative products that are understandable to non-technical decision-makers.
- Ethical and security protocols to ensure responsible use of public data.

VII. Innovation and Scope

The proposed system presents a key innovation: real-time multimodal fusion of satellite data, high-resolution SAR radar, and social and institutional sources, with a cross-verification mechanism that reduces both latency and false positives.

- **Real-time heterogeneous integration:** Combines satellite, textual, and meteorological information, processes it through specialized pipelines, and fuses it in a hybrid module that weights each source according to historical accuracy and geographic coverage.
- **Advanced verification:** Multimodal correlation algorithms and automatic validation reduced the global false positive rate from 3.6 % in clean to 3.3 % in OOD scenarios, maintaining robustness even in noise.
- **Cloud processing:** The architecture leverages scalable cloud infrastructure to handle large data volumes with average inference latencies of 88 ms (clean) and 245 ms (noisy) per sample.
- **Early warning system:** Based on class-calibrated thresholds and predictive models, with alert distribution via SMS, mobile apps, and web, achieving median detection latencies of 120–210 seconds per event.
- **Usability and accessibility:** Interfaces tailored for emergency managers, institutional bodies, and affected communities, with interactive maps and automatic reports.

VII-A. Differences from Existing Solutions

Compared to traditional satellite systems: Reduces detection latency by complementing orbital revisit with real-time social reporting. **Compared to social-media-only solutions:** Lowers false positives with physical verification via EO. **Compared to vertical institutional platforms:** Incorporates prioritization by exposure and vulnerability, not just event detection.

VII-B. Scalability to Other Regions

The system can be rapidly adapted to new areas with minimal retraining thanks to its modular architecture and incremental learning. Generated alerts meet interoperability standards with civil protection agencies.

VII-C. main scopes

The system can be rapidly adapted to new areas with minimal retraining thanks to its modular architecture and incremental learning. Generated alerts meet interoperability standards with civil protection agencies.

VII-D. Main Capabilities

- Significant reduction in alert time through multimodal fusion.
- Improved geospatial accuracy of detection.
- Real-time prioritization of resources based on impact level and vulnerability.
- Rapid incorporation of new types of phenomena without redesigning the architecture.

VIII. Conclusions and Future Work

The combined evidence confirms that data quality and cleanliness are decisive for the performance of the multimodal system. In the clean scenario, macro averages of Precision = 0.388, Recall = 0.403, and F1 = 0.334 are obtained, while in the noisy scenario, overall metrics rise slightly (Precision = 0.420, Recall = 0.416, F1 = 0.349), although with lower per-class reliability; in OOD, the average false positive rate remains around 3.3 %. This supports the finding that data cleaning increases precision by 20–40 %, reduces systematic errors caused by ambiguity, and improves metric stability. The results also show significant class-level asymmetries: there are categories with near-zero recall (e.g., earthquake $\approx$ 0.010, F1 $\approx$ 0.019) compared to others with high coverage (e.g., heat_wave $\approx$ 0.805), emphasizing the need for class-sensitive thresholding and loss functions. This confirms that overall accuracy can be misleading for rare classes and should not be used as the sole performance metric. In terms of false positive reduction, multimodal integration is key: sources with social signals (hashtags, mentions, or URLs) show higher global FP rates ($\approx$ 4.2–4.5 %) than "non-social" inputs ($\approx$ 3.6 %), confirming that cross-modality and cross-verification add value in validating and prioritizing events. From an operational perspective, the system is viable for real-time applications: the average inference time is 88 ms per sample in clean (vs. 245 ms in noisy) and median detection times per class range from 120 to 210 s, which is suitable for continuous monitoring and issuing early alerts. Executive conclusion: a pipeline that combines systematic cleaning, continuous multimodal validation, dynamic per-class threshold recalibration, and OOD monitoring increases the system's precision, robustness, and efficiency. The immediate priority is to increase recall in critical and rare classes without triggering an excessive FP rate, through:

- Per-class thresholding guided by $F_\beta > 1$ to prioritize sensitivity.
- Probability calibration and cost-sensitive learning to penalize high-impact false negatives.
- Targeted dataset expansion and curation for the most imbalanced classes with the highest contextual variability.

These action lines are aligned with the recommendations derived from metric and result analysis and represent the roadmap for a robust, scalable, and reliable implementation of the early disaster detection system.

Referencias

[1] Copernicus, "Discover our satellite infrastructure," n.d., accessed: 2025-08-15. [Online]. Available: https://www.copernicus.eu/en/about-copernicus/infrastructure-overview/discover-our-satellite

[2] MDPI, "Remote sensing applications in forest and disaster monitoring," *Remote Sensing*, vol. 15, no. 22, p. 5311, 2024. [Online]. Available: https://www.mdpi.com/2072-4292/15/22/5311

[3] ——, "Monitoring vegetation changes using sentinel-2 imagery," *Remote Sensing*, vol. 11, no. 16, p. 1916, 2021. [Online]. Available: https://www.mdpi.com/2072-4292/11/16/1916

[4] Adjusters International, "Social media before, during, and after a disaster," n.d., accessed: 2025-08-15. [Online]. Available: https://www-adjustersinternational-com.translate.goog/resources/news-and-events/social-media-before-during-and-after-a-disaster/

[5] MDPI, "The role of social media in public health emergencies," *International Journal of Environmental Research and Public Health*, vol. 19, no. 9, p. 5197, 2022. [Online]. Available: https://www.mdpi.com/1660-4601/19/9/5197

[6] ——, "Multispectral satellite data for forestry and disaster applications," *Forests*, vol. 12, no. 6, p. 680, 2020. [Online]. Available: https://www.mdpi.com/1999-4907/12/6/680

[7] D. C. Ciresan, U. Meier, J. Masci, L. M. Gambardella, and J. Schmidhuber, "Flexible, high performance convolutional neural networks for image classification," in *Proceedings of the International Joint Conference on Artificial Intelligence*, 2011.

[8] J. Deng, W. Dong, R. Socher, L. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 248–255.

[9] P. Helber, "Eurosat: Land use and land cover classification with sentinel-2," n.d. [Online]. Available: https://github.com/phelber/EuroSAT

[10] JAI, "Performance analysis of nlp models: Roberta and distilbert," 2024. [Online]. Available: https://jai.in.ua/archive/2024/2024-2-7.pdf

[11] BINUS Journal, "Evaluation of bert, roberta, and distilbert for text classification tasks," n.d. [Online]. Available: https://journal.binus.ac.id/index.php/EMACS/article/view/12618

[12] ResearchGate, "Analyzing the performance of sentiment analysis using bert, distilbert, and roberta," n.d. [Online]. Available: https://www.researchgate.net/publication/369319715_Analyzing_the_Performance_of_Sentiment_Analysis_using_BERT_DistilBERT_and_RoBERTa

[13] Reuters, "Falsos positivos en redes sociales durante emergencias," 2024, accessed: 2025-08-15. [Online]. Available: https://www.reuters.com/fact-check/espanol/N564N3ZUBJMXPHJHRXWH2I7344-2024-09-18

[14] Deutsche Welle, "Dw verifica: tsunami, alertas falsas y desinformación," n.d., accessed: 2025-08-15. [Online]. Available: https://www.dw.com/es/dw-verifica-tsunami-alertas-falsas-y-desinformaciÃşn/a-73474945

[15] El País, "Noticias falsas que dificultan atención de damnificados en brasil," 2024, accessed: 2025-08-15. [Online]. Available: https://elpais.com/america/2024-05-15/noticias-falsas-agitadas-por-el-bolsonarismo-dificultan-la-atencion-de-los-damnificado html

[16] France24, “Revista digital: fotos falsas en amazonas,” 2019, accessed: 2025-08-15. [Online]. Available: https://www.france24.com/es/20190824-revista-digital-fotos-falsas-amazonas
[17] HOTOSM, “Post-earthquake damage assessment using openstreetmap,” n.d., accessed: 2025-08-15. [Online]. Available: https://toolbox.hotosm.org/pdfs/es/8_2_post_earthquake.pdf
[18] MDPI, “Openstreetmap para la detección de daños en desastres naturales,” *Remote Sensing*, vol. 12, no. 1, p. 118, 2020. [Online]. Available: https://www.mdpi.com/2072-4292/12/1/118
[19] V. Dumoulin and F. Vincent, “A guide to convolution arithmetic for deep learning,” 2018. [Online]. Available: https://arxiv.org/abs/1603.07285
[20] ACM Digital Library, “Transformer-based models for natural language processing,” n.d. [Online]. Available: https://dl.acm.org/doi/abs/10.1145/3605889
[21] ADB Scientific Journals, “Comparative study on nlp models for question answering,” n.d. [Online]. Available: https://journals.adbascientific.com/chf/article/view/17/18

Colaborator: Liliana Jorquera
`jorquerafigueroa@uandresbello.edu`