

Multimodal Spatiotemporal Neural Digital Twin Framework for Early Natural Disaster Prediction and Resilient Decision Support

Dannuri Mounika

SR University

Johnson Kolluri

SR University

Sunil Srinivas

TKR College of Engineering and Technology

Kiran Siripuri

`skiran@sithyd.siu.edu.in`

Symbiosis International University

Research Article

Keywords: Natural disaster risk assessment, Environmental monitoring, Multimodal data fusion, Earth observation, Neural digital twin, Dynamic graph reasoning, Disaster resilience, Early warning systems

Posted Date: September 17th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10894859/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Multimodal Spatiotemporal Neural Digital Twin Framework for Early Natural Disaster Prediction and Resilient Decision Support

, Dannuri Mounika^{1,2}, Johnson Kolluri¹, Sunil Srinivas²,
Kiran Siripuri^{3*}

¹School of Computer Science and Artificial Intelligence, SR University,
Warangal, 506371, Telangana, India .

²CSE-AIML, TKR College of Engineering and Technology, Hyderabad,
500097, Telangana, India.

^{3*}Symbiosis Institute of Technology, Hyderabad Campus, Symbiosis
International (Deemed University), Pune, India.

*Corresponding author(s). E-mail(s): skiran@sithyd.siu.edu.in;

Contributing authors: dmounikaresearch@gmail.com;
johnson.kolluri@gmail.com; sunilsrinivasb@gmail.com;

Abstract

Background:

The growing occurrence and intensifying severity of natural disasters are intensifying environmental degradation, ecosystem disruption, infrastructure damage, and risks to vulnerable populations, creating an urgent need for reliable environmental monitoring and resilient disaster-management systems. Although deep learning has improved hazard prediction, many existing approaches remain limited to isolated forecasting tasks, depend on individual or homogeneous data sources, and use static learning mechanisms that inadequately represent rapidly evolving disaster conditions. Moreover, converting predictive outputs into timely and context-aware response actions remains a major challenge. This study proposes RADIANT-Net, a resilient adaptive disaster intelligence with agentic neural twins network framework that unifies multimodal environmental data integration, dynamic risk reasoning, disaster evolution modeling, and intelligent decision support.

Results:

The proposed framework combines multimodal representation learning, dynamic graph reasoning, Neural Digital Twin technology, and agentic artificial intelligence to analyze satellite imagery, meteorological observations, sensor measurements, geospatial information, population exposure data, and disaster-related textual information jointly. The evaluation was conducted via publicly available Earth observations, precipitation, weather, population, geographic, and disaster-related datasets. RADIANT-Net achieved an overall accuracy 92.00% accompanied by a precision 92.68% and a recall of 92.00%, a F1 score of 91.96%, and 97.68% area under the receiver operating characteristic (ROC) curve (AUC). It further obtained an intersection over union of 0.8576, a mean absolute error (MAE) of 0.054, a root mean square error (RMSE) of 0.169, while maintaining a decision latency of 0.058 s. The findings demonstrate effective spatial risk identification, robust predictive performance, and rapid decision support under heterogeneous and incomplete environmental observations.

Conclusions:

RADIANT-Net provides an integrated approach that extends conventional disaster forecasting beyond hazard prediction toward dynamic environmental risk assessment and adaptive emergency decision-making. By modeling evolving interactions among hazards, affected environments, critical infrastructure, and exposed populations, the framework can support early warning, evacuation planning, resource allocation, rescue prioritization, and postdisaster recovery. Importantly, its integration of heterogeneous environmental observations enables more comprehensive identification of vulnerable areas and evolving risk conditions, supporting timely interventions that can reduce environmental and societal impacts. These findings demonstrate the potential of intelligent multimodal environmental monitoring to strengthen disaster resilience, climate adaptation, ecosystem protection, and sustainable environmental risk management.

Keywords: Natural disaster risk assessment, Environmental monitoring, Multimodal data fusion, Earth observation, Neural digital twin, Dynamic graph reasoning, Disaster resilience, Early warning systems

1 Introduction

Natural disasters and environmental risk

Natural disasters are a growing threat to environmental sustainability, infrastructure resilience, and human well-being. Floods, cyclones, wildfires, landslides, droughts, and earthquakes can disrupt ecosystems, degrade natural resources, damage critical infrastructure, and expose vulnerable communities to prolonged social and economic consequences. Climate change, rapid urbanization, land-use modification, and increased population exposure further intensify these risks Windheuser et al. [2023], Wang et al. [2025a], Tiggeloven et al. [2025]. The impacts of disasters are rarely isolated; disruptions in transportation, energy, water and communication networks can amplify

environmental damage and hinder emergency response Wang et al. [2025a]. disaster risk reduction therefore requires monitoring systems that can identify emerging hazards, track changing environmental conditions, and provide timely information for early warning and response planning.

Evolution of Artificial Intelligence and machine learning in disaster management

Bibliometric and methodological evaluations have demonstrated the growing integration of artificial intelligence (AI) and machine learning (ML) in hazard risk assessment, disaster prediction, and emergency monitoring Lin et al. [2025]. Traditional disaster management relies on empirical observations, historical statistical models and numerical simulations based on physics Akhyar et al. [2024]. Although they are fundamental to understanding physical hazard processes, conventional models struggle to digest heterogeneous high-speed large-scale data streams from satellite constellations, meteorological networks and IoT sensor arrays Wibowo et al. [2025]. Although early ML techniques improved risk forecasting and emergency analysis Chamola et al. [2021], the literature emphasizes that most current frameworks remain task-specific and prediction-oriented. They prioritize standalone forecast accuracy over real-time operational decision support, leaving a significant gap in automated disaster response and resource allocation.

Deep learning, multimodal sensing, and spatiotemporal hazard perception

To process complex environmental data streams, domain-specific deep learning architectures have advanced rapidly. The macroscale atmospheric dynamics for the forecasting of weather and the modeling of tropical cyclones is captured through global 3D neural networks Bi et al. [2023] and specialized deep architectures Rahman et al. [2025], Huang et al. [2025]. For inundation dynamic modeling, transformer-based models (e.g. floodSformer) Pianforini et al. [2024], explainable LSTM networks Li et al. [2024], and hybrid transformer-LSTM frameworks Cao et al. [2024], Zhang et al. [2025] effectively model nonlinear spatiotemporal flood behavior. Furthermore, cross-modal satellite fusion – integrating Sentinel-1 synthetic aperture radar (SAR) observations with Sentinel-2 optical imagery Portalés-Julià et al. [2025] alongside multisensor deep learning Fawakherji and Hashemi-Beni [2025] significantly improves flood boundary mapping under severe weather constraints.

Currently, massive machine-type communication networks combined with temporal convolutional networks (mTCNs) facilitate high-frequency IoT disaster monitoring Umadevi et al. [2025]. The spatial topologies of sensor networks and drone services are modeled using dynamic graph neural networks (GNNs) Shafiee Roudbari et al. [2024], Akram et al. [2025], spatial graph architectures Taghizadeh et al. [2025], Wang et al. [2025b], Yao et al. [2025], and multimodal wildfire detection frameworks [34]. Parallel predictive advancements span landslide susceptibility forecasting via iterative and transformer-based models Zhu et al. [2020], Shadrin et al. [2024], as well as seismic event detection via hybrid ETAS models Zhang et al. [2024], P-wave convolutional neural networks Zhexebay et al. [2025], explainable uniform hazard acceleration

models Shokrgozar-Yatimdar and Chen [2025], and knowledge-guided neural architectures. Nevertheless, these spatiotemporal perception models function strictly as isolated predictive tools; they lack embedded mechanisms for real-time evacuation planning, automated decision intelligence, or adaptive response execution.

Physics-guided simulation, digital twins, and autonomous decision intelligence

To enforce physical consistency in data-driven models including physics-informed machine learning (PIML) and physics-informed neural networks (PINNs) integrate governing conservation laws directly into neural loss functions, improving simulation fidelity for multiscale geoscience hazards Karniadakis et al. [2021], Habib et al. [2025], Mustajab et al. [2024]. Currently, neural digital twin architectures combine 3D city models, physical simulations, and real-time IoT adjustments to support climate resilience and hazard visualization Yun et al. [2022], Riaz et al. [2023]. On the operational front, multiagent coordination platforms (such as ResQConnect) facilitate human-centered emergency response and logistics management.

Despite these developments, a critical operational gap persists. Existing digital twins remain simulation oriented and conceptual, lacking integration with multimodal foundation models or agentic AI. Furthermore, emerging multiagent platforms operate without direct coupling to real-time spatiotemporal deep learning, dynamic graph intelligence, or neural digital twin simulations Aththanayake et al. [2026]. An end-to-end framework that unifies multimodal spatiotemporal perception, physics-guided future hazard simulation, and resilient multiagent decision intelligence remains an unfulfilled necessity in environmental hazard management.

Limitations of existing approaches

The current study has several important limitations:

- **Fragmented disaster modeling:** Many systems focus on a single task, such as hazard prediction, mapping, damage assessment, or evacuation planning.
- **Static data integration:** Conventional multimodal fusion strategies have limited ability to adapt to changing environmental conditions, missing observations, and variations in data reliability.
- **Restricted spatial adaptability:** Static graph structures cannot adequately represent the evolving relationships among hazardous processes, infrastructure, affected regions, and exposed populations.
- **Separation of prediction and simulation:** Many digital twin applications emphasize scenario simulation without continuous integration with learning-based hazard prediction.
- **Limited operational intelligence:** Predictive output is rarely directly connected with adaptive decision-making for evacuation, resource allocation, rescue, and recovery.

Objectives and contributions

To address these gaps, the proposed RADIANT-Net objectives and contributions are as follows:

- Multimodal environmental intelligence integrates satellite imagery, meteorological observations, Internet of Things sensor measurements, geospatial information, population data, and textual information related to disasters within a unified framework.
- Complementary feature learning: This introduces a multimodal foundation encoder to learn spatial, temporal, and semantic features derived from heterogeneous environmental data.
- Dynamic spatial intelligence: A dynamic graph intelligence module is developed to capture the evolving relationships between affected regions, critical infrastructure, and exposed populations.
- Disaster evolution modeling: This method employs a neural digital twin to represent changing disaster conditions and estimate future hazard states across multiple time horizons.
- Integrated risk prediction: Jointly estimates the occurrence of disasters, the type of hazard, the spatial risk, and the severity through multitask prediction.
- Adaptive decision intelligence: This method incorporates agentic artificial intelligence to transform predictive information into recommendations for evacuation, resource allocation, rescue prioritization, emergency communication, and recovery.
- Environmental contribution: This study provides an integrated mechanism for converting heterogeneous environmental observations into actionable risk information, with the aim of strengthening early warning, protecting environmentally vulnerable areas, reducing disaster impacts, and supporting climate-resilient environmental resource coordination.

Research gap

Taking into account limitations, the following research gaps are identified.

- A unified approach to integrate multimodal environmental observations, dynamic spatial reasoning, hazard prediction, and disaster evolution modeling remains limited.
- Existing systems do not adequately represent the changing environmental and spatial dependencies as disaster conditions evolve.
- The connections among predictive intelligence, disaster simulation, and actionable emergency decision support remain insufficiently explored.

Addressing these gaps is necessary to move beyond isolated hazard prediction toward adaptive environmental intelligence capable of supporting resilient disaster management under uncertain and rapidly changing conditions.

Summary of the Research Gap

The reviewed studies, as shown in Table 1, indicate that AI, deep learning, multimodal learning, GNNs, digital twins, physics-informed learning, explainable AI, and agentic AI individually improve disaster management.

Table 1: Table 1 Comparative overview of recent studies on multimodal disaster prediction methods

Research Direction	Representative Studies	Methods Used	Key Limitation
AI-based disaster management	Tiggeloven et al. [2025]–Chamola et al. [2021]	ML, AI reviews, early warning frameworks	Often fragmented and task-specific
Deep learning prediction	Fawakherji and Hashemi-Beni [2025]– Li et al. [2024], Zhang et al. [2024]	CNN, LSTM, Transformer, hybrid models	Limited decision-support capability
Multimodal learning	Portalés-Julià et al. [2025], Umadevi et al. [2025]	Remote sensing fusion, sensor fusion, TCN	Mostly static data fusion
Graph intelligence	Shafiee Roudbari et al. [2024], Akram et al. [2025]	GNN, Dynamic GNN, IoT/drone data	Limited simulation and response planning
Digital Twin systems	Yun et al. [2022], Riaz et al. [2023]	Hybrid modeling, 3D city modeling, simulation	Weak integration with deep learning and AI agents
Physics-informed and explainable AI	Karniadakis et al. [2021], Shokrgozar-Yatimdar and Chen [2025]	PINNs, Explainable Deep Learning	Often separated from operational decision-making
Agentic AI	Aththanayake et al. [2026]	Multiagent platform, human-centered response	Limited multimodal prediction and disaster simulation

2 Materials and Methodology

2.1 Data Acquisition and Sources

The prediction of natural disasters is a difficult task that is based on information from various domains, since no one dataset can completely describe the complex processes leading to the occurrence of disasters. Disaster development is determined by a

combination of environmental conditions, variability in weather, terrain characteristics, historical disaster behavior, and human activities. Therefore, the proposed RADIANT-Net is designed to fuse heterogeneous datasets from publicly available and trusted repositories to provide a comparative representation of evolving disaster conditions. The primary source of spatial information is satellite remote sensing data. Changes in the Earth’s surface before, during, and after disaster events are assessed via optical imagery from Sentinel-2, Landsat-8/9, and MODIS completed by Sentinel-1 synthetic aperture radar (SAR) data. These datasets are useful for providing information on land cover, vegetation conditions, water bodies, flood inundations, burned areas, urban development, and surface temperature. Unlike optical sensors, Sentinel-1 can acquire images under cloud cover and at night, making it particularly useful during severe weather events when optical observations are not available. The meteorological data are provided by NOAA, ERA5, the India Meteorological Department (IMD) and NASA POWER. The dataset includes rainfall, relative humidity, air temperature, atmospheric pressure, wind velocity and direction, soil moisture, and incoming solar radiation. These variables describe the atmospheric conditions that cause floods, cyclones, droughts, heatwaves, and other extremes. Historical weather observations also provide useful temporal patterns that enhance the model’s ability to identify conditions prior to disaster occurrence. The acquired datasets are organized in a unified spatiotemporal and multimodal source where information from different modalities is fused for further analysis, as illustrated in Figure 1.

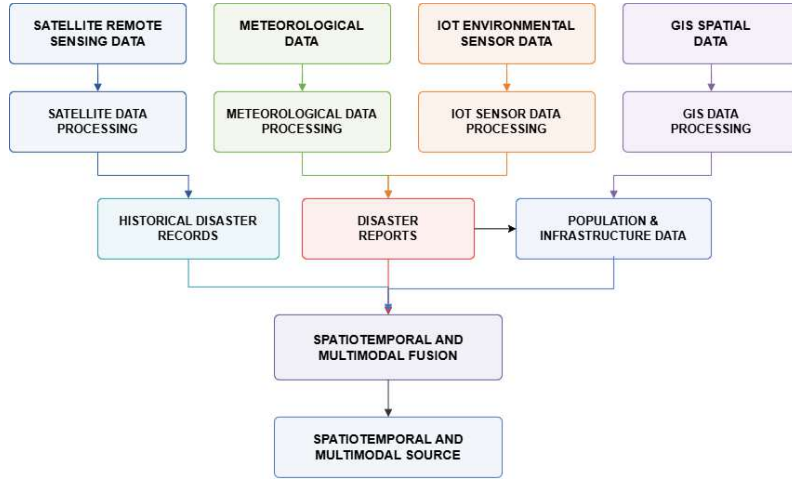


Fig. 1 Data acquisition framework

As depicted in Fig.1, the RADIANT-Net framework uses large-scale observations, and the framework includes multimodal sources such as satellites, metrological data, and Internet of Things (IoT) monitoring systems. Local environmental conditions are continuously monitored by river gage stations, rainfall gages, seismic sensors, soil moisture sensors, wildfire temperature sensors, and air-quality monitoring stations.

These real-time observations improve situational awareness by detecting fast changes that are not immediately visible in satellite images, thus improving the early warning capability of the proposed framework. Geographical characteristics are represented by GIS spatial datasets (OpenStreetMap, SRTM and HydroSHEDS). datasets consist of digital elevation models, land-use maps, slopes, aspects, river networks, drainage systems, transportation networks, and building footprints. This type of information allows us to understand the morphology of the terrain , the hydrological connectivity, and the distribution of the infrastructure, all of which are very important for disaster susceptibility and emergency access. Historical disaster information is collected from internationally recognized repositories, including EM-DAT, DesInventar, the USGS and NASA FIRMS. These databases contain past disaster event records, including location, date of occurrence, disaster category, number of casualties, economic losses, and infrastructure damage. Historical observations are used as reference data in the development of the models so that the learning algorithm can learn the correlation between environmental conditions and recorded disaster events. The proposed framework uses not only numerical and spatial data but also textual data related to disasters collected from ReliefWeb, official government advisories, disaster management agencies, authentic news agencies, and emergency alerts. These reports provide contextual descriptions of disaster progression, evacuation activities, rescue operations, and damage assessments. Adding textual information allows us to enrich the dataset with semantic knowledge, which is often missing from satellite imagery or environmental measurements alone.

Table 2: Multimodal data sources integrated into the proposed RADIANT-Net framework

Data Source	Used in Study	Purpose in RADIANT-Net
Sentinel-1	SAR/VV/VH/flood or land-surface information	Product level, polarization, acquisition period, processing steps
Sentinel-2	Optical/NDVI/NDWI	Product level, bands/indices, acquisition period, preprocessing
NASA GPM	Precipitation	Version, Early/Late/Final run, temporal resolution, period
IMERG		
NOAA meteorological data	Temperature, rainfall, wind, etc.	Exact NOAA dataset/product, variables, period
NASA FIRMS	Active fire/wildfire information	Product, date range, spatial/temporal resolution
OpenStreetMap	Roads, buildings, infrastructure/exposure	Data extraction date and OSM attribution/license
WorldPop	Population/exposure information	Dataset/version, year and spatial resolution
CrisisNLP/crisis reports	Disaster-event/textual information	Exact dataset, source, collection period and license

Sentinel-1 and Sentinel-2 data are made available systematically and free of charge through the Copernicus ecosystem. NASA's IMERG provides precipitation estimates and multiple processing runs and formats. NOAA provides data-access services for its datasets. Multimodal data sources for the assessment of the impact of disasters on society As illustrated in Table 2, population and infrastructure data are collected from WorldPop and OpenStreetMap. These data provide information on population distribution, hospitals, schools, emergency shelters, transportation networks, power plants, and communication infrastructure. This information allows the framework to estimate human exposure, identify vulnerable areas, and assist in emergency planning by providing information on critical facilities that need to be protected as a priority in response to a disaster. The seven data sources provide complementary information on the physical environment, atmospheric conditions, geographical characteristics, historical disaster behavior and societal exposure. By integrating these heterogeneous datasets into a unified spatiotemporal and multimodal source, RADIANT-Net can capture complex relationships between multiple disaster influencing factors. This multimodal representation provides a solid basis for accurate disaster prediction, spatial risk assessment, severity estimation, and intelligent decision support.

2.2 Data Preprocessing

The heterogeneous datasets incorporated in the proposed study are from different sources with different spatial resolutions, temporal frequencies, coordinate systems, data formats, and measurement characteristics. These inconsistencies must be addressed to combine the datasets into a common learning framework. Therefore, we design an integrated preprocessing framework designed to improve the consistency of the data while preserving the information in each modality. Preprocessing includes data cleaning, spatial alignment, temporal synchronization, feature normalization, multimodal integration, and event-based training sample generation.

2.2.1 Data Cleaning and Quality Control

First, the reliability of the collected data set should be improved. Duplicate observations, bad records, missing entries, and measurements that fail quality control are flagged and removed. This process reduces the impact of unreliable observations and helps to ensure consistency across all data sources. The cleaned data set is written as

$$D_{clean} = D - (D_{noise} \cup D_{missing} \cup D_{duplicate}) \quad (1)$$

The collected multimodal datasets are filtered using Equation (1) on uncertain observations D_{noised} denotes refers to noisy observations, $D_{missing}$ corresponds to incomplete records, and $D_{duplicate}$ represents repeated samples. The resulting dataset D_{clean} contains only observations that meet the predefined quality requirements.

2.2.2 Spatial Preprocessing

As satellite imagery, GIS layers, and socioeconomic datasets are obtained from different platforms, they initially do not have the same spatial reference. All spatial datasets are thus projected to coordinate the reference system and resampled to a common

spatial resolution. The next step is to extract the study area from each data set so that all observations refer to the same geographical region.

The spatial transformation can be written as

$$S^* = T(S) \quad (2)$$

Eq. (2) aligns satellite imagery, GIS layers, infrastructure maps and population datasets to a common spatial reference, enabling consistent spatial analysis, where S is the original spatial dataset, $T(\cdot)$ is the spatial transformation function and S^* is the aligned dataset.

2.2.3 Temporal Synchronization

The temporal characteristics of the collected datasets vary considerably. Satellite observations are taken from time to time, meteorological measurements at fixed intervals, IoT sensors create observations in a continuous fashion, and textual reports are created as major events occur. To establish temporal consistency, all the dynamic datasets are synchronized within a common observation window.

Synchronized multimodal observation at the time step t is expressed as

$$X_t = \{X_{sat}, X_{met}, X_{iot}, X_{txt}\}_t \quad (3)$$

Equation (3) synchronizes all observations in the same temporal window before multimodal learning, where X_{sat} , X_{met} , X_{iot} , and X_{txt} denote the meteorological, IoT, textual and satellite observations associated with the same time interval.

2.2.4 Feature normalization and encoding

The numerical ranges of the collected variables differ substantially. To prevent high-magnitude variables with larger magnitudes from dominating the learning process, continuous features are normalized before model training. Min-max normalization is adopted to map all numerical variables to a common interval.

$$x^* = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (4)$$

Equation (4) scales heterogeneous features into a common numerical range, improving optimization stability during model training, where x represents the original value of the characteristic, $x_{\min}$ and $x_{\max}$ are the minimum and maximum values of that characteristic, respectively, and x^* corresponds to the normalized value.

2.2.5 Integration of multimodal data

Once individual datasets have been preprocessed, they are combined into a unified representation. Each sample contains synchronized information describing the same geographical location and observation period. This integration preserves the complementary characteristics of each modality while enabling joint learning.

The integrated feature representation is defined as

$$X_{fusion} = [X_{sat}, X_{met}, X_{iot}, X_{gis}, X_{hist}, X_{txt}, X_{soc}] \quad (5)$$

Equation (5) integrates heterogeneous modalities into a unified spatiotemporal and multimodal source, preserving complementary spatial, temporal, semantic and socioeconomic information, with X_{sat} , X_{met} , X_{iot} , X_{gis} , X_{hist} , X_{txt} , and X_{soc} corresponding to the satellite, meteorological, IoT, GIS, historical disaster, textual and socioeconomic features, respectively.

2.2.6 Training Sample Generation

The last preprocessing step generates the samples for supervised learning by associating each integrated multimodal observation with the corresponding disaster label. Each sample consists of the prediction target and the environmental conditions observed before or during a disaster event.

The training dataset is then represented as

$$D = \{(X_i, Y_i)\}_{i=1}^N \quad (6)$$

The supervised dataset for training, testing and validating the proposed model is defined by Eq. (6). X_i : Integrated multimodal feature vector. Y_i : Corresponding ground-truth label. N : Total number of samples. The dataset is then split into training, test and validation sets to facilitate model development and objective performance evaluation.

2.3 RADIANT-Net architecture

In this paper, we propose RADIANT-Net (Resilient Adaptive Disaster Intelligence with Agentic Neural Twins Network), a multimodal spatiotemporal framework for early natural-disaster prediction and emergency decision support. It consists of five connected components, as shown in Figure 2, a multimodal foundation encoder, a dynamic graph intelligence module, a neural digital twin, multitask prediction heads and an agentic decision intelligence module. Together, these components estimate the occurrence of the disaster, the type of disaster, the spatial risk, severity, and the appropriate response actions over future prediction horizons.

As shown in Fig. 2, the integration of the RADIANT-Net framework architecture includes disaster event records collected from publicly available disaster repositories, including GDACS, USGS, and NOAA, and are subsequently integrated with complementary environmental information obtained from Open-Meteo, NASA GPM IMERG, NOAA meteorological observations, and data acquired from Sentinel-1 and Sentinel-2 satellite imagery. These heterogeneous data sources are combined to construct a balanced multimodal data set covering five disaster categories. The multimodal foundation encoder extracts and fuses heterogeneous features. The dynamic graph intelligence module further empowers the fused heterogeneous features to capture spatial dependencies. The neural digital twin module simulates spatiotemporal evolution for early

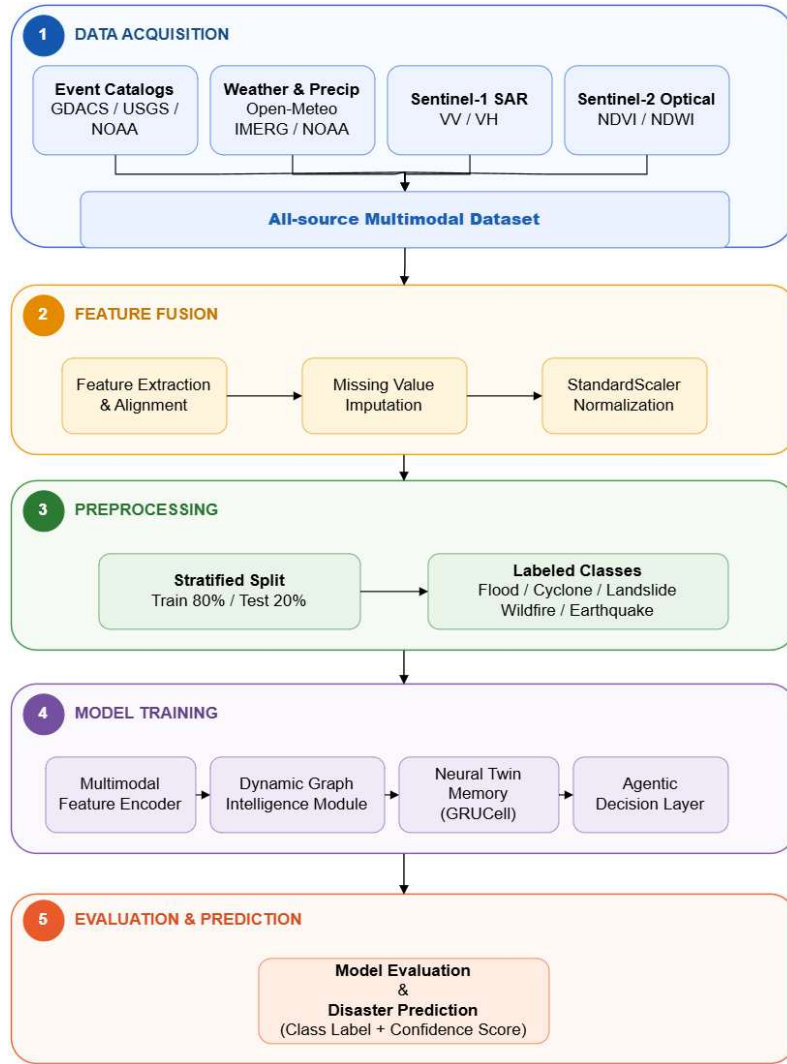


Fig. 2 Overall End-to-end architecture of the proposed model

disaster prediction, and the agentic decision intelligence module generates actionable responses. The system produces outputs, including disaster occurrence prediction, disaster type classification, spatial risk mapping, severity estimation.

2.3.1 Multimodal input representation

Satellite imagery, precipitation, weather, active fires, infrastructure, population, and crisis-reported data are assembled for each geographical region and prediction time. The heterogeneous inputs are quality controlled, spatially aligned, and temporally synchronized within a common observation window. The resulting sample provides a

unified multimodal input to the foundation encoder for subsequent feature extraction and fusion.

For the geographical region i at the warning time t , the multimodal input is

$$X_{i,t} = \{X_{i,t_s}^{S1}, X_{i,t_o}^{S2}, X_{i,t-L:t}^{GPM}, X_{i,t-L:t}^{NOAA}, X_{i,t-L:t}^{FIRMS}, X_i^{OSM}, X_i^{WP}, X_{i,t-L:t}^{TXT}\} \quad (7)$$

Equation (7) is used to construct the complete multimodal observation associated with a specific geographical region and prediction time. where X^{S1} denotes Sentinel-1 and X^{S2} denotes Sentinel-2 observations; X^{GPM} represents GPM IMERG precipitation; X^{NOAA} contains meteorological variables; X^{FIRMS} contains active-fire observations; X^{OSM} describes roads and infrastructure; X^{WP} represents the population distribution; and X^{TXT} contains crisis-related textual reports. The interval $t - L : t$ specifies the historical lookback window. Only observations available at or before t are included, preventing future information from entering the early warning model. Historical disaster inventories are used primarily as reference labels for training and evaluation.

2.3.2 Multimodal foundation encoder

The multimodal foundation encoder converts the heterogeneous inputs into a common latent representation while preserving the distinctive information contained in each modality. SAR imagery, optical imagery, weather sequences, active-fire observations, infrastructure–population data, and crisis reports are processed through modality-specific encoders. The resulting representations are transformed into projected feature tokens and integrated through cross-modal attention. The fused output forms a unified multimodal embedding for subsequent spatial reasoning and disaster forecasting in RADIANT-Net. Each data modality is processed via a suitable encoder. Sentinel-1 SAR observations and Sentinel-2 multispectral imagery are processed by separate image encoders because radar backscatter and optical reflectance have different physical characteristics. Radar and optical images are processed via CNN- or Vision Transformer-based encoders, whereas precipitation and meteorological sequences are represented via a temporal transformer, temporal convolutional network, or recurrent unit. Fire detections, infrastructure attributes, population data, and crisis reports are encoded via modality-specific neural layers.

For modality m , the encoded representation is defined as

$$h_{i,t}^{(m)} = P_m f_m \left(X_{i,t}^{(m)} \right), \quad \mathbf{h}_{i,t}^{(m)} \in \mathbb{R}^d \quad (8)$$

Equation (8) is used to transform every input modality into a common d -dimensional representation space. In Eq. (8), $f_m(\cdot)$ is the encoder assigned to modality m , whereas P_m is a learnable projection layer. The common dimensionality allows heterogeneous image, time series, vector, demographic, and textual representations to be combined within the same latent space.

The projected features are integrated through availability-aware cross-modal attention:

$$\alpha_{i,t}^{(m)} = \frac{q_{i,t}^{(m)} \exp(\mathbf{w}^T \mathbf{h}_{i,t}^{(m)})}{\sum_{r=1}^M q_{i,t}^{(r)} \exp(\mathbf{w}^T \mathbf{h}_{i,t}^{(r)}) + \varepsilon}, \quad \mathbf{h}_{i,t}^F = \sum_{m=1}^M \alpha_{i,t}^{(m)} \mathbf{h}_{i,t}^{(m)} \quad (9)$$

Equation (9) is employed to integrate the available modalities into a unified multimodal representation. Here, M is the number of modalities, $\alpha_{i,t}^{(m)}$ is the weight attention, and $\mathbf{w}$ is a learnable vector. The indicator $q_{i,t}^{(m)} \in \{0, 1\}$ identifies whether modality m is available and satisfies the quality requirements. The small constant ε prevents division by zero. This formulation allows RADIANT-Net to operate when one or more data sources are missing.

2.3.3 Dynamic graph intelligence module

Natural hazards are not treated as independent events within isolated grid cells. Instead, the study region is defined as a dynamic graph:

$$G_t = (V, E_t), \mathbf{h}_{i,t}^{G,(0)} = \mathbf{h}_{i,t}^F \quad (10)$$

Equation (10) is used to represent geographical regions as graph nodes, with their evolving spatial relationships as dynamic constructed edges. The node set V contains geographical regions or grid cells, whereas E_t contains their relationships at time t . This module represents geographical regions as graph nodes, while their spatial, environmental, and infrastructural relationships as dynamic constructed edges. Graph attention layers capture dependencies between neighboring regions and learn their relative influence on disaster propagation. This produces a spatially enhanced representation that improves the modeling of disaster spread and regional interactions. An edge may represent geographical proximity, hydrological connectivity, road connectivity, weather similarity, wind-driven interaction, or population movement. The fused multimodal representation initializes the feature vector of each node.

The contribution of adjacent node j to target node i is calculated via graph attention:

$$\beta_{ij,t}^{(l)} = \text{softmax}_{j \in \mathcal{N}(i)} \left[\text{LeakyReLU} \left(\mathbf{a}_l^T \left[W_l \mathbf{h}_{i,t}^{G,(l)} \parallel W_l \mathbf{h}_{j,t}^{G,(l)} \parallel \mathbf{e}_{ij,t} \right] \right) \right] \quad (11)$$

Equation (11) is used to quantify the relative influence of neighboring geographical regions on the target region. In Eq. (11), $\mathcal{N}(i)$ is the neighboring nodes associated with node i , $\mathbf{e}_{ij,t}$ contains edge attributes, W_l is the transformation mapping, $\mathbf{a}_l$ is the attention vector and $\parallel$ denotes concatenation.

The graph-enhanced node representation is updated as

$$\mathbf{h}_{i,t}^{G,(l+1)} = \phi \left(\sum_{j \in \mathcal{N}(i) \cup \{i\}} \beta_{ij,t}^{(l)} W_l \mathbf{h}_{j,t}^{G,(l)} \right) \quad (12)$$

Equation (12) is used to aggregate information from influential neighboring regions and to generate a spatially improved representation. The nonlinear function $\phi(\cdot)$ can be implemented via ReLU, GELU, or another activation function. Equations (11) and (12) allow the model to learn hazard-specific propagation patterns. For example, upstream rainfall and drainage connections can dominate flood prediction, while wind direction and vegetation continuity may be more important to the progression of wildfires.

2.3.4 Neural digital twin module

The graph-enhanced representation is passed to the neural digital twin, which maintains a latent approximation of the evolving disaster environment. Its initial state is obtained from the graph-enhanced features:

$$\mathbf{z}_{i,t} = \psi(\mathbf{h}_{i,t}^G) \quad (13)$$

Equation (13) is used to initialize the virtual disaster state of each geographical region. The mapping function $\psi(\cdot)$ converts the graph representation into the latent state used for future simulation. As new observations become available, this state is updated to reflect the evolving physical environment. The neural digital twin module initializes a latent disaster state from graph-enhanced representations and models its spatiotemporal evolution over multiple forecast horizons. The latent state transition function iteratively predicts future states by incorporating temporal dynamics and spatial interactions between neighboring regions. This enables simulation of the progression of disasters and early prediction of future risk conditions across regions.

Future disaster states are recursively estimated as

$$\hat{\mathbf{z}}_{i,t+k} = F_\theta \left(\hat{\mathbf{z}}_{i,t+k-1}, \mathbf{u}_{i,t+k|t}, \sum_{j \in \mathcal{N}(i)} A_{ij,t+k|t} \hat{\mathbf{z}}_{j,t+k-1} \right), \quad k = 1, \dots, K \quad (14)$$

Equation (14) is used to forecast the progression of the disaster environment on multiple future horizons. where F_θ is a learnable transition model, $\mathbf{u}_{i,t+k|t}$ represents the external forcing available at warning time t , and $A_{ij,t+k|t}$ represents the predicted interaction between neighboring regions. External forcings may include rainfall, wind, pressure, temperature, or other forecast variables. The resulting sequence of latent states models how risk may increase, decrease, or spread spatially.

2.3.5 Multitask prediction heads

The multitask module generates multiple outputs from shared representations, including probability of disaster occurrence, classification of disaster types, spatial risk mapping, and severity estimation. Each prediction head is optimized jointly to capture complementary aspects of disaster forecasting and improve overall model performance.

The digital-twin states are passed to task-specific output heads. The probability of disaster-occurrence and the probability of disaster-type are calculated as

$$\hat{p}_{t+k} = \sigma(W_o \vec{z}_{t+k} + b_o), \hat{c}_{t+k} = \text{softmax}(W_c \vec{z}_{t+k} + b_c) \quad (15)$$

Equation (15) is used to predict whether a disaster will occur and to identify its most likely category. The vector $\vec{z}_{t+k}$ is the pooled digital-twin representation. The scalar $\hat{p}_{t+k} \in [0, 1]$ represents the probability of occurrence, while $\hat{c}_{t+k}$ contains probabilities for disaster classes such as floods, cyclones, wildfires, or landslides.

Spatial risk and event severity are estimated via

$$\hat{R}_{i,t+k} = \sigma[g_R(\hat{z}_{i,t+k})], \hat{s}_{t+k} = g_S(\vec{z}_{t+k}) \quad (16)$$

Equation (16) is used to produce location-specific risk values and estimate the expected severity of the predicted event. The function $g_R(\cdot)$ generates grid or region risk, while $g_S(\cdot)$ estimates severity. Depending on the disaster category, severity can represent the extent of the flood, the intensity of the fire, the exposed population, the impact of the infrastructure or the level of ordinary damage.

2.3.6 Agentic decision intelligence module

The agentic decision intelligence module converts predictive outputs into structured emergency recommendations. The prediction outputs are combined with population, infrastructure, road network, shelter, and resource information. Situation assessment, resource allocation, evacuation planning, and communication agents evaluate candidate emergency responses. The module performs multiagent reasoning, decision generation, and actionable decision output for disaster response. Multiple cooperative agents, including risk assessment agents, resource planning agents, response strategies, and communication agents, interact through a shared context to analyze the predicted disaster conditions. The decision generation stage ranks candidate actions, selects optimal strategies, and provides justification with confidence scores.

The recommended action is selected as

$$a_t^* = \operatorname{argmax}_{a \in A(C_t)} \left[U(a | \hat{\mathcal{Y}}_{t+1:t+K}) - \lambda C(a) \right] \quad (17)$$

Equation (17) is used to select an emergency action that maximizes the expected benefit while considering the operational cost and feasibility. Here, $\hat{\mathcal{Y}}_{t+1:t+K}$ contains the predicted events, type, risk, and severity. The set $A(C_t)$ contains actions that satisfy current constraints, where $U(\cdot)$ denotes the utility for emergency-management, $C(a)$ is the cost of action, and λ balances utility and cost. The generated actions remain recommendations subject to validation by authorized emergency personnel.

2.4 Model optimization

RADIANT-Net is trained through a multitasking joint objective that combines losses associated with disaster occurrence, disaster category, spatial risk, severity, and future-state simulation. The architecture is trained via a joint multitask objective:

$$\mathcal{L}_{total} = m_o\lambda_o\mathcal{L}_{occ} + m_c\lambda_c\mathcal{L}_{type} + m_r\lambda_r\mathcal{L}_{risk} + m_s\lambda_s\mathcal{L}_{sev} + m_d\lambda_d\mathcal{L}_{twin} \quad (18)$$

Equation (18) is used to jointly optimize all the prediction and the digital-twin simulation tasks. The terms represent losses for disaster occurrence, disaster type, spatial risk, severity, and future-state simulation. The coefficients $\lambda_o, \lambda_c, \lambda_r, \lambda_s, \lambda_d$ control their relative importance. Indicators $m_o, m_c, m_r, m_s, m_d \in \{0, 1\}$ exclude a task loss when the corresponding ground-truth label is not available.

2.5 Early warning generation

At operational time t , the latest available observations are processed, and forecasts are generated for every future horizon k . A location-specific warning is issued when the estimated probability of occurrence and spatial risk satisfy validation-derived thresholds:

$$A_{i,t+k} = \mathbb{I} \left[\hat{p}_{t+k} \geq \tau_k \wedge \hat{R}_{i,t+k} \geq \rho_k \right] \quad (19)$$

Equation (19) is used to determine whether an early warning should be issued for a particular region and forecast horizon. Here, $\mathbb{I}$ is the indicator function, τ_k is the probability threshold of occurrence, and ρ_k is the spatial risk threshold. These values should be selected via validation data to achieve an acceptable balance between missed events and false alarms.

The warning lead time is calculated as

$$LT = t_{onset} - t_{alert} \quad (20)$$

Equation (20) is used to measure how far RADIANT-Net generates an alert before the documented onset of the disaster in advance. In Eq. (20), t_{alert} denotes the initial warning time, and t_{onset} indicates the recorded onset time. The lead time should be reported together with precision, recall, false-alarm rate, and spatial-risk accuracy because classification performance alone does not indicate whether a warning is operationally timely.

3 Experimental Setup

The experimental evaluation of the proposed RADIANT-Net framework was carried out on a workstation running Microsoft Windows 11 Home Single Language, Build 10.0.26200. The system was equipped with a Qualcomm Snapdragon X (X126100) Oryon processor, 15.6 GB of RAM, and an integrated Qualcomm Adreno X1-45 GPU. All experiments were executed on a single CUDA-enabled GPU (Google Colab runtime). The framework was implemented in Python 3.12.0 within a dedicated environment to

ensure consistency and reproducibility of the software. The deep learning model was developed via PyTorch. Data preparation and numerical processing were performed with NumPy (v2.5.1) and Pandas (v3.0.3), while Scikit-learn (v1.9.0) was used for the partition of the data set and the calculation of evaluation metrics, including precision, recall, precision, F1 score, confusion matrix, and ROC-AUC. All performance graphs and visual analyzes were performed using Matplotlib 3.11.0.

The experimental data set was constructed from eight complementary sources representing diverse environmental and social information: Sentinel-1 synthetic aperture radar data, Sentinel-2 multispectral observations, NASA GPM IMERG precipitation products, NOAA meteorological records, NASA FIRMS active-fire detections, OpenStreetMap geospatial and infrastructure features, WorldPop population-density estimates, and disaster-related textual reports obtained from CrisisNLP. The multimodal remote sensing data recovered via the planetary computer were processed with Rasterio to extract features. The collected data were preprocessed via normalization, resizing, augmentation, feature extraction, etc., before training. The dataset was partitioned into training, test, and validation sets, and the models' performance was evaluated on the independent test set via standard classification metrics. To allow for fair comparison, reproducibility and consistency of the reported results, the same hardware, software environment, and experimental settings were used throughout the entire study. A summary of the experimental configuration is provided in Table 3.

Table 3: Table 3 Computational environment, hardware - software resources used for the Proposed implementation.

Component	Specification
Operating System	Microsoft Windows 11 Home Single Language (Build 10.0.26200)
Processor	Qualcomm Snapdragon X (X126100) Oryon CPU
RAM	15.6 GB
GPU	Qualcomm Adreno X1-45 (Integrated)
Python	3.12.0
Deep Learning Framework	PyTorch 2.13.0
Data Processing	NumPy 2.5.1, Pandas 3.0.3
Machine Learning	Scikit-learn 1.9.0
Visualization	Matplotlib 3.11.0
Geospatial Processing	Rasterio
Satellite Data Access	pystac-client, Planetary Computer SDK

3.1 Algorithm:

Algorithm: Summarizes complete training work flow

Input: Multimodal observations X , modality masks q ,
spatial graph G , labels Y , history length L ,
forecast horizon K , epochs E , learning rate η ,
alert thresholds τ and ρ

Output:

Trained parameters Θ^* , occurrence probability p ,
disaster type c , risk map R , severity s ,
warning A , lead time LT , and optimal action a^*

```
1: Clean, normalize, and temporally align  $X$ 
2: Initialize modality encoders, attention-fusion module,
   graph network, digital twin, and prediction heads

3: for epoch = 1 to  $E$  do
4:   Encode each available modality
5:   Fuse encoded features using mask-aware attention
6:   Construct  $G$  and learn spatial dependencies
7:   Initialize the digital-twin state
8:   for  $k = 1$  to  $K$  do
9:     Predict the future regional state
10:    Estimate  $p$ ,  $c$ ,  $R$ , and  $s$ 
11:   end for
12:   Compute the joint multi-task loss
13:   Update model parameters using Adam
14: end for

15:  $\Theta^* \leftarrow$  parameters with the best validation performance
16: Generate forecasts from the latest observations

17: for  $k = 1$  to  $K$  do
18:    $A[i, t+k] \leftarrow 1$  if  $p[t+k] \geq \tau[k]$ 
   and  $R[i, t+k] \geq \rho[k]$ ; otherwise,  $A[i, t+k] \leftarrow 0$ 
19: end for

20:  $LT \leftarrow$  predicted disaster-onset time - alert time
21: Rank feasible response actions using predicted risk,
   exposure, resources, operational cost, and constraints
22:  $a^* \leftarrow$  highest-utility feasible action

23: return  $\Theta^*$ ,  $p$ ,  $c$ ,  $R$ ,  $s$ ,  $A$ ,  $LT$ , and  $a^*$ 
```

3.2 Hyperparameter settings

The RADIANT-Net model was trained on an all-source multimodal disaster dataset comprising 500 balanced samples across five classes. Nineteen multimodal features were standardized via StandardScaler. A stratified training–test split of 80:20 was applied, producing 400 training samples and 100 testing samples, with 42 fixed random seeds for reproducibility. The network configured a hidden representation of 64 and a dropout probability of 0.20. The multimodal encoder applies fully connected layers with batch normalization and ReLU activation. The neural-twin module updates an internal state through a GRUCell, and the agentic decision layer maps the refined representation to five disaster classes. Training was performed for 30 epochs via the Adam optimization with a learning rate of 0.001 and a mini-batches of 64 samples. While cross-entropy loss was employed as the informative function. All the experiments were executed on a CPU. These hyperparameter settings were kept constant during training and evaluation to ensure fair and reproducible results, as listed in Table 4.

Table 4: Table 4 Hyperparameter configuration used for training the proposed model

Hyperparameter	Value
Dataset	All-source multimodal (500 samples)
Input features	19
Number of classes	5
Train–test split	80%/20% (stratified)
Training samples	400
Testing samples	100
Feature scaling	StandardScaler
Random seed	42
Architecture	RADIANT-Net (Encoder + Neural Twin + Decision)
Hidden dimension	64
Dropout rate	0.2
Normalization	BatchNorm1d(encoder)
Activation	ReLU
Neural-twin cell	GRUCell
Optimizer	Adam
Learning Rate	0.001

Hyperparameter	Value
Batch Size	64
Number of Epoch	30
Loss function	CrossEntropyLoss
Train Shuffle	TRUE
Device	CPU

3.3 Distribution of data sets

The proposed RADIANT-Net framework was evaluated using a multimodal disaster data set developed for the classification of five-class natural disasters . The dataset consists of 500 georeferenced disaster events that are evenly distributed among five disaster categories, namely, flood, wildfire, landslide, cyclone, and earthquake, with 100 samples allocated to each class to maintain a balanced class distribution. Rather than relying on a single data modality, each disaster event is represented by a comprehensive multimodal feature vector that integrates environmental, meteorological, geospatial, and remote sensing information. The input features include rainfall, temperature, humidity, wind speed, atmospheric pressure, river level, soil moisture, slope of the terrain, vegetation dryness, seismic activity, ground displacement, satellite-derived water index, population density, accessibility of roads and historical risk score. To enhance the representation of disaster conditions, the dataset was further enriched with Sentinel-1 synthetic aperture radar (SAR) observations, backscatter coefficients (VV and VH), and Sentinel-2 multispectral features, incorporating the normalized difference vegetation index (NDVI) and the normalized difference water index (NDWI). Meteorological information, including precipitation and weather variables, was obtained from NASA GPM IMERG, NOAA, and Open-Meteorology. After preprocessing and integration of features, the final dataset comprised 19 input features for subsequent model development and performance assessment.

For experimental evaluation, the data set was partitioned into 80:20 stratified training and test split, ensuring that each disaster category retained the same class proportion in both subsets. This resulted in 400 training samples and 100 testing samples, corresponding to 80 training instances and 20 testing instances per disaster class. In this study, no separate validation set was considered; therefore, model performance was assessed through the independent test set. Before training, all input features were standardized using the standard scaler technique and a constant random seed of 42 was used to ensure the reproducibility of the experimental analysis. Fig. 3 presents the distribution of the samples across the data set.

4 Results

4.1 Training Performance

The heterogeneous data sources are combined to build a balanced multimodal data set, with 500 disaster events and 19 input features that cover five disaster categories.

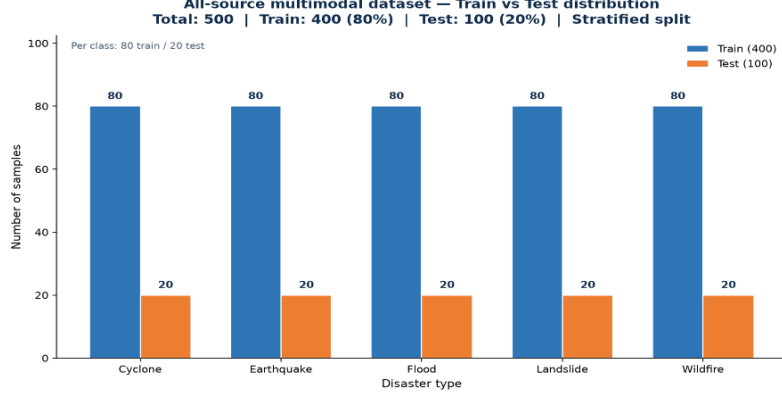


Fig. 3 Distribution of training and testing samples across all the disaster classes

Data preprocessing, which includes multimodal information alignment, null value filling with suitable imputation techniques, and feature value normalization with the Standard-Scaler method, is performed. The data set was partitioned using an 80:20 training-to-testing to preserve the distribution of the classes being maintained during model development. The training split has 400 samples, and the test split has 100 samples. In the next stage, the processed features are passed to the proposed RADIANT-Net architecture. The framework consists of a multimodal feature encoder, a GRU cell-based neural twin memory module for adaptive feature representation, and an agentic decision layer that responds to the final classification of the disaster. The model is optimized with the Adam with a learning rate of 0.001, a mini-batch size of 64, a dropout rate of 0.20 and a cross-entropy loss function for 30 training epochs. The trained data set is then tested on the independent test data set with multiple classification measures, including accuracy, precision, recall, and F1 score together with the analysis of the confusion matrix. The framework also enables confidence scores for each predicted disaster category, which helps to make reliable decisions. The experimental results under the experimental configuration adopted show that the overall classification accuracy of the proposed RADIANT-Net is 92%, and it can effectively learn multimodal disaster characteristics and make accurate predictions of the type of disaster.

4.2 Training precision

RADIANT-Net achieves stable model convergence and higher training accuracy after 30 epochs. The initial accuracy of the training was 49.00% and that of the seventh epoch was 87.25%, which reflected the ability to learn features early. After improvement, the training accuracy converged to 92.00%. The final training accuracy of the model was 92.00%, which is very close to the accuracy of the independent test data set and indicates that a high generalizability and a slight degree of overfitting are observed in the model. The results presented in Fig. 4 indicate that the multimodal learning paradigm is effective in capturing complex associations in heterogeneous disaster information.

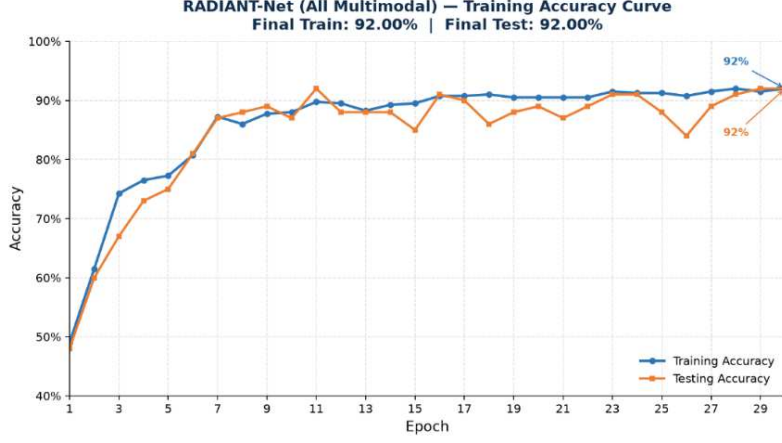


Fig. 4 Evolution of training accuracy across 30 training epochs

4.3 Validation accuracy

The proposed RADIANT-Net was evaluated on an independent validation set (20% validation set, 100 samples) after each training epoch. Effective early training increases validation accuracy from 48.00% in the first epoch to 89.00% in the ninth epoch. The highest validation accuracy is 92.00% at Epoch 11. The model reached 92.00% validation accuracy and stabilized after 30 training epochs, as shown in Fig. 5, with minor fluctuations. The peak and final validation accuracies were similar, indicating stable convergence and good generalizability for multimodal disaster classification.

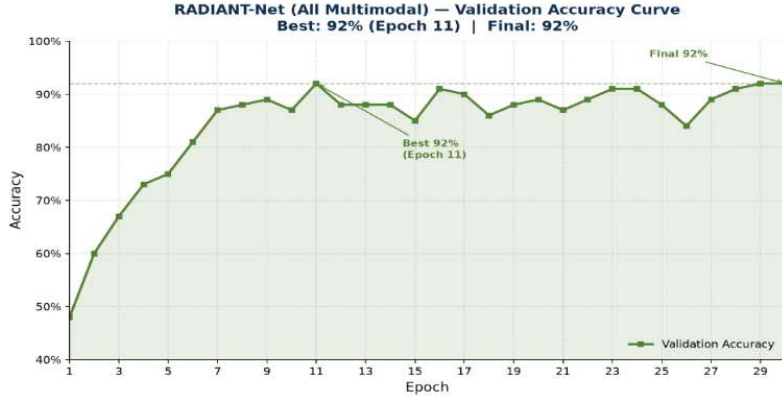


Fig. 5 Validation accuracy progression over 30 training epochs

Table 5 summarizes the training and validation performance in 30 training epochs

Table 5: Table 5 validation accuracy summery across individual training epochs

Epoch	Val Acc	Epoch	Val Acc	Epoch	Val Acc
1	0.4800	11	0.9200	21	0.8700
2	0.6000	12	0.8800	22	0.8900
3	0.6700	13	0.8800	23	0.9100
4	0.7300	14	0.8800	24	0.9100
5	0.7500	15	0.8500	25	0.8800
6	0.8100	16	0.9100	26	0.8400
7	0.8700	17	0.9000	27	0.8900
8	0.8800	18	0.8600	28	0.9100
9	0.8900	19	0.8800	29	0.9200
10	0.8700	20	0.8900	30	0.9200

The validation accuracy shifted from **48%** to **89%** during the first nine epochs, indicating effective learning and fast convergence in the early stages of training. Small fluctuations were observed between 12% and 26%, and the model achieved a final validation accuracy of **92%**.

4.4 Training Loss Progression



Fig. 6 Training loss progression over 30 training epochs

Fig. 6 presents the training loss of the proposed RADIANT-Net trended downward throughout the 30 training epochs, which indicates stable convergence in model optimization. The first loss is 1.5844, which is expected for a randomly initialized model. In the early training phase, a significant reduction in loss was achieved, with

a value of approximately 0.48 at epoch 8, indicating effective learning of multimodal disaster features. The loss still decreased even after the first epoch of learning, but at a slower pace, and after approximately epoch 15, it seemed to have stabilized with minor fluctuations. The training loss at the last epoch (epoch 30) is 0.2612. The consistent reduction in training loss and the 92% training and validation accuracies prove that the proposed framework converges stably and is successfully optimized in the absence of any notable training instability.

4.5 Validation Loss

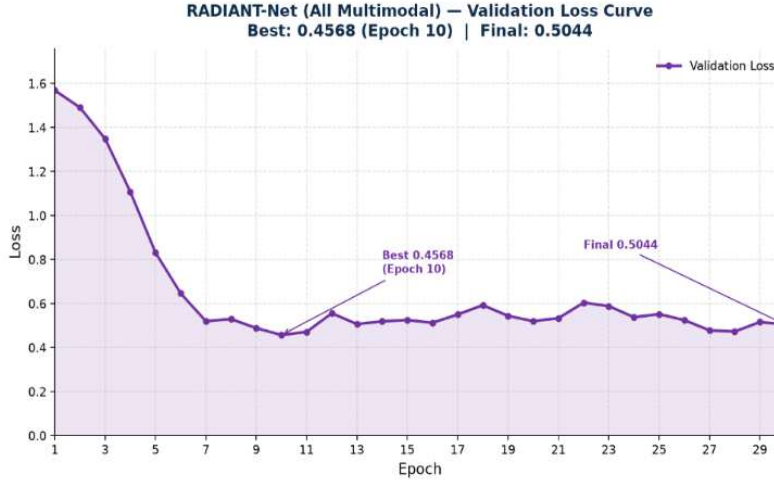


Fig. 7 Validation loss progression over 30 training epochs

The observed validation loss of the proposed RADIANT-Net always decreased during the training process, as shown in Fig. 7, indicating effective learning and good generalizability. In the first epoch, we observed a validation loss of 1.5685, which is expected from an untrained model with limited predictive power. We observed a rapid decrease in the early training phase, reaching 0.5196 at epoch 7 and the lowest value of 0.4568 at epoch 10, indicating that the learned feature representations generalized well to unseen multimodal disaster samples. After the 10th epoch, the validation loss fluctuated slightly between 0.47 and 0.60 and was relatively stable for the remaining training epochs. The final validation loss achieved is 0.5044 at epoch 30, which is consistent with the validate accuracy achieved 92%. Stable validation loss and corresponding training performance demonstrate that the proposed framework converged successfully and maintained good generalizability without significant overfitting.

4.6 Testing performance

According to Table X, the proposed RADIANT-Net classifies 100 multimodal disaster datasets from an independent test set classwise. The macro-averaged precision, recall, and F1 score values of 0.9268, 0.9200, and 0.9196, respectively, indicate the same

performance in all types of disasters. The model classification accuracy was 92.00%. Earthquakes classified all earthquakes with a precision of 1.00 accompanied by recall and an F1 score. The F1 scores for Landslide and Wildfire were 0.9744 and 0.9268, respectively, suggesting a good prediction. Floods had good recall (0.95) but low precision (0.7917), indicating that nonflood events were labeled floods. Several cyclone events were misclassified due to high precision (0.9375) but low recall (0.75). Given their comparable environmental and meteorological properties, floods and cyclones may still be confused. The proposed RADIANT-Net correctly and reliably classifies all five disaster categories. The strong performance of the framework for earthquakes, landslides, and wildfires and its high general precision of 92.00% validate its efficacy but also reveal areas for development in flood and cyclone event discrimination.

4.7 Classification Report

Table 6: Table 6 Classwise classification performance of the proposed method for the independent test dataset

Class	Precision	Recall	F1-score	Support
Cyclone	0.9375	0.7500	0.8333	20
Earthquake	1.0000	1.0000	1.0000	20
Flood	0.7917	0.9500	0.8636	20
Landslide	1.0000	0.9500	0.9744	20
Wildfire	0.9048	0.9500	0.9268	20
Accuracy			0.9200	100
Macro avg	0.9268	0.9200	0.9196	100
Weighted avg	0.9268	0.9200	0.9196	100

The macro average computes the arithmetic mean of the evaluation metric across every class and contributes equally to the final result irrespective of class size.

$$\text{Macro Average} = \frac{1}{C} \sum_{i=1}^C M_i$$

where C = total number of classes and M_i = metric value (precision, recall, or F1 score) for class i

The weighted average computes the average metric by considering more samples in each class.

$$\text{Weighted Average} = \frac{\sum_{i=1}^C N_i M_i}{\sum_{i=1}^C N_i}$$

where N_i = support of class i and where M_i = metric value of class i

Since the dataset contains 20 testing samples for each of the five classes, the weighted average and macro average are nearly identical because the class distribution is balanced.

Overall Classification Accuracy

For the experiment,

$$\text{Accuracy} = \frac{92}{100} \times 100 = 92\%$$

indicates that 92 out of 100 test samples were correctly classified by the proposed RADIANT-Net framework.

The **confusion matrix** shows that RADIANT-Net can distinguish most of the disaster categories.

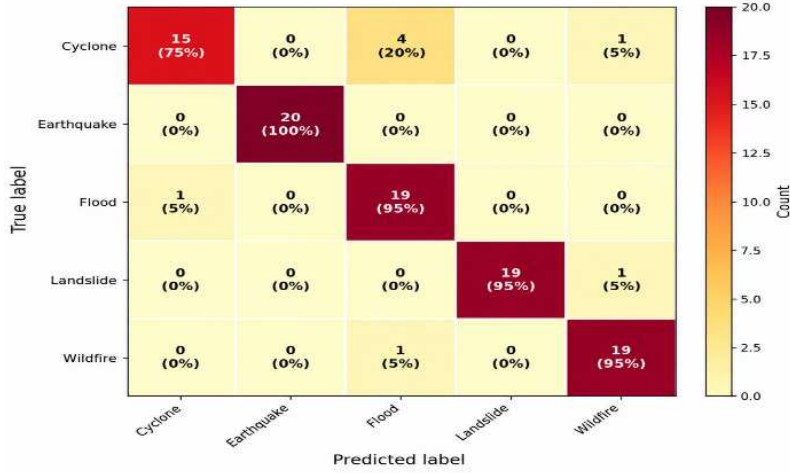


Fig. 8 Confusion matrix illustrating the classwise prediction performance

As illustrated in Fig. 8, the confusion matrix provides a detailed evaluation of the classification performance of the proposed RADIANT-Net in all five disaster categories. The classification accuracy was 92.00%, with 92 out of 100 test samples correctly classified. All 20 earthquake samples were correctly classified. Flood, landslide, and fire performed well, classifying 19 out of 20 samples. The most difficult category was Cyclone, with 15 of 20 samples identified. Four cyclone samples were misclassified as floods and one as a wildfire. One flood sample was incorrectly classified as a cyclone. In addition, a landslide sample was misclassified as a wildfire and a wildfire sample was predicted as a flood. The rest of the misclassifications are related mainly to cyclones and floods, which have similar meteorological and hydrological characteristics. Despite these challenging scenarios, the model still performs well for all classes of hazards, demonstrating its efficacy for multimodal prediction of natural disasters .

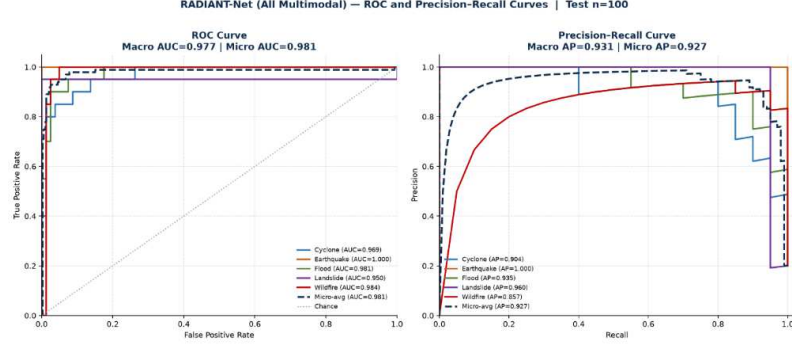


Fig. 9 ROC curves of the classes, precision–recall curves and corresponding average precision (AP) scores

As illustrated in Fig. 9, the precision recall (PR) curves for each disaster category were obtained by applying the one-versus-rest evaluation strategy to the independent test dataset. PR curves show the balance between precision and recall across different classification thresholds. The average precision (AP) score offers an overall measure of the model’s ranking ability for each disaster class. The proposed RADIANT-Net achieves a microaverage AP of 0.927, reflecting a good overall classification ability. The best results were obtained for the five categories of earthquake disasters, with a perfect AP of 1.000, followed by landslides (0.960), floods (0.935), cyclones (0.904), and wildfires (0.857). The relatively lower AP values for Cyclone and Wildfire imply that these classes are harder to classify, which is consistent with the classification report and the misclassification trends observed in the confusion matrix. In summary, the analysis of the PR curve indicates that the proposed RADIANT-Net has a good trade-off between precision and recall across different decision thresholds, demonstrating the strong probability estimation and reliable performance of multimodal disaster classification.

4.8 Comparative Analysis

Comparative evaluation of the proposed RADIANT-Net model against representative deep learning baseline models, namely, the CNN, LSTM, transformer, Vision transformer (ViT) and Swin transformer. To make a fair difference, all models were trained and evaluated on a common multimodal disaster data source, the same data partitioning strategy, and the same experimental settings. The proposed RADIANT-Net achieves the highest classification accuracy of 92.0%, outperforming all the baseline approaches. The other methods, Vision Transformer (ViT), achieved 80.0%, Swin Transformer (77.0%), LSTM (69.0%), CNN (62.0%) and the traditional Transformer (50.0%). The transformer-based vision models outperformed CNN and LSTM, but the classification accuracy was still significantly lower than that of the proposed framework. Table 7 compares the proposed approach with the latest methods.

Table 7: Table 7 Performance Comparison with deep learning baseline models

Model	Test Accuracy	Macro-F1
CNN	62.0%	0.5939
LSTM	69.0%	0.6699
Transformer	50.0%	0.4650
ViT	80.0%	0.7858
Swin Transformer	77.0%	0.7511
RADIANT-Net	92.0%	0.9196

4.9 Comparative evaluation with state-of-the-art models

This comparative analysis assesses RADIANT-Net with widely used deep learning architectures on the same all-source multimodal disaster dataset and the same training/test split.

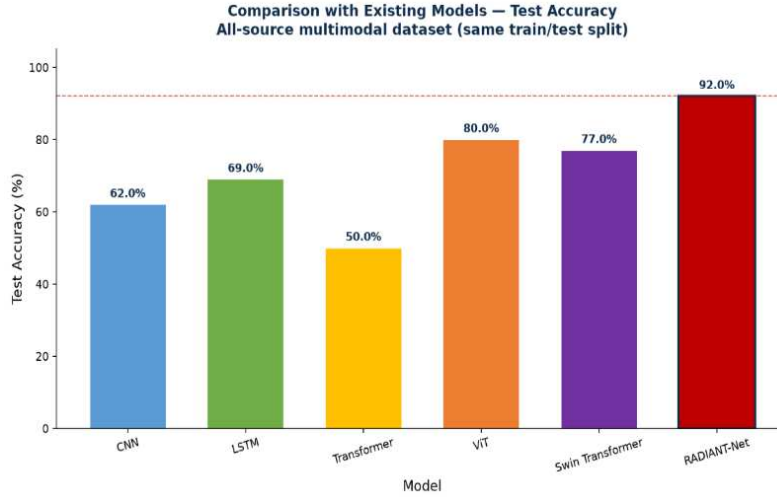


Fig. 10 Comparison of the accuracies of RADIANT-Net, CNN, LSTM, transformer, ViT and Swin transformer on the all-source multimodal test set

Fig. 10 highlights that the enhanced performance of RADIANT-Net is primarily attributed to its multimodal feature integration, the neural twin memory mechanism, and the agentic decision layer, which improve the feature representation and discrimination ability of heterogeneous disaster data. The experimental findings demonstrate that

the proposed framework is a more effective solution for the classification of multimodal natural disasters than the baseline models evaluated.

4.10 Computational Complexity

Table 8 summarizes the comparison of the computational complexity of the proposed RADIANT-Net with several representative deep learning models in terms of the trainable parameters, inference, floating point operations (FLOPs), and model size.

Table 8: Table 8 Comparison of trainable parameters, FLOPs, inference times, and model sizes for the proposed and baseline models

Model	Parameters	FLOPs	Inference Time (ms)	Model Size (MB)
CNN	10821	124256	0.691	0.045
LSTM	50757	977984	1.513	0.198
Transformer	67461	677376	2.175	0.272
ViT	69125	880512	2.402	0.279
Swin Transformer	69829	1219392	2.699	0.281
RADIANT-Net	45189	45472	0.868	0.183

The proposed framework has only 45,189 trainable parameters and requires only 45,472 FLOPs per sample, making it a computationally efficient architecture for multimodal disaster classification. The average inference time of RADIANT-Net was 0.868 ms with a compact model size of 0.183 MB, which was much lighter than the evaluated LSTM-, Transformer-, ViT- and Swin Transformer-based models. The proposed framework achieved the best classification accuracy among the compared methods, with relatively low computational complexity. The results show that RADIANT-Net can balance predictive performance and computational efficiency for real-time multimodal disaster prediction and resource-constrained deployment settings.

4.11 Comparison of training times

The training time for the performance of the proposed RADIANT-Net was evaluated against representative deep learning models, as shown in Table 9, in the same experimental settings, i.e., 30 training epochs, a batch size of 64 and CPU-based execution. The training of the proposed framework was performed within 3.43 s, which is better than that of all the baseline models evaluated based on computational efficiency. Among the compared approaches, the Swin transformer required the longest training time (19.11 s), followed by the remaining baseline models, while RADIANT-Net achieved the shortest training duration while maintaining the highest classification accuracy.

Table 9: Table 9 Comparative analysis of model training time under identical experimental settings

Model	Training	
	Epochs	Time
CNN	30	3.61 s
LSTM	30	4.18 s
Transformer	30	6.44 s
ViT	30	8.05 s
Swin	30	19.11 s
Transformer		
RADIANT-Net	30	3.43 s

These results, as shown in Table 9, indicate that the proposed architecture not only enhances predictive performance but also offers substantially reduced computational overhead, thereby making it suitable for real-time multimodal disaster prediction and resource-constrained deployment environments.

4.12 Ablation Study

We present an ablation study in Table X that evaluates the contribution of each architectural component and input modality under the same training conditions. Removing batch normalization, dropout or substituting the proposed architecture with a standard MLP led to a degradation in classification performance, thus confirming the effectiveness of the complete RADIANT-Net design. Similarly, modality-specific experiments revealed that meteorological features alone (64.7%) and satellite features alone (46.0%) were not sufficient to accurately predict disasters. These results suggest that multimodal feature fusion and the proposed neural twin memory and agentic decision modules are crucial for robust disaster classification. The findings obtained through the multiseed ablation studies presented in Table 10 indicate an average accuracy of 88.0%, while the main evaluation of the entire RADIANT-Net attained an accuracy of 92.0% in the independent test data set.

Table 10: Table 10 Ablation study evaluating the contributions of architectural components and inputs

Configuration	Accuracy		Macro-F1		Accuracy (%)	Macro-F1	Accuracy Plot (%)
	Mean	Std.	Mean	Std.		(Mean $\pm$ Std.)	
Full	0.880	0.024	0.881	0.024	88.0	0.881	88.0
RADIANT-Net					± 2.4	± 0.024	

Configuration	Accuracy		Macro-F1		Macro-F1		Accuracy Plot (%)
	Mean	Std.	Mean	Std.	Accuracy (%)	(Mean $\pm$ Std.)	
w/o Neural Twin	0.893	0.025	0.893	0.026	89.3 $\pm$ 2.5	0.893 $\pm$ 0.026	89.3
w/o Multimodal Encoder	0.883	0.009	0.883	0.010	88.3 $\pm$ 0.9	0.883 $\pm$ 0.010	88.3
w/o BatchNorm	0.870	0.036	0.871	0.035	87.0 $\pm$ 3.6	0.871 $\pm$ 0.035	87.0
w/o Dropout	0.873	0.026	0.869	0.031	87.3 $\pm$ 2.6	0.869 $\pm$ 0.031	87.3
MLP Only (No Encoder/Neural Twin)	0.863	0.038	0.863	0.039	86.3 $\pm$ 3.8	0.863 $\pm$ 0.039	86.3
w/o Sentinel-1/2 Features	0.883	0.021	0.881	0.023	88.3 $\pm$ 2.1	0.881 $\pm$ 0.023	88.3
Meteorological Features Only	0.647	0.009	0.627	0.009	64.7 $\pm$ 0.9	0.627 $\pm$ 0.009	64.7
Satellite Features Only	0.460	0.016	0.425	0.032	46.0 $\pm$ 1.6	0.425 $\pm$ 0.032	46.0

Comparison of the training times of the proposed RADIANT-Net with those of representative deep learning baseline models under the same experimental conditions: 30 training epochs, a batch size of 64, and CPU-based execution. The proposed framework ended the training process in 3.43 s, which is much faster than all the baseline models evaluated. In contrast, the Swin transformer has the longest training time (19.11 s), which is consistent with the greater computational complexity. Compared to the other methods, RADIANT-Net had the best classification accuracy with a significantly shorter training time. These findings demonstrate that the proposed architecture offers an efficient trade-off between computational cost and predictive performance, which makes it well suited for practical multimodal disaster prediction applications.

5 Conclusion

This study proposed RADIANT-Net, a multimodal deep learning framework for natural disaster classification with heterogeneous environmental, meteorological, remote sensing, and geospatial information. The proposed framework, by integrating complementary data from multiple sources, can effectively capture the complex relationships among disaster-related features and provide a reliable classification for five major disaster categories, namely flood, cyclone, landslide, wildfire, and earthquake. The

results of the experiments revealed that RADIANT-Net achieved an overall precision of 92.0%, accompanied by high precision, recall, and F1-scores for all disaster classes. Furthermore, comparative experiments have demonstrated that the proposed framework outperforms representative deep learning approaches, such as CNN, LSTM, transformer, Vision transformer (ViT) and Swin transformer, with lower computational complexity, shorter training times and smaller model sizes. The ablation study verified that both the proposed architectural components and the multimodal feature fusion mechanism contribute significantly to the observed performance gains, emphasizing the importance of integrating diverse data sources for disaster prediction. The results indicate that RADIANT-Net offers good coordination between prediction accuracy and computational efficiency and can be used in intelligent disaster monitoring and early warning applications. Although the current framework shows promising performance, future research will focus on validating the model with larger and more diverse real-world disaster datasets collected from multiple geographic regions; incorporating additional data modalities such as IoT sensor streams and social media reports; and extending the framework with physics-informed learning, graph neural reasoning, and explainable artificial intelligence techniques. The improvements are expected to increase the robustness, real-world applicability, and scalability of the proposed framework for real-time disaster management and decision support.

6 Future Scope

RADIANT-Net will be enhanced by embedding real-time multimodal data streams, including satellite imagery, Internet of Things (IoT) sensors, weather observations, and social media, to enable continuous disaster monitoring and early warning. Graph neural networks, physics-informed learning, and explainable AI will be added to enhance prediction accuracy, model interpretability, and decision support. Other future work will include the integration of the Digital Twin methodology and Agentic AI for autonomous disaster monitoring, validation of the framework on larger benchmark datasets from multiple geographic regions, and deployment on edge and cloud computing platforms for real-time operational use.

ETHICS DECLARATION

This manuscript represents original research and has been prepared in accordance with accepted standards of research and publication ethics. The manuscript has not been previously published and is currently not being considered for publication by another journal. The authors have taken appropriate measures to ensure the originality of the submitted work and to avoid plagiarism, inappropriate text reuse, data fabrication, falsification, and redundant publication. All authors approved and reviewed the final version of the manuscript and accept responsibility for the integrity, accuracy, and reliability of the work.

DATA AVAILABILITY

The datasets used in this study were obtained from publicly accessible sources, including Sentinel-1, Sentinel-2, GPM IMERG, NOAA, FIRMS, OpenStreetMap, WorldPop, and

CrisisNLP. The processed datasets and class labels are available from the corresponding author upon reasonable request.

FUNDING

This research did not receive external funding.

Acknowledgments

The authors acknowledge both institutions for their support

AUTHOR CONTRIBUTIONS

DM designed conceptualization and implemented the model architecture, methodology and training pipeline, conducted all experiments and statistical analyzes, and wrote the manuscript. JK supervised the work, data collection, and formal analysis, and SS performed formal analysis. SK contributed to the acquisition of funding and participated in the review and editing of the manuscript.

COMPETITIVE INTERESTS

The authors declare no conflict of interest.

Supplementary Information

Supplementary Information accompanies this article and includes additional figures and table details, and experimental results.

References

- Akhyar Akhyar, Mohd Asyraf Zulkifley, Jaesung Lee, Taekyung Song, Jaeho Han, Chanhee Cho, Seunghyun Hyun, Youngdoo Son, and Byung Woo Hong. Deep artificial intelligence applications for natural disaster management systems: A methodological review. *Ecological Indicators*, 163:112067, 2024. doi: 10.1016/j.ecolind.2024.112067.
- J. Akram, W. Hussain, R. H. Jhaveri, R. S. Rathore, and A. Anaissi. Dynamic GNN-based multimodal anomaly detection for spatial crowdsourcing drone services. *Digital Communications and Networks*, 11(5):1639–1656, 2025. doi: 10.1016/j.dcan.2025.03.012.
- S. Aththanayake, C. Mallikarachchi, J. Wickramasinghe, S. Kugarajah, D. Meedeniya, and B. Pradhan. ResQConnect: An AI-powered multi-agentic platform for human-centered and resilient disaster response. *Sustainability*, 18(2):1014, 2026. doi: 10.3390/su18021014.
- Kaifeng Bi, Lingxi Xie, Hengheng Zhang, Xin Chen, Xiaotao Gu, and Qi Tian. Accurate medium-range global weather forecasting with 3D neural networks. *Nature*, 619(7970):533–538, 2023. doi: 10.1038/s41586-023-06185-3.
- K. Cao, T. Zhang, and J. Huang. Advanced hybrid LSTM-transformer architecture for real-time multi-task prediction in engineering systems. *Scientific Reports*, 14: 5568, 2024. doi: 10.1038/s41598-024-55483-x.
- Vinay Chamola, Vikas Hassija, Sarthak Gupta, Abhinav Goyal, Mohsen Guizani, and Biplab Sikdar. Disaster and pandemic management using machine learning: A survey.

- IEEE Internet of Things Journal*, 8(21):16047–16071, 2021. doi: 10.1109/JIOT.2020.3044966.
- M. Fawakherji and L. Hashemi-Beni. Flood detection and mapping through multi-resolution sensor fusion: Integrating UAV optical imagery and satellite SAR data. *Geomatics, Natural Hazards and Risk*, 16(1):2493225, 2025. doi: 10.1080/19475705.2025.2493225.
- M. Habib, A. Habib, and B. Alibrahim. Applications of physics-informed neural networks in geosciences: From basic seismology to comprehensive environmental studies. *Open Geosciences*, 17(1):20250853, 2025. doi: 10.1515/geo-2025-0853.
- C. Huang, P. Mu, J. Zhang, S. Chan, S. Zhang, H. Yan, S. Chen, and C. Bai. Benchmark dataset and deep learning method for global tropical cyclone forecasting. *Nature Communications*, 16:5923, 2025. doi: 10.1038/s41467-025-61087-4.
- George Em Karniadakis, Ioannis G. Kevrekidis, Lu Lu, Paris Perdikaris, Sifan Wang, and Liu Yang. Physics-informed machine learning. *Nature Reviews Physics*, 3: 422–440, 2021. doi: 10.1038/s42254-021-00314-5.
- W. Li et al. An interpretable hybrid deep learning model for flood forecasting based on transformer and LSTM. *Journal of Hydrology: Regional Studies*, 54:101873, 2024. doi: 10.1016/j.ejrh.2024.101873.
- S. Lin, Z. Liang, H. Guo, Q. Hu, X. Cao, and H. Zheng. Application of machine learning in early warning system of geotechnical disaster: A systematic and comprehensive review. *Artificial Intelligence Review*, 58:169, 2025. doi: 10.1007/s10462-025-11175-0.
- A. H. Mustajab, H. Lyu, Z. Rizvi, and F. Wuttke. Physics-informed neural networks for high-frequency and multi-scale problems using transfer learning. *Applied Sciences*, 14(8):3204, 2024. doi: 10.3390/app14083204.
- M. Pianforini, S. Dazzi, A. Pilzer, and R. Vacondio. Real-time flood maps forecasting for dam-break scenarios with a transformer-based deep learning model. *Journal of Hydrology*, 635:131169, 2024. doi: 10.1016/j.jhydrol.2024.131169.
- E. Portalés-Julià, G. Mateo-García, and L. Gómez-Chova. Understanding flood detection models across Sentinel-1 and Sentinel-2 modalities and benchmark datasets. *Remote Sensing of Environment*, 328:114882, 2025. doi: 10.1016/j.rse.2025.114882.
- S. Rahman, M. F. Faisal, P. Mondal, and M. M. Rahman. Tropical cyclone track prediction harnessing deep learning algorithms: A comparative study on the northern indian ocean. *Results in Engineering*, 26:105009, 2025. doi: 10.1016/j.rineng.2025.105009.
- K. Riaz, M. McAfee, and S. S. Gharbia. Management of climate resilience: Exploring the potential of digital twin technology, 3D city modelling, and early warning systems. *Sensors*, 23(5):2659, 2023. doi: 10.3390/s23052659.
- D. Shadrin, S. Illarionova, F. Gubanov, K. Evteeva, M. Mironenko, I. Levchunets, R. Belousov, and E. Burnaev. Wildfire spreading prediction using multimodal data and deep neural network approach. *Scientific Reports*, 14:2606, 2024. doi: 10.1038/s41598-024-52821-x.
- N. Shafiee Roudbari, S. R. Puneekar, Z. Patterson, U. Eicker, and C. Poullis. From data to action in flood forecasting leveraging graph neural networks and digital twin visualization. *Scientific Reports*, 14:18571, 2024. doi: 10.1038/s41598-024-68857-y.

- E. Shokrgozar-Yatimdar and P. Chen. Explainable deep learning for real-time prediction of uniform hazard spectral acceleration for on-site earthquake early warning. *Geophysical Journal International*, 243(2):ggaf345, 2025. doi: 10.1093/gji/ggaf345.
- M. Taghizadeh, Z. Zandsalimi, M. A. Nabian, M. Shafiee-Jood, and N. Alemazkoor. Interpretable physics-informed graph neural networks for flood forecasting. *Computer-Aided Civil and Infrastructure Engineering*, 40:2629–2649, 2025. doi: 10.1111/mice.13484.
- Timothy Tiggeloven, Samira Pfeiffer, Alessia Matanó, Marc van den Homberg, Lisa Thalheimer, Markus Reichstein, and Silvia Torresan. The role of artificial intelligence for early warning systems: Status, applicability, guardrails, and ways forward. *iScience*, 28(11):113689, 2025. doi: 10.1016/j.isci.2025.113689.
- M. Umadevi, J. A. Kumar, S. V. Priyan, and C. Vivek. Design of mTCN framework for disaster prediction: A fusion of massive machine type communications and temporal convolutional networks. *Scientific Reports*, 15:28624, 2025. doi: 10.1038/s41598-025-15397-8.
- Bo Wang, Han Shi, Yueming Lucy Qiu, Nana Deng, Destenie Nock, Xingchi Shen, Zhaohua Wang, and Yi David Wang. Unequal power outages induced by natural disasters. *Nature Communications*, 16(1):8947, 2025a. doi: 10.1038/s41467-025-64012-x.
- H. Wang, J. Chen, Y. Zheng, and X. Song. Accelerating flood warnings by 10 hours: The power of river network topology in AI-enhanced flood forecasting. *npj Natural Hazards*, 2:45, 2025b. doi: 10.1038/s44304-025-00083-6.
- A. Wibowo, I. Amri, A. Surahmat, and R. Rusdah. Leveraging artificial intelligence in disaster management: A comprehensive bibliometric review. *Jamba: Journal of Disaster Risk Studies*, 17(1):1776, 2025. doi: 10.4102/jamba.v17i1.1776.
- Lucas Windheuser, R. Karanjit, R. Pally, S. Samadi, and Nina C. Hubig. An end-to-end flood stage prediction system using deep neural networks. *Earth and Space Science*, 10(1):e2022EA002385, 2023. doi: 10.1029/2022EA002385.
- M. Yao, D. Tian, G. Hu, Q. Kong, Y. Li, and J. Shu. Research on landslide displacement prediction based on transformer. *ce/papers*, 8(2):613–619, 2025. doi: 10.1002/cepa.3150.
- S.-J. Yun, J.-W. Kwon, and W.-T. Kim. A novel digital twin architecture with similarity-based hybrid modeling for supporting dependable disaster management systems. *Sensors*, 22(13):4774, 2022. doi: 10.3390/s22134774.
- H. Zhang, S. Ke, W. Liu, and Y. Zhang. A combining earthquake forecasting model between deep learning and epidemic-type aftershock sequence (ETAS) model. *Geophysical Journal International*, 239(3):1545–1556, 2024. doi: 10.1093/gji/ggae349.
- Z. Zhang, D. Wang, Y. Mei, J. Zhu, and X. Xiao. Developing an explainable deep learning module based on the LSTM framework for flood prediction. *Frontiers in Water*, 7:1562842, 2025. doi: 10.3389/frwa.2025.1562842.
- D. Zhexebay et al. Deep learning for early earthquake detection: Application of convolutional neural networks for P-wave detection. *Applied Sciences*, 15(7):3864, 2025. doi: 10.3390/app15073864.
- L. Zhu, L. Huang, L. Fan, J. Huang, F. Huang, J. Chen, Z. Zhang, and Y. Wang. Landslide susceptibility prediction modeling based on remote sensing and a novel

deep learning algorithm of a cascade-parallel recurrent neural network. *Sensors*, 20 (6):1576, 2020. doi: 10.3390/s20061576.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryInformation.docx](#)