A Novel AI-Based Framework for Remote Sensing Data Analysis

Nguyen Ngoc Quang[1], Do Anh Huy[2], Nguyen Hoang Long[2], Nguyen Quang Minh[0000-0003-2951-8332], 3,*

[1] Head of VegaCosmos, Vietnam
quangnnvs@vegastar.com.vn
[2] Technical consultant team of Geospatial Department, VegaCosmos, Vietnam
dohuyvs@vegastar.com.vn
longnhvs@vegastar.com.vn
[3] Hanoi University of Mining and Geology, Vietnam
nguyenquangminh@humg.edu.vn

Abstract. Remote sensing data plays a crucial role in environmental monitoring, disaster assessment, and urban planning. However, efficiently acquiring, analyzing, and visualizing such data from multiple sources remains a challenge. This paper introduces GeoHUB (https://geohub.ai/), an AI-powered platform designed to streamline remote sensing data analysis by integrating conventional and artificial intelligence-based processing techniques. GeoHUB supports data acquisition from diverse satellite sources, including SPOT, Sentinel, Landsat, Pleiades, KOMPSAT, COSMO-SkyMed, and UAV imagery, ensuring comprehensive data accessibility. The platform provides advanced analytical tools for image classification, change detection, and feature extraction, leveraging deep learning and machine learning models to enhance interpretation accuracy. Furthermore, GeoHUB seamlessly integrates with Geographic Information Systems (GIS) to facilitate spatial analysis, querying, and result visualization. A case study is presented to demonstrate GeoHUB's capabilities in assessing the damage caused by Typhoon Yagi (2024) in Hai Phong, Vietnam, highlighting its effectiveness in rapid disaster impact evaluation. The results underscore the potential of GeoHUB as a powerful decision-support system based on remote sensing applications across various domains.

Keywords: Remote Sensing, Artificial Intelligence, GeoHUB, GIS Integration, Disaster Assessment, Satellite and UAV Imagery.

1 Introduction

The rapid advancement of remote sensing technology has led to the acquisition of vast amounts of multi-source Earth observation data, encompassing various types, time periods, and scales [1]. The abundance of available imagery enables remote sensing to provide reliable data for various sectors, including agriculture, urban and environmental management, as well as disaster monitoring and mitigation [2]. Conventionally, the processing and analysis of remotely sensed imagery have relied on desktop-based software, which requires users to have highly specialized knowledge and skills—skills that are not always available across different domains. Another limitation of desktop-based platforms is their restricted capacity to handle large volumes of data. This has led to the development of various cloud-based platforms. Some well-known examples include Google Earth Engine [3, 4], NASA Earth Exchange [5], Descartes Labs (https://descarteslabs.com/), Data Cube [6], Planetary Computer (https://planetarycomputer.microsoft.com/), the CASEarth EarthDataMiner [7], PIE-Engine (https://www.piesat.cn/en/PIE-Engine.html), Amazon (https://aws.amazon.com/earth/). These platforms are used widely for remote sensing data analysis in various areas such as water management [8], agriculture [9, 10], deforestation [11], climate monitoring [12] and environmental monitoring [13].

With the advancement of deep learning, many remote sensing imagery analysis tasks such as scene classification, object detection, land-use classification, change detection, and multi-view 3D reconstruction can now be performed automatically without the need for training sample [14]. Pre-trained deep learning models and tools hosted on cloud platforms enable the interpretation of raw data, transforming it into more meaningful data for various applications across different domains [15]. However, these platforms are limited in that they primarily focus on data interpretation, while other crucial steps such as data acquisition, analysis, and visualization remain separate, making it challenging for users to efficiently integrate and utilize the data. GeoHUB is designed to overcome these limitations by integrating all three key processes such as remote sensing imagery acquisition, interpretation, and spatial analysis into a unified platform. Additionally, it leverages GIS functionality to seamlessly present results, providing a comprehensive solution for remote sensing data analysis.

2 Conceptual design of GeoHUB

The vegastar remote sensing and data analysis online model is designed based on technologies such as: (1) WebGIS, providing web-based access to GIS tools and data repositories, where users can interact with geospatial information through a simple browser without resorting to specialized software and required expertise; (2) Spatial Data Infrastructure (SDI), framework to support geospatial data sharing and integration across different systems and organizations, employed to ensure data interoperability and accessibility to form a cohesive geospatial ecosystem; (3) Geographic Information System, the foundational technology for GeoHub, enabling the acquisition, storage, management, analysis, and visualization of spatial data, allowing users to create interactive queries, edit map data, and present results in different formats, i.e., maps, charts, or reports; (4) Remote Sensing, integrated into GeoHub to enable data collection of Earth's surface from databases of operating satellites worldwide, essential for the exploitation of wide coverage, cost-effective spatial data used in monitoring and analysis applications over remote areas, providing critical insight to environmental challenges; (5) Artificial Intelligence (AI) and Machine Learning (ML), complementing geospatial data workflows by automating complex processing, reserved for advanced applications, i.e., object detection, image classification, and spatial pattern recognition, improving data processing speed, allowing faster decision-making in a timely manner; (6) 3D visualization and modeling, enabling presentation of geospatial data in three dimensions, particularly useful for applications which require high degree of spatial awareness and relationships between entities, i.e., urban planning, infrastructure architecture, and environmental simulation.

GeoHUB consists of five functional modules as presented in conceptual design of the system as in Figure 1 and the interface is presented in Figure 2.

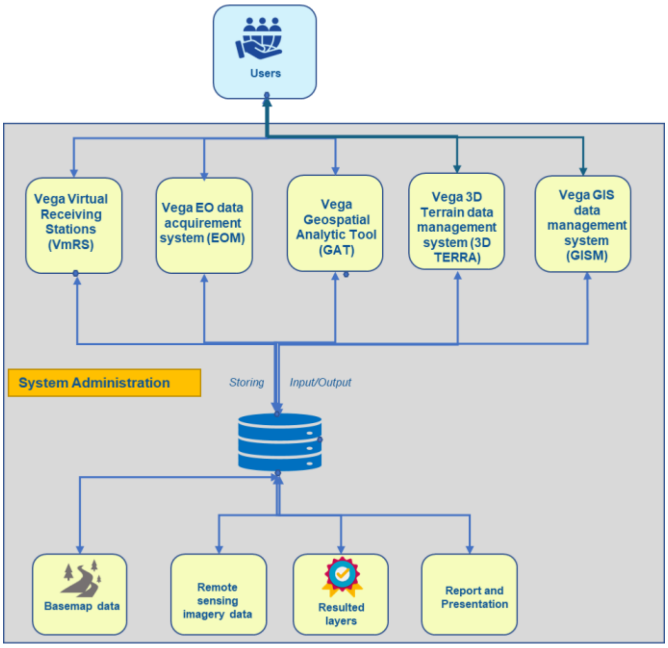

Figure 1 Conceptual design of GeoHUB

2.1 Vega Virtual multi Receiving Stations (VmRS)

VmRS is a module which comprises of functions for accessing the satellite data. By connecting to satellite databases from the world's leading satellite data providers, these solutions ensure seamless access to high-quality remote sensing imagery. They also optimize data management and processing, enabling efficient handling of large datasets for various applications. Additionally, advanced geospatial platforms offer project planning support, assisting users in analyzing spatial data, making informed decisions, and enhancing the overall efficiency of geospatial projects across different domains.

2.2 Vega EO data management system (EOM)

Providing comprehensive multi-source satellite data management solutions, the system integrates advanced automation features and seamless compatibility with the

STAC 1.0 standard. It enables efficient satellite data synchronization, ensuring up-to-date and accessible information. Additionally, users can perform advanced data searches based on spatial parameters, attributes, and classification for precise data retrieval. The platform also supports various additional storage sources, including FTP and WebDAV, offering flexible and scalable data management capabilities.

2.3 Vega Geospatial Analytic Tool (GAT)

GAT is a set of advanced tools for object detection, semantic segmentation and monitoring by conventional and AI methods. The system supports tracking, mapping, and digital modeling for a wide range of applications. It enables the identification of objects on the Earth's surface and facilitates the monitoring of terrain and environmental changes. Additionally, it allows for the rapid detection of changes in structures, forest cover, shoreline dynamics, and other critical features over time. The system also supports agricultural mapping, including rice crop identification, crop yield forcasting, and land cover classification. Furthermore, it provides essential monitoring capabilities for assessing water, forest, air quality, and land use, ensuring comprehensive environmental management and analysis. Table 1 presents analytic tools integraded in GeoHUB.

***Table 1**: GeoHUB's analytic tools*

Analytic Tools		**Data used**	**Description**
Processing	Interferometry Generation	Synthetic Aperture Radar	Creates interferograms from SAR remote sensing data based on values of phase difference between acquisition instances. Determine the corresponding elevation changes or surface displacement of the area of interest during a time interval.
	DEM Generation		Generates a 30-m resolution Digital Elevation Model (DEM) from a pair of SAR images. Allows loading auxiliary data to perform an accuracy assessment of the results.
	Pansharpening	Optical	Fusion technique that leverages the desirable traits of very high-resolution panchromatic images and lower-resolution multispectral images. Produces a single very high-resolution multispectral images.
	Mosaic		Combines multiple remote sensing images having similar resolution and overlapping regions to create a single composite image. Used for expanding the total area and scope of the area of interest.
Detection & Classification	Aerial-photo Detection	Optical	Utilizes diverse remote sensing data from both satellites and modern object detection modules such as YOLOv5/v8. Includes identifying and locating positions, and determining aircraft quantity.

	Change Detection		Quantifies surface transformation across different period (s) using satellite image time series of an area of interest. Evaluate current conditions, identify significant changes, and develop urban development strategies based on empirical data.
	Building Detection		Leverages very high-resolution remote sensing imagery combined with advanced AI models to automatically detect and identify buildings. Resulted information is integrated into urban planning proposals and land use maps.
	Ship Detection		Combines remote sensing data and advanced deep learning algorithms for automatic identification and classification of ship variants. Supports further tracking objects of interest in maritime, coastal, and island regions.
	Surface Classification		Employs Random Forest machine learning model for classification, analysis, and identification of primary land cover types, mainly consisting of: constructions, bare soil, vegetation, and water bodies.
Visualization	Basemap	Open Basemap Database	Allows users to display available processing results obtained using our tool suite on a provided base map. Enables the selection of different base map types, data upload, and visualization of remote sensing data. Supports common geospatial data formats such as *.shp, *.geojson, *.json, *.tif, *.tiff, *.kml, and *.wkt.

2.4 Vega 3D Terrain Data Management System (3D TERRA)

3D TERRA is a powerful solution for managing and processing 3D terrain data from various formats, including **OBJ, FBX, and LAS**. It enables the automatic generation of **3D maps** from raw data while efficiently managing multiple **3D layers**. The system provides **interactive visual displays**, allowing users to seamlessly explore and interact with 3D terrain models. Additionally, it offers **advanced analytics** capabilities, including **terrain measurement, visibility analysis, and real-time simulation** integrated with weather data, making it an essential tool for geospatial analysis and decision-making.

2.5 Vega GIS Data Management System (GISM)

Vega GIS Data Management System (GISM) is a comprehensive platform for managing and editing map data, offering powerful tools for handling vector data layers and generating detailed reports and statistics through SmartReport. It enables users to efficiently manage and edit vector-based map layers, ensuring accurate representation of geospatial data. Additionally, the system supports document management, report creation, and statistical analysis, enhancing data-driven decision-making. With robust spatial and attribute data management capabilities, GISM provides a seamless and efficient solution for GIS professionals across various domains.

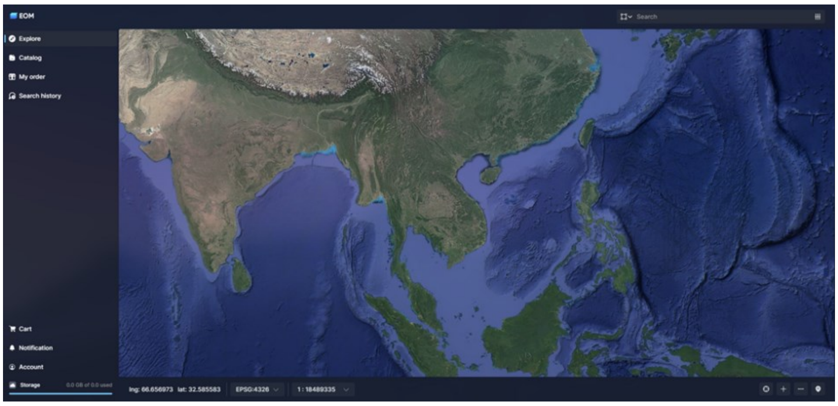

Figure 2 Interface of GeoHUB

3 Case study: Evaluation of the damage of Typhoon Yagi in Hai Phong Province by GeoHUB

3.1 Yagi Typhoon and research location

Typhoon Yagi was an exceptionally powerful Category 4 storm, ranking as one of the strongest tropical cyclones to ever strike northern Vietnam in recorded history. Its impact on the Red River Delta Region was unprecedented, both in terms of its intensity and the widespread damage it caused. After crossing southern China, Yagi maintained its strength over the warm waters of the Gulf of Tonkin before making landfall near Hai Phong City on September 7th, 2024. With wind gusts reaching nearly 200 km/h, the typhoon left a trail of destruction across the northern coast of Vietnam. In just a 24-hour period, it affected multiple cities, including Haikou and Hanoi, showcasing its rapid and devastating journey. The aftermath of Typhoon Yagi brought significant economic and environmental tolls to Hai Phong, a coastal city now facing the long-term consequences of this catastrophic event. This case study is for Eevaluation of the impact of Typhoon Yagi to some areas Hai Phong City using GeoHUB.

3.2 Data

Sentinel-1 Synthetic Aperture Radar (SAR) images were obtained both before and after Typhoon Yagi, specifically on August 31st, 2024, and September 10th, 2024, using VH polarization. These images provided crucial data for assessing the storm's impact on the region. Additionally, a Digital Elevation Model (DEM) derived from commercial SRTM data was utilized to calculate the slope degree, which proved essential in enhancing the distinguishing water surfaces from higher terrains such as hills and mountains. To support the analysis, auxiliary data included the administrative boundaries of Hai Phong City and damage statistics from the Hai Phong City Information Portal, which provided detailed insights into the extent of the devastation caused by Typhoon Yagi. These datasets collectively contributed to a comprehensive assessment of the typhoon's impact on the city’s landscape and infrastructure

3.3 Method

Figure 3 depicts a workflow using Google Earth Engine (GEE) as the underlying data processing and analysis platform, with GeoHUB acting as an interface that provides specific functionalities to end-users.

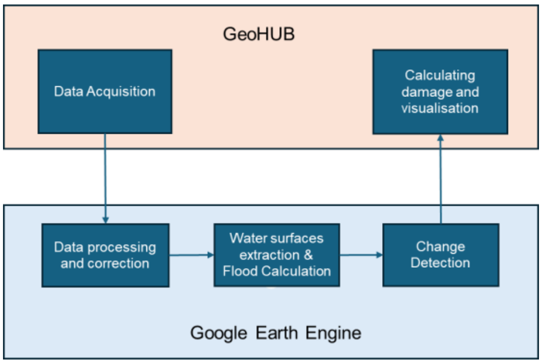

Figure 3 Workflows for Typhoon damage estimation

Data Acquisition: This function uses the GeoHUB VmRS to query and download the Sentinel data and stores the data in GeoHUB account for processing and analysis.

Data Processing and correction: The Sentinel-1 preprocessing process involves several key steps to prepare the data for analysis. First, speckle filtering is applied to reduce noise and improve image clarity. Next, the data undergoes georeferencing and terrain correction to align it accurately with geographic coordinates and account for topographic effects. Finally, the image is converted to dB (decibels), a logarithmic scale used to enhance the contrast of the radar backscatter, making it more suitable for interpretation and further analysis. These steps ensure the data is precise and ready for detailed processing.

Water surfaces extraction and flood calculation: The process of identifying water surfaces using radar signals of polarization VH involves several steps to ensure accurate detection. First, thresholding is applied to the VH signals in dB, where water surfaces typically exhibit low VH values. These values are determined either empirically or through histogram analysis. The preliminary water surface detection step involves identifying areas where the VH values exceed the set threshold, classifying them as potential water surfaces. To refine the identification, DEM data is incorporated to eliminate false positives, such as hillsides mistakenly identified as water surfaces. This step uses slope data (representing elevation) to remove areas with high elevation, which may have low VH values due to radar shadows but are not water surfaces. Finally, a water binary mask is constructed, where pixels identified as true water surfaces are labeled with a value of 1, accurately distinguishing real water bodies from other features.

The process of calculating flood depth on the Google Earth Engine (GEE) platform using Harvard University's FwDET model (FwDETv2.0) involves several key steps. First, a conceptual and technical replication of the FwDET model is performed to ensure accuracy in the analysis. Local outlier filtering is then applied to flood margin elevation pixels using the modified Z-score, ensuring that extreme elevation values do not skew the results. Following this, a new estimated water surface elevation is constructed using a cost accumulation algorithm, which accounts for terrain variation and flow patterns. The modeled flood surface elevation is then subtracted from the smoothed DEM to calculate the flood depth. Finally, a low-pass convolution is applied to the resulting flood depth estimation layer to smooth the data and enhance its accuracy, providing a clear representation of flood depth across the affected area.

Change detection: The process of evaluating changes in reflective objects on the land surface using multi-temporal combination of two Sentinel-1 images on the Google Earth Engine (GEE) platform involves creating an RGB composite image. This composite combines two images taken before and after the Typhoon, where the image before the Typhoon is assigned to the Red (R) and Green (G) bands, and the image after the Typhoon is assigned to the Blue (B) band. The resulting image allows for easy identification of areas that have undergone significant change. For instance, regions that were present before the Typhoon but are no longer visible afterward will appear yellow on the composite, indicating objects that may have been destroyed or relocated. This multi-temporal analysis effectively highlights areas affected by the Typhoon and provides valuable insights into the extent of damage.

3.4 Results and discussions

With GeoHUB, the resulting data can be presented in a highly intuitive manner. For instance, the points and areas impacted by Typhoon Yagi can be linked with available videos, photos, and manually input data. Additionally, reports can be customized to enhance visual appeal, making them more engaging and accessible to the audience. Figure 4 is an example of the data visuallization using GeoHUB.

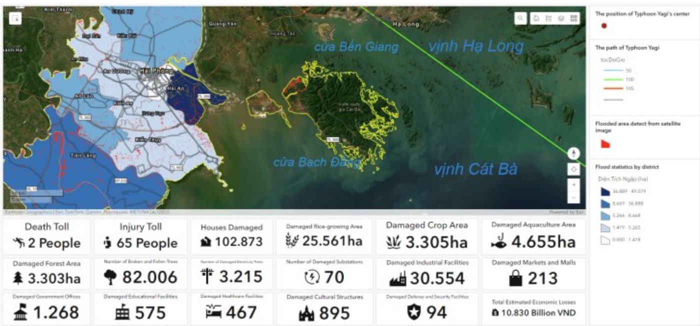

Figure 4 Visualization of resulting from remote sensing and General Statistic Office of Vietnam

This data can be presented alongside the estimated damage calculated from the previously described process. GeoHUB provides valuable information for disaster mitigation and response. By using forecasts of rainfall and Typhoon Yagi’s path, it is possible to predict the location and extent of the typhoon's impact. With this information, both local and central governments can make informed decisions on the appropriate measures for disaster response. GeoHUB also provides a simulation tools to visualize the pathway of the typhoon as well as the water level and flooding area and geolocation of the possible most heavily affected areas as in Figure 5 and Figure 6.

Figure 5 Simulation of typhoon path and flooding area

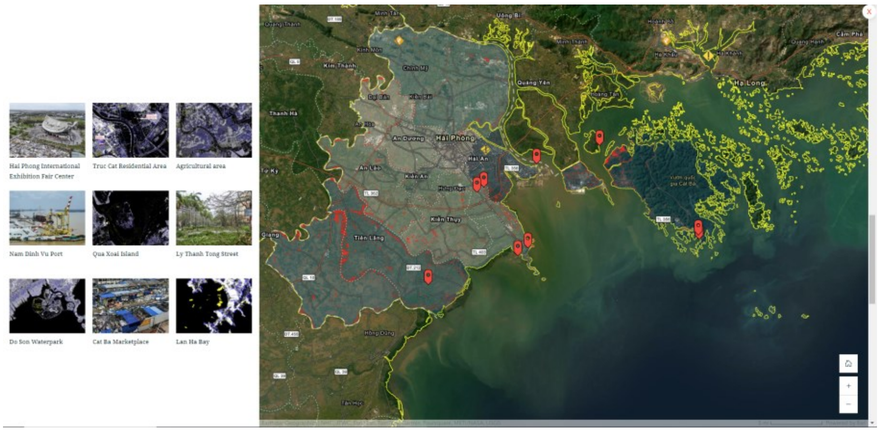

Figure 6 Anticipated locations of most heavily affected areas

GeoHUB can also be used to assess damage following a typhoon. An example of damage detection in remote sensing imagery is observed in Truc Cat Residential Area, where a comparison of two Sentinel-1 images taken before (August 31st, 2024) and after (September 10th, 2024) the typhoon highlights significant changes (Figure 7). The bright yellow areas indicate locations where objects were present before the storm but have since disappeared or shifted. Meanwhile, the blue areas represent regions that were previously flat but became more rugged, or where new objects have emerged or changed positions.

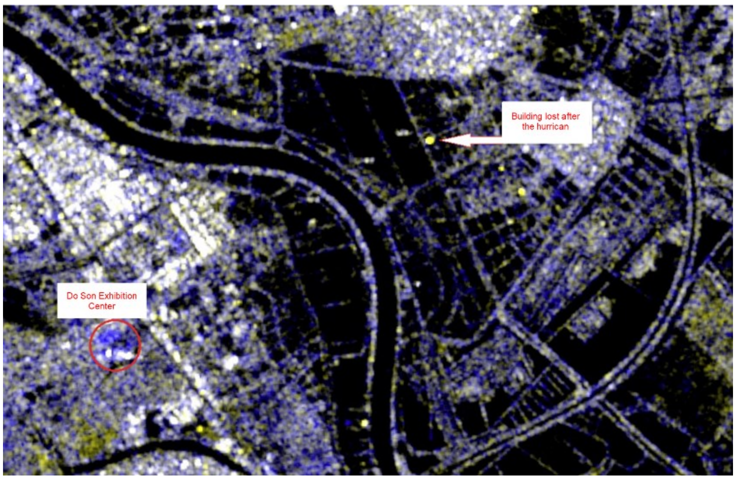

Figure 7 Damages in Truc Cat Residential Area

Another example of damage assessment is in the Agricultural Area, where a comparison of two Sentinel-1 images taken before (August 31st, 2024) and after (September 10th, 2024) the storm reveals significant changes as in Figure 8. The blue areas indicated regions that had relatively flat surfaces before the storm but became more rugged or where new objects appeared, suggesting structural changes. Meanwhile, the dark yellow areas represent regions with substantial object changes or partially flooded areas. The analysis results show that most fields appear in blue, indicating that crops may have been damaged, broken, or flattened, leading to noticeable changes in the satellite imagery. The dark yellow regions may correspond to areas that remain flooded after the storm.

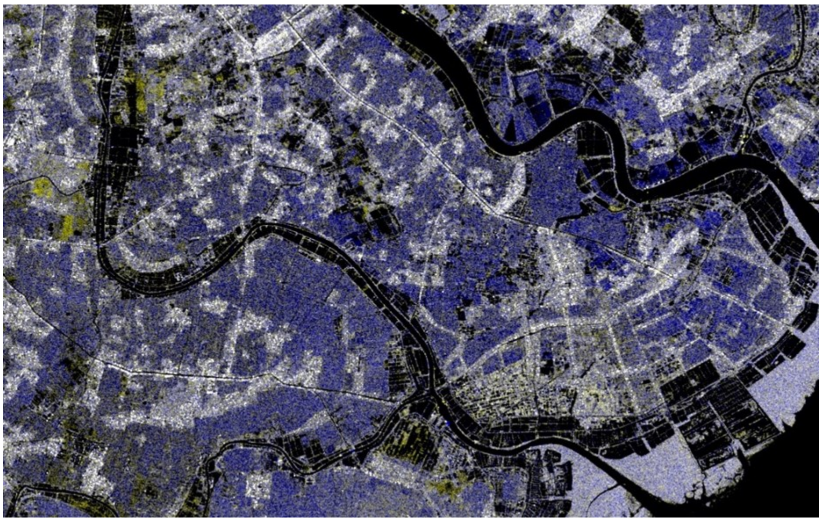

Figure 8 Assessment of damage for agricultural area

Figure 9 Disappearance of floating fish rafts in Lanha Bay

Due to the limited availability of data at the time, a formal accuracy assessment using empirical data and detailed damage statistics was not conducted. However, a preliminary evaluation of the platform's accuracy was performed through visual comparisons of pre- and post-storm imagery, demonstrating a strong correlation with observed damage. A comparison of 10 locations where changes were detected showed a 100% accuracy rate, with fish raft destruction (as illustrated in Figure 9) serving as a representative example.

Despite the lack of a formal validation process, the integration of radar data in our analysis provides significant credibility to the results. Radar data delivers direct measurements of physical properties, enabling precise detection of real-world flood dynamics. While its complexity may introduce interpretive challenges, this report effectively harnesses its strengths to generate reliable and insightful assessments of the disaster's impact.

4 Conclusions

This paper presents the conceptual design and methodologies for developing the GeoHUB platform, a cloud-based system for data acquisition, analysis, and presentation of remote sensing data, including both satellite and aerial imagery. The platform leverages AI models, conventional algorithms, and GIS spatial analysis to enhance data processing and interpretation.

The case study shows that GeoHUB is a powerful and versatile system that seamlessly integrates the processes of data retrieval, data analysis, and GIS-based result analysis. Its open architecture facilitates continuous advancements, allowing for the integration of new algorithms and machine learning models, which enhances its capability in data analysis. This flexibility makes GeoHUB an invaluable tool for a wide range of applications, particularly in fields such as disaster response, environmental monitoring, and urban planning, where dynamic and scalable data processing is crucial. Currently, GeoHUB still has limited computational capacity, especially for functions involving deep learning. However, these limitations will soon be resolved through the enhancement of computing resources for AI training and deployment.

References

1. Samadzadegan F, Toosi A, Dadrass Javan F (2025) A critical review on multi-sensor and multi-platform remote sensing data fusion approaches: current status and prospects. Int J Remote Sens 46:1327–1402. https://doi.org/10.1080/01431161.2024.2429784
2. Kumar L, Mutanga O (2018) Google Earth Engine applications since inception: Usage, trends, and potential. Remote Sens 10:1–15. https://doi.org/10.3390/rs10101509
3. Tamiminia H, Salehi B, Mahdianpari M, et al (2020) Google Earth Engine for geo-big data applications: A meta-analysis and systematic review. ISPRS J Photogramm Remote Sens 164:152–170
4. Mutanga O, Kumar L (2019) Google earth engine applications. Remote Sens. 11:591
5. Nemani R (2011) Nasa earth exchange: Next generation earth science collaborative. ISPRS-International Arch Photogramm Remote Sens Spat Inf Sci 3820:17
6. Lewis A, Oliver S, Lymburner L, et al (2017) The Australian geoscience data cube—foundations and lessons learned. Remote Sens Environ 202:276–292
7. Liu J, Wang W, Zhong H (2020) EarthDataMiner: a cloud-based big earth data intelligence analysis platform. In: IOP Conference Series: Earth and Environmental Science. IOP Publishing, p 12032
8. Pekel J-F, Cottam A, Gorelick N, Belward AS (2016) High-resolution mapping of global surface water and its long-term changes. Nature 540:418–422
9. Kontgis C, Warren MS, Skillman SW, et al (2017) Leveraging Sentinel-1 time-series data for mapping agricultural land cover and land use in the tropics. In: 2017 9th International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp). IEEE, pp 1–4
10. Xiong J, Thenkabail PS, Gumma MK, et al (2017) Automated cropland mapping of continental Africa using Google Earth Engine cloud computing. ISPRS J Photogramm Remote Sens 126:225–244
11. Brovelli MA, Sun Y, Yordanov V (2020) Monitoring forest change in the amazon using multi-temporal remote sensing data and machine learning classification on Google Earth Engine. ISPRS Int J Geo-Information 9:580
12. Huntington JL, Hegewisch KC, Daudert B, et al (2017) Climate engine: Cloud computing and visualization of climate and remote sensing data for advanced natural resource monitoring and process understanding. Bull Am Meteorol Soc 98:2397–2410
13. Abdelmajeed AYA, Albert-Saiz M, Rastogi A, Juszczak R (2023) Cloud-based remote sensing for wetland monitoring—a review. Remote Sens 15:1660
14. Li J, Huang X, Gong J (2019) Deep neural network for remote-sensing image interpretation: Status and perspectives. Natl Sci Rev 6:1082–1086
15. Zhang Z, Zhang M, Gong J, et al (2023) LuoJiaAI: A cloud-based artificial intelligence platform for remote sensing image interpretation. Geo-Spatial Inf Sci

26:218–241