Estimation of Landslide Hazard Zones Using Deep Learning Based on Diverse Geospatial Data

Kohei Arai, Kengo Oiwane, Hiroshi Okumura
Information Science Dept., Saga University, Saga City, Japan

***Abstract*—Traditional landslide hazard mapping in Japan relies on labor-intensive field surveys, which are slow, costly, and fail to update dynamically amid rising climate-driven disasters like the 2018 Heavy Rain Event, leaving gaps in timely evacuations. This study addresses these challenges by proposing a semantic segmentation framework using ResUNet to fuse Sentinel-2 optical, Sentinel-1 SAR amplitude, DEM-derived Terrain Ruggedness Index (TRI), and JAXA land cover data, tackling class imbalance with BCE + Dice loss and providing probability/uncertainty maps via 4-TTA for robust hazard delineation under adverse weather. The principal aim is to enable operational, weather-robust hazard zone extraction with AUC upto 0.89 (best multimodal configuration), outperforming single-modality baselines (e.g., optical-only AUC 0.74; SAR-only 0.69) through synergistic feature fusion, while highlighting multimodal SAR's edge for cloud-obscured scenarios. Validated on Hiroshima Prefecture data—Japan's highest-risk region with ~32,000 hazard spots—this approach demonstrates pre/post-disaster change detection, but reveals limitations in spatial generalization due to region-specific training.**

Keywords—SAR; Optical; ResUNet++; land cover classification; Digital Elevation Model; landslide hazard zones; The Heavy Rain Event of July 2018

I. INTRODUCTION

In recent years, the incidence of sediment-related disasters in Japan has increased significantly, and the resulting damage has become more severe. According to statistics from the Ministry of Land, Infrastructure, Transport and Tourism (MLIT), the annual average number of sediment-related disasters in Japan was 1,064 from 1995 to 2004, but this figure has increased to 1,523 over the most recent 10-year period (2015 to 2024) [1]. The main factors cited for this increase are the frequent occurrence of record-breaking heavy rainfall caused by linear precipitation zones and typhoons. According to observational data published by the Japan Meteorological Agency (JMA) from approximately 1,300 locations nationwide, the annual average number of occurrences of hourly rainfall exceeding 50 mm was 285 from 1995 to 2004. In contrast, the annual average for the most recent 10-year period (2015 to 2024) was 333, also showing an increasing trend [2]. Thus, a strong correlation is suggested between climate change and the increased risk of landslides and debris flows.

To prevent and mitigate damage from such disasters, national and prefectural governments maintain hazard maps. These hazard maps visualize the risk levels not only for landslides and debris flows but also for other natural disasters like tsunamis, storm surges, and floods. They are used to provide residents with evacuation criteria and information on evacuation sites. However, the selection of landslide hazard zones currently relies heavily on on-site surveys conducted by experts. Consequently, the selection process requires extensive time and cost, making it extremely difficult to comprehensively cover large areas and continuously reflect changes in land use, environmental changes due to climate change, and topographic conditions. Furthermore, cases have been reported where landslides occurred outside designated hazard zones, indicating the potential for hazards to occur beyond these areas [3]. Insufficient update frequency thus risks undermining the reliability of hazard map warning zones. To enhance residents' disaster preparedness awareness, it is essential to provide reliable warning zone information updated at near-real-time.

This study contributes: 1) optimal multimodal input analysis showing AUC gains (e.g., +0.15 from optical to optical+DEM+land cover); 2) probability/uncertainty visualization for operational trust; 3) pre/post-2018 Heavy Rain Event application revealing surface change sensitivity; 4) BCE+Dice loss for imbalance handling. Focused on Hiroshima's granite-prone terrain (31,987 hazard spots), it prioritizes methodological validation over broad generalization, with restrained claims pending cross-region tests.

This study aims to extract dynamically changing landslide hazard zones. To achieve this, we propose a hazard zone extraction method that integrates deep learning models with multiple geospatial datasets. Specifically, we use optical imagery or Synthetic Aperture Radar (SAR) amplitude imagery as the base input data. By combining this with a Digital Elevation Model (DEM) and land cover imagery, we quantitatively compare the improvement in extraction accuracy compared to using a single data type alone. Furthermore, we analyze the reliability of the extraction results by visualizing not only a probability map (which displays the model's confidence) as a heatmap, but also an uncertainty map calculated from the variability of the predictions. Additionally, we apply the proposed model to satellite images before and after The Heavy Rain Event of July 2018 to compare changes in the landslide hazard zones.

As mentioned earlier, the occurrence of landslides and debris flows is not only caused by prolonged heavy rain or torrential downpours but also faces increased risk from natural disasters such as earthquakes and volcanic activity. However, it is also pointed out that besides natural disasters, deforestation and changes in land use alter soil stability, increasing the risk of landslides and debris flows [4]. Illegal logging and clear-cutting often occur in remote mountain areas where humans rarely venture. Consequently, detection is delayed, potentially leading to the occurrence of unknown landslides and debris flows. Given this background, utilizing satellite imagery suitable for wide-area observation is effective.

The remainder of this study is organized as follows: Section II reviews related work; Section III details the proposed method and the materials used; Section IV presents the results; and Section V offers concluding remarks.

II. RELATED RESEARCH WORKS AND DIFFERENCES FROM THIS STUDY

A. Related Research Works

Satellite remote sensing has become an indispensable technology for landslide detection and monitoring. Scaioni et al. categorized the applicability of remote sensing technology into three categories: landslide recognition, landslide monitoring, and landslide hazard analysis and prediction, explaining the potential of remote sensing through further advancements [5].

Research on landslide detection is advancing through various approaches. Kanai et al. performed landslide detection from aerial photographs using the deep learning model Mask R-CNN [6]. Qi et al. evaluated the landslide detection accuracy of U-Net and ResUNet models, demonstrating that ResUNet outperforms U-Net in terms of precision, recall, and F1 score [7]. Furthermore, regarding verification focused on versatility, Prakash et al. constructed a convolutional neural network (CNN) training method using composite inventories and proposed an approach applicable to new areas [8]. Liu et al. proposed the Multi-Scale Complex Background Emphasis (MSSCBE) method to address the challenge of landslides being easily confused with complex backgrounds, leading to false detections and missed detections [9].

Research on landslide hazard zone detection using various deep learning models has also been conducted. However, optical imagery alone has limitations in capturing elevation and topographic features. Consequently, research integrating geospatial data is advancing. Liu et al. proposed a semantic segmentation network (FFS-Net) that integrates high-resolution remote sensing images and DEM data [10]. Similarly, Yoshida et al. proposed a decision tree model using explanatory variables derived from optical satellite imagery and DEM [11].

Regarding landslide hazard zones, Ado et al. analyzed articles on machine learning (ML) models used to create landslide hazard maps, reporting a comprehensive review while emphasizing the importance of landslide hazard maps [12]. To mitigate future landslide risks and enhance the applicability of sustainable land use planning, Habumugisha et al. (2022) [13] constructed multiple deep learning models (DNN/LSTM/CNN/RNN) for landslide susceptibility mapping. Similarly, Chen et al. proposed an improved U-Net model addressing the limitation that conventional machine learning models focused solely on individual pixels and failed to consider the surrounding environment (spatial structural information) [14]. By utilizing diverse inputs, including satellite imagery, topographic data, and geological information, they reported improved accuracy and the generation of landslide susceptibility maps.

Several publications provide important insights regarding applications in SAR, and numerous studies have been conducted on disaster detection using SAR data. Mondini et al. published a comprehensive review on SAR imagery for landslide detection, detailing SAR's utility, operational frameworks, and future applications [15]. Furthermore, Mondini et al. used SAR amplitude imagery from the Sentinel-1 satellite to identify landslides by leveraging SAR-specific changes and photo interpretation techniques, aiming for rapid landslide detection [16].

Regarding research utilizing deep learning, Liang et al. proposed a landslide detection method using a morphological model (DLM) for single-polarization SAR images after disasters [17]. Arai et al. proposed an approach to identify non-vegetated areas indicating landslide-prone zones using the Radar Vegetation Index (RVI) derived from Sentinel-1 SAR data and the EfficientNetV2 deep learning model [18].

As a study combining terrain data, Nava et al. proposed an Attention U-Net model utilizing Sentinel-1 satellite SAR amplitude images and slope angles [19]. By utilizing phase information and interferometrically combining phase data from multiple time periods, Interferometric Synthetic Aperture Radar (InSAR) analysis technology can detect displacement. As research applying InSAR technology, Li et al. proposed an MSFD-Net model using displacement images and tilt angles as input data [20].

Numerous studies utilizing InSAR have also been conducted regarding landslide hazard zones. Yuan et al. proposed an approach to extract landslide susceptibility factors and predict landslide susceptibility using InSAR technology by constructing a hybrid model combining a convolutional neural network (CNN) and a recurrent neural network (RNN) strong in handling time-series data [21]. Additionally, Yang et al. proposed an approach to identify potential landslide sites in specific regions by combining InSAR technology with multiple topographic and meteorological datasets [22].

B. Differences from this Study

Much current research focuses primarily on landslide detection and landslide hazard zone extraction after disasters occur. This study aims to extract landslide hazard zones from satellite imagery observed during routine periods, regardless of whether they preceded or followed a disaster. Incorporating changes that increase landslide risks-such as Land Use/Land Cover changes and deforestation—is expected to enhance the reliability of hazard maps and support their revision. Furthermore, to verify its effectiveness in rapidly assessing post-disaster damage and mitigating secondary disaster risks, satellite imagery before and after The Heavy Rain Event of July 2018 is input and compared.

III. PROPOSED METHOD AND MATERIALS

A. Intensive Study Area

This study focuses on Hiroshima Prefecture in Japan. This region is characterized by the steep terrain of the Chugoku Mountains, presenting an extremely high risk of sediment-related disasters. During The Heavy Rain Event of July 2018, the prefecture suffered extensive damage due to numerous slope failures and debris flows. Table I lists the top five prefectures in Japan with the highest number of designated sediment-related disaster hazard zones, based on statistics from the Ministry of Land, Infrastructure, Transport and Tourism [23]. Hiroshima Prefecture recorded approximately 9,000 more hazard zones than Shimane Prefecture, which ranked second, making it the region with the highest landslide risk nationwide. Against this backdrop, Hiroshima Prefecture was selected as the test site for verifying the proposed extraction method.

The reason for focusing solely on Hiroshima Prefecture rather than the whole of Japan lies in its distinct geological characteristics [24]. Granite, widely distributed in Hiroshima Prefecture, is particularly prone to weathering into sand and soil (decomposed granite, locally known as *Masa soil*) due to rain and wind, and has been reported to have a high potential for triggering landslides [25,26].

Fig. 1 shows the geographical scope of the study area, depicting Hiroshima Prefecture and its surrounding regions. The optical satellite imagery (Sentinel-2[1]) and SAR satellite imagery (Sentinel-1[2]) used in this analysis were obtained from the Copernicus Browser[3] within the Copernicus Data Space Ecosystem.

TABLE I. Top 5 Prefectures in Japan with the Highest Number of Landslide Hazard Spots

Rank	Prefecture	Debris flow	Landslide	Steep Slope Failure	Total
1	Hiroshima	9,964	80	21,943	31,987
2	Shimane	8,120	264	13,912	22,296
3	Yamaguchi	7,532	285	14,431	22,248
4	Hyogo	6,912	286	13,550	20,748
5	Oita	5,125	222	14,293	19,640

Fig. 1. Map of the study area, Hiroshima Prefecture, Japan.

B. Data Used

1) Optical imagery: For optical imagery, satellite data from the European Space Agency (ESA) Sentinel-2 mission were acquired. Sentinel-2 provides multispectral data with a spatial resolution of 10 meters, enabling detailed observation of surface spectral reflectance characteristics. Its high revisit frequency (approximately every 5 days) is advantageous for rapid monitoring, although observations are constrained by cloud cover and nighttime conditions. For this study, true-color images (TCI) synthesized from RGB bands were acquired. To minimize cloud effects, images captured during daylight hours on a clear day, August 8, 2024, were obtained. The datasets were created using QGIS[4] by saving images while shifting the location of the target area. Fig. 2 shows a portion of the Sentinel-2 optical image of Hiroshima Prefecture acquired on August 8, 2024, as exported from QGIS.

Fig. 2. Sentinel-2 optical image acquired on August 8, 2024, near Hiroshima City.

2) SAR imagery: For SAR imagery, amplitude (Ground Range Detected, GRD) data from the Sentinel-1 mission operated by the ESA were acquired. Sentinel-1 is an active C-band sensor that transmits microwaves, enabling observation of the Earth's surface regardless of weather conditions or the time of day. However, raw GRD data contain geometric distortions such as radar shadows and speckle noise, particularly in mountainous areas of the study region. To eliminate these distortions and ensure geometric consistency with other geospatial data, preprocessing was performed using the ESA Sentinel Application Platform (SNAP[5]).

The processing steps are as follows:

- Apply Orbit File: Correction of orbital errors using precise orbit ephemerides.
- Thermal Noise Removal: Reduction of additive thermal noise from the sensor.
- Calibration: Conversion of digital pixel values to physical backscatter coefficients.
- Range Doppler Terrain Correction: Orthorectified using Copernicus 30m global DEM to precisely align SAR images with geographic coordinates.

In this study, images were acquired on August 7 and August 19, 2024, to align with the timing of optical data acquisition. Fig. 3 shows a portion of the Sentinel-1 SAR image acquired over Hiroshima Prefecture on August 7, 2024, as exported from QGIS.

---

[1] https://sentinels.copernicus.eu/copernicus/sentinel-2

[2] https://sentinels.copernicus.eu/copernicus/sentinel-1

[3] https://dataspace.copernicus.eu/ecosystem/services/copernicus-browser

[4] https://qgis.org/

[5] https://step.esa.int/main/toolboxes/snap/

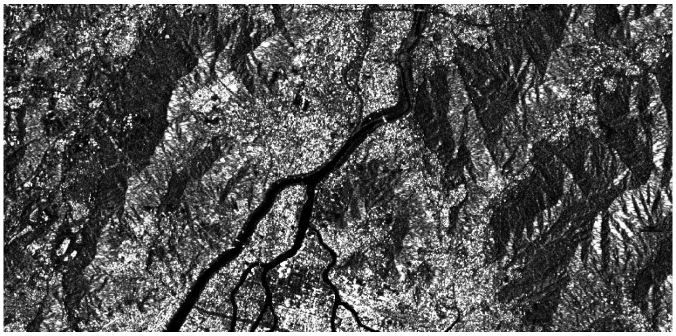


Fig. 3. Sentinel-1 SAR image acquired on August 7, 2024, near Hiroshima City.

3) Digital Elevation Model (DEM)

- Topographic features are key factors influencing landslide hazard potential. We utilized the 5-meter mesh DEM provided by the Geospatial Information Authority of Japan (GSI). To quantify the local terrain complexity correlated with slope instability, we calculated the Terrain Ruggedness Index (TRI). The TRI for a central pixel is defined as the square root of the sum of the squares of the elevation differences between the central pixel and its eight adjacent pixels, as shown in Eq. (1) [27]:

$$TRI = \sqrt{\sum_{i=1}^{8}(X_{center} - X_i)^2} \tag{1}$$

where, X_{Center} represents the elevation of the central pixel, and Xi represents the elevation of the adjacent pixels. TRI was calculated using Raster Terrain Analysis tool available in QGIS. This index functions as a feature that quantitatively represents surface ruggedness.

In this study, 5-meter mesh DEMs (5A) were obtained from the GSI Base Map Information[6] download service. The creation date is June 20, 2025. Under the assumption that this date is close to the acquisition dates of the optical and SAR images, the data were used accordingly. Fig. 4 shows a portion of DEM converted to TRI format for Hiroshima Prefecture, as exported from QGIS.

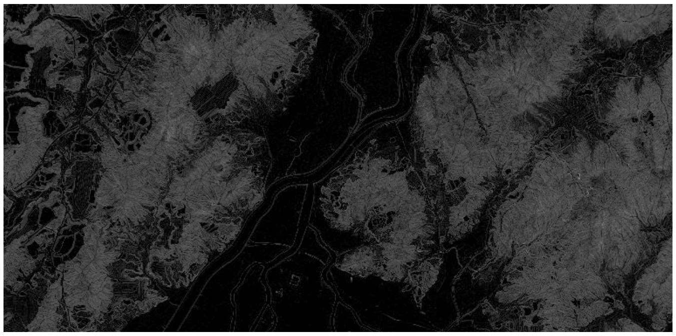


Fig. 4. TRI image derived from the DEM for the area around Hiroshima City.

4) Land Cover Classification Map

- This study utilized the high-resolution Land Use/Land Cover (LULC[7]) map provided by the Japan Aerospace Exploration Agency (JAXA). This dataset features a 10-meter spatial resolution and classifies the land surface into 15 categories (e.g., water bodies, grasslands, and bare land). In the proposed model, these 15 classification categories were converted into one-hot-encoded vectors and treated as independent feature channels.

- These data from 2024 were released in April 2025. As with the DEM, it was assumed that the observation dates of the optical and SAR images were close to the release date. Fig. 5 shows a portion of the LULC map for Hiroshima Prefecture, as exported from QGIS.

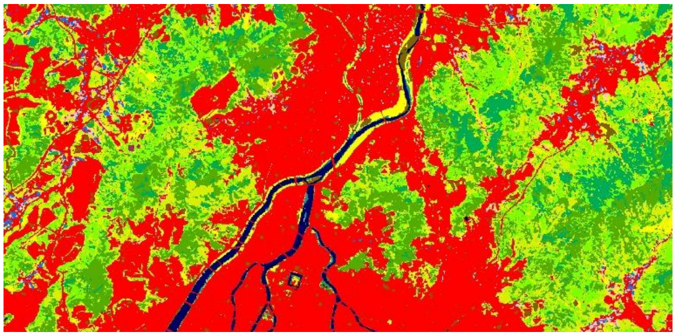


Fig. 5. Land cover classification image for the area around Hiroshima City.

5) Landslide Hazard Zone

- The ground truth data labels used for model training and evaluation were obtained from the "Landslide Hazard Areas (Steep Slope Collapse)" raster data[8] (2023 edition) provided by the Geospatial Information Authority of Japan. These areas are designated under the Sediment-Related Disaster Prevention Act as zones where steep slopes (30 degrees or more) are prone to collapse and pose a danger to residents. Fig. 6 shows a portion of the landslide hazard zones within Hiroshima Prefecture, as exported from QGIS.

- To create correct labels, we performed a binarization process. Designated areas were extracted, and the "landslide danger zone" (yellow area) and the "special landslide danger zone" (red area) were defined as "danger zone" (value = 1) and merged into a single positive class. All other areas were defined as "nondanger zone" (value = 0).

Fig. 6. Image of landslide hazard warning zones around Hiroshima City.

---

[6] https://service.gsi.go.jp/kiban/app/

[7] https://www.eorc.jaxa.jp/ALOS/jp/dataset/lulc/lulc_v2504_j.htm

[8] https://disaportal.gsi.go.jp/index.html

C. Proposed Methodology and Model Architecture

- The methodology consists of data preparation, dataset generation, model training using various input combinations, and comparative evaluation.
- For deep learning dataset generation, the aligned data were cropped into 256×256-pixel patches. These patches were subsequently divided into training, validation, and test datasets. Note that during dataset creation, to capture only the topographic structure and geological features within Hiroshima Prefecture, the model was comprehensively trained by changing the image positions in all areas each time. Therefore, areas included in training may also be present within test cases (though not as identical images). This is because the research aims to identify models that maximize accuracy when combining geospatial data and to establish specialized methods for each prefecture; the input of new areas during testing is not anticipated.
- For the model implementation phase, ResUnet++ architecture was adopted as the architecture. To evaluate the contribution of various environmental factors to landslide detection accuracy, multiple models were implemented using different combinations of input data (e.g., optical data only, SAR + DEM, etc.). Each model was trained individually, and the extraction accuracy was evaluated and compared using the test dataset.
- Finally, based on quantitative results, we conducted a comparative analysis between the highest-performing model and a baseline model that used only sensor data. This analysis involved acquiring sensor data before and after the disaster to detect and compare changes in the estimated landslide hazard zones.
- We employed ResUNet++, a semantic segmentation method, as a model for extracting landslide hazard zones [28]. This ResUNet++ is an architecture that improves upon the standard U-Net and ResUNet. In this study, extraction is performed using multiple channels. Therefore, it is necessary to evaluate the importance of channels while performing detailed extraction of complex boundary regions. For these reasons, ResUNet++ was adopted.

D. Experimental Setup

- Table II shows the model types and input channel counts. The dataset consisted of 2,390 pairs in total, which were randomly split into three subsets: 1,673 pairs (70%) for training, 478 pairs (20%) for validation, and 239 pairs (10%) for testing. Furthermore, for all models, using the same seed value, vertical/horizontal flipping and scaling were applied as identical affine transformations during training.
- We designed eight experimental models with varying input combinations, ranging from single-sensor inputs to fused multimodal inputs. The input image size is standardized to 256×256 pixels for all models. The number of input channels was varied based on a configuration combining optical data (3 channels), DEM (1 channel), SAR (1 channel), and one-hot encoded land cover data (15 channels) in the channel direction.
- Table III shows the main hyperparameters used for model construction. The loss function employs a combination of BCE (binary cross-entropy) and Dice loss, termed "BCE Dice Loss". This combines BCE, which determines positive/negative on a per-pixel basis and tends to stabilize gradients, with Dice loss, which evaluates the overlap between the predicted mask and the ground truth mask and is effective for class-imbalanced images.
- The total loss (L) is defined as follows [see Eq. (2)]:

$$\Lambda = \langle \cdot \Lambda_{\mathrm{BXE}} + ® \cdot \Lambda_{\Delta\mathrm{IXE}} \qquad (2)$$

where, L_{BCE} represents the BCE loss and L_{DICE} represents the Dice loss. In this study, to unify pixel-level accuracy and region-based overlap, equal weights (α=0.5, β=0.5) were assigned to both components.

TABLE II. EXPERIMENTAL MODELS AND INPUT CHANNEL CONFIGURATIONS

Model	Number of Input Channels
(a) Optical	3
(b) Optical + DEM(TRI)	4
(c) Optical + Land Cover	18
(d) Optical + DEM(TRI) + Land Cover	19
(e) SAR	1
(f) SAR + DEM(TRI)	2
(g) SAR + Land Cover	16
(h) SAR + DEM(TRI) + Land Cover	17

TABLE III. HYPERPARAMETERS

Parameter	Value
Image Size	256×256
Batch Size	8
Loss Function	BCE + Dice
Optimizer	Nadam
Learning Rate Scheduler	ReduceLROnPlateau
Initial Learning Rate	1.0×10^{-4}
Patience (LR Decay)	10 epochs
Decay Factor	0.2
Early Stopping	30 epochs

E. Evaluation Metrics

1) Model Performance Metrics

- To quantitatively evaluate the extraction accuracy of the proposed model, we adopted precision, recall, and F1 score as accuracy metrics. These metrics are calculated based on the components of the confusion matrix: true positives (TP), false positives (FP), and false negatives (FN) [see Eq. (3) to Eq. (5)]:

$$\text{Πρεχισιον} = \frac{TP}{TP+FP} \tag{3}$$

$$\text{Ρεχαλλ} = \frac{TP}{TP+FN} \tag{4}$$

$$\text{Φ1} = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall} \tag{5}$$

- In addition to comparing accuracy using a single threshold, we created Precision-Recall (PR) curves showing the relationship between precision and recall as the prediction probability threshold was continuously varied from 0.0 to 1.0. We also calculated the area under the curve (AUC). This enabled the comparison of overall model performance independent of any specific threshold. The AUC is commonly used as a performance evaluation metric in landslide susceptibility mapping [12].

2) Visualization of Probability and Uncertainty Maps

- To quantify the model's confidence for each pixel in the segmentation results, probability maps and uncertainty maps were output and visualized. While typical segmentation results output binary values (0 or 1) per pixel, this segmentation model also outputs pixel-level prediction probabilities during testing. Visualizing these alongside the binary results enabled the evaluation of prediction reliability. For calculating both maps, we adopted an inference method based on four-pattern Test-Time Augmentation (4-TTA). Specifically, we applied four transformation patterns to the input image: no transformation (p0), horizontal flip (p1), vertical flip (p2), and horizontal plus vertical flip (p3). Inference is performed on each transformed image. The average (μ) of the four obtained probability values is calculated as the probability map, and the standard deviation (σ) is calculated as the uncertainty map. The probability map indicates the expected probability for each pixel, while the uncertainty map shows that larger standard deviations indicate greater instability in the model's prediction. Fig. 7 shows an overview of the estimation method. p0 to p3 represent the probability values for each transformed image. This method not only reduces prediction instability associated with single-image inference but also visualizes areas of model uncertainty, making it useful for decision-making in practical applications.

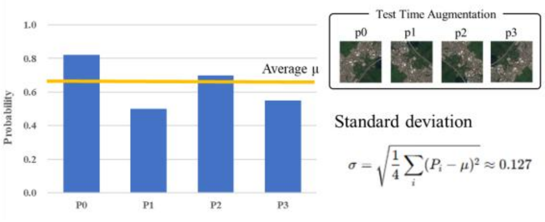


Fig. 7. Overview of inference method using 4-TTA.

IV. RESULTS

A. *Quantitative Evaluation of Extraction Accuracy*

1) *Performance of optical-based models*: Table IV shows the evaluation results for each model based on optical images. Note that the values in Table IV were obtained by searching for a threshold where precision and recall were balanced and then running the model using that threshold. The optical-only model (a) achieved P=0.68, R=0.67, F1=0.67, and AUC=0.74. In contrast, the model incorporating DEM (b) showed significant accuracy improvements: P=0.80, R=0.79, F1=0.79, and AUC=0.88. On the other hand, the model incorporating land cover (c) showed P=0.71, R=0.71, F1=0.71, and AUC=0.79, indicating less improvement than the DEM model. The model incorporating all data (d) achieved the best results across all metrics: P=0.81, R=0.82, F1=0.81, and AUC=0.89, demonstrating the most balanced performance among the metrics.

Fig. 8 shows the PR curves for each optical-based model. In all comparisons, adding input data shifted the entire PR curve upward and to the right, indicating improved accuracy. Adding a DEM contributed significantly to a substantial improvement in AUC (+0.14, from 0.74 to 0.88). In contrast, adding land cover results in a smaller improvement in AUC (+0.05, from 0.74 to 0.79). This can be attributed to the fact that many of the ground surface elements represented by land cover were already captured by the spectral information in the optical images. Nevertheless, since both precision and recall improved, land cover is considered to supplement surface features to a certain extent for areas that are difficult to distinguish using optical images alone.

TABLE IV. RESULTS OF EACH MODEL BASED ON OPTICAL IMAGES

Model	Precision	Recall	F1 score	AUC
(a)	0.68	0.67	0.67	0.74
(b)	0.80	0.79	0.79	0.88
(c)	0.71	0.71	0.71	0.79
(d)	0.81	0.82	0.81	0.89

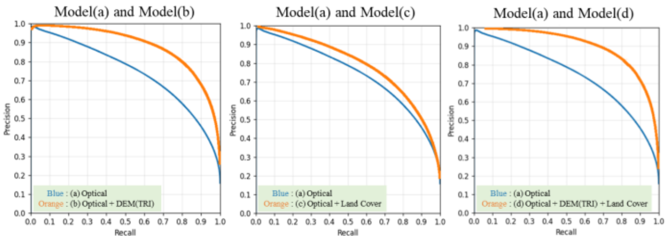


Fig. 8. Comparison of precision–recall (PR) curves for optical-based models with different input configurations.

2) *Performance of SAR-based models*: Table V shows the evaluation results for each model based on SAR images. As in the case of optical imagery, the values in Table V were obtained by searching for a threshold at which precision and recall were balanced and then applying that threshold. For the SAR-only

model (e), P=0.63, R=0.62, F1=0.63, and AUC=0.69, respectively. In contrast, the model incorporating a DEM (f) showed a substantial improvement, with P=0.75, R=0.73, F1=0.74, and AUC=0.82. The model incorporating land cover (g) also achieved P=0.74, R=0.73, F1=0.73, and AUC=0.81, indicating an improvement comparable to model (f). The model using all inputs (h) achieved the best performance, with P=0.81, R=0.80, F1=0.81, and AUC=0.89, matching the accuracy of the best optical-based model (d). This model also exhibited the most favorable balance among the metrics.

Fig. 9 shows the PR curves for the SAR-based models. As with the optical imagery case, adding input data led to improved performance across the entire curve. In addition, for SAR-based models, unlike the optical imagery case, the inclusion of land-cover data contributed to a substantial improvement in accuracy comparable to that achieved by adding DEM. This was likely because discriminating surface types solely based on microwave backscattering intensity is challenging, as diverse land covers can exhibit similar backscatter signatures depending on surface roughness and dielectric properties. Therefore, the inclusion of land cover data provided essential categorical context, resolving these ambiguities and substantially complementing the SAR-derived spatial features.

TABLE V. RESULTS OF EACH MODEL BASED ON SAR IMAGES

Model	Precision	Recall	F1 score	AUC
(e)	0.63	0.62	0.63	0.69
(f)	0.75	0.73	0.74	0.82
(g)	0.74	0.73	0.73	0.81
(h)	0.81	0.80	0.81	0.89

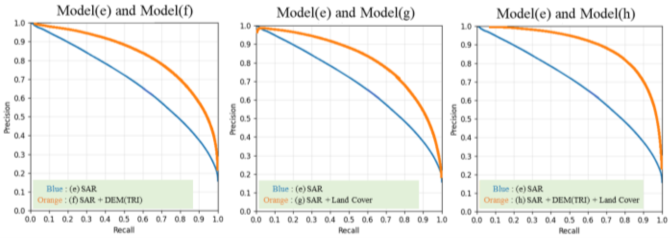


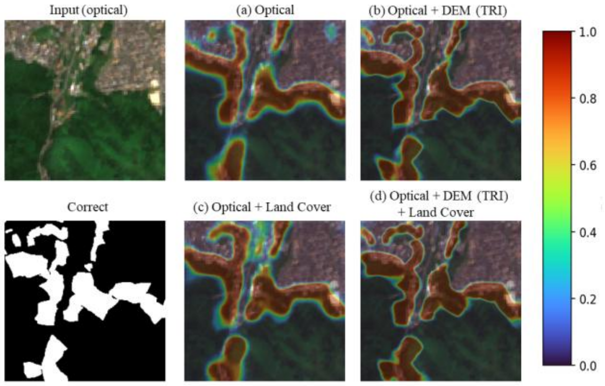
Fig. 9. Comparison of precision–recall (PR) curves for SAR-based models with different input configurations.

B. Visualization of Probability Maps and Uncertainty Maps

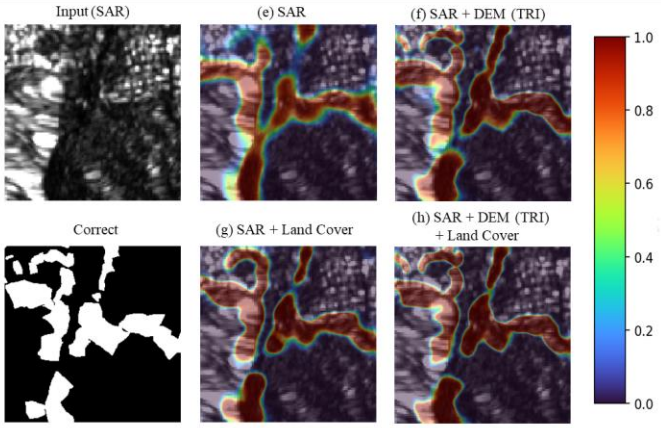
The output results of the probability maps and uncertainty maps are shown in Fig. 10. Comparing the results across all models, a consistent tendency was observed in both categories: high confidence in the central parts of the predicted areas and increased uncertainty along the boundaries. In the probability maps, the expansion of the red regions indicated higher reliability, whereas in the uncertainty maps, a reduction in edge width indicated less ambiguity in the prediction.

For the single-sensor models (a) and (e), many locations exhibited low confidence even in the central parts of the extracted areas, and the area of high-uncertainty edges along the boundaries became larger, indicating greater model instability. In contrast, the models incorporating DEM (b) and (f) showed a substantial improvement in confidence. In the models incorporating land cover (c) and (g), false detections decreased, particularly in urban areas and in locations away from slopes where sediment-related disasters are less likely to occur. In the models incorporating both DEM and land cover (d) and (h), high confidence in the central parts was maintained while the area of high-uncertainty edges near the boundaries became smaller.

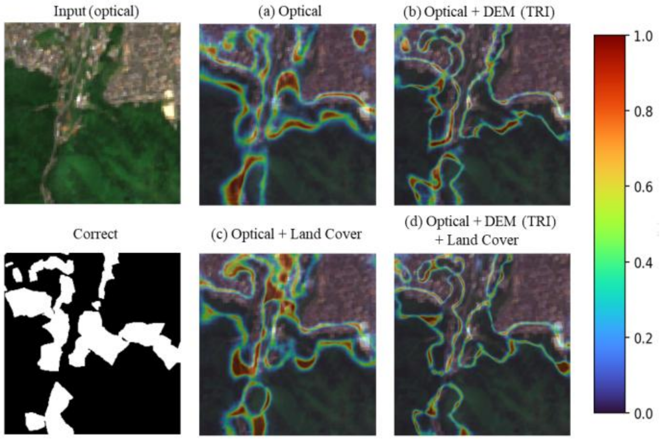
- Thus, presenting deep learning outputs not only as binary results (positive/negative) but also simultaneously as probability maps and uncertainty maps can facilitate an intuitive understanding of the situation for specific areas and can be utilized for rapid and accurate decision-making.



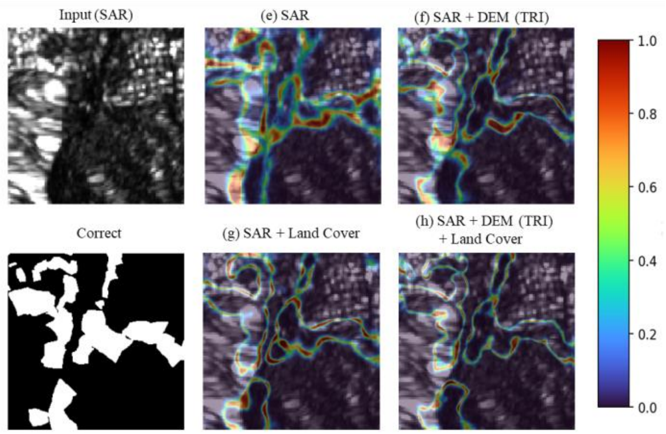
(a) Probability map of the optical-based model.



(b) Probability map of the SAR-based model.



(c) Uncertainty map of the optical-based model.



(d) Uncertainty map of the SAR-based model.

Fig. 10. Probability maps and uncertainty maps.

C. Comparison of Estimation Results before and after Disaster Events

1) Overview of the disaster: Model (d) (optical + DEM + land cover) achieved the highest accuracy for extracting sediment-disaster hazard zones. In this section, we evaluated the practical applicability of this model by applying it to optical images acquired before and after an actual disaster event and analyzing the resulting changes in the probability maps. For comparison, the same analysis was also conducted using the baseline model (a) (optical only). In this study, Nakahata, Yasuura Town, Kure City, Hiroshima Prefecture, which suffered severe damage during The Heavy Rain Event of July 2018 in western Japan, was selected as the validation area. In this district, heavy rainfall triggered sediment-related disasters in multiple mountainous locations, causing road disruptions to settlements and extensive building damage across the area. Fig. 11 shows a photograph of the damage in the nearby Ichihara district[9]. In addition, Fig. 12 shows the optical satellite images acquired for the study area (Nakahata district) before the disaster (June 1, 2018) and after the disaster (July 16, 2018).

Fig. 11. Ichihara district, Kure City

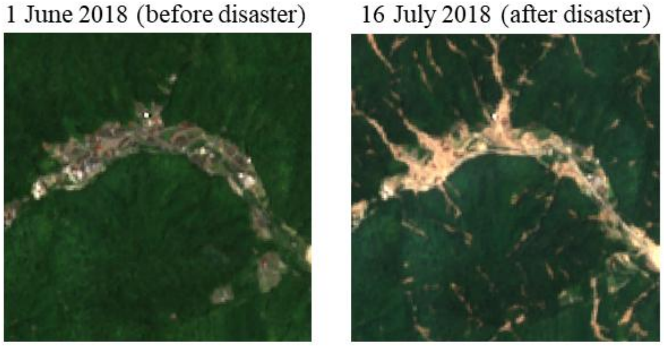


Fig. 12. Optical satellite images of Nakahata, Yasuura Town.

2) Analysis of probability map change: Fig. 13 shows the probability maps produced before and after the disaster, and Fig. 14 shows the difference between the before-disaster and after-disaster probability maps. In the after-disaster probability map (July 16, 2018), both models reflected changes in surface conditions caused by heavy rainfall, with the predicted areas concentrated mainly near the boundary between mountainous and lowland areas. However, the spatial distributions differed clearly between the models.

In model (a), the high-probability regions were spread over a wide area, and sporadic responses were also observed on surrounding slopes. Because the input relied only on spectral information, distinguishing the presence or absence of sediment-related disasters became difficult; as a result, the model responded to apparent changes unrelated to the disaster, leading to unstable extraction. In contrast, in model (d), the high-probability regions were concentrated mainly along boundary areas consistent with the disaster-affected zone, and unnecessary extraction on surrounding slopes is suppressed. This can be interpreted as the model not treating all apparent after-rainfall changes as disasters but instead narrowing the candidate areas by considering topography and land cover information. The ability to output only high-confidence regions is beneficial for practical operations such as rescue activities and risk assessment of secondary disasters.

- Because the DEM and land cover data used in this study are static datasets updated every few years, the same pre-disaster datasets were used for both the before-disaster and after-disaster analyses in this experiment. However, because actual landslides can cause substantial topographic changes, achieving more dynamic and accurate extraction requires topographic and land cover data that reflect conditions immediately after the disaster. In recent years, studies have advanced in directly estimating up-to-date land cover and geospatial data from satellite imagery using deep learning [29,30]. By integrating these techniques into the proposed method, further improvements in the timeliness and accuracy of landslide hazard extraction are expected.
- Prob/uncertainty maps show high center confidence, edge uncertainty reduced by fusion. Pre/post-2018: multimodal pinpoints changes vs. optical scatter. No baselines (e.g., DeepLabV3+, RF on slope/cover) or disjoint tests yet.

---

[9] https://www.city.kure.lg.jp/site/kure-city-reconstruction-memorial-museum/kure-city-reconstruction-memorial-museum-higaijokyo.html

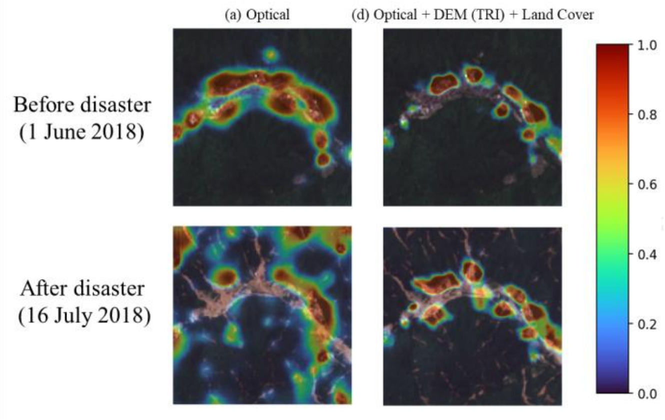


Fig. 13. Comparison of probability maps before and after the disaster.

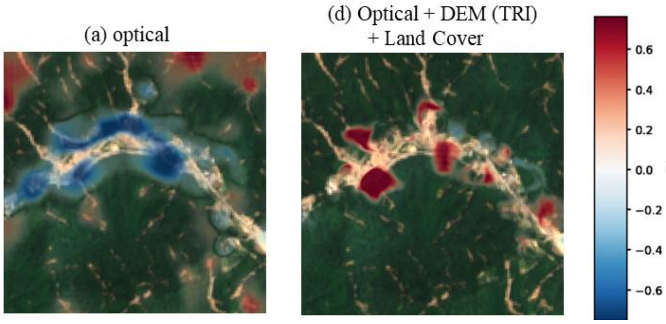


Fig. 14. Differential probability maps (before and after the disaster) visualizing the changes in the probability map.

V. CONCLUSION

We proposed a deep learning model for extracting sediment disaster hazard zones by integrating satellite imagery and geospatial data. Models were developed by combining diverse information sources, including optical imagery, SAR imagery, DEM, and land cover maps. The accuracy comparison results confirmed that integrating geospatial data substantially improved extraction performance compared with models using a single data source. The best-performing configuration was "optical/SAR + DEM + land cover", achieving an AUC of 0.89. The effectiveness of DEM and land cover was also demonstrated. Because optical and SAR imagery had complementary characteristics that compensated for each other's limitations, obtaining high extraction accuracy with both image types suggested that the proposed approach is highly generalizable and is not constrained by weather conditions or day–night differences. In addition, the case study of The Heavy Rain Event of July 2018 indicated that inputting before-disaster and after-disaster optical imagery may enable the detection of changes in hazard zones.

Multimodal fusion boosts robustness, SAR aiding clouds, but Hiroshima training limits broader claims—no cross-prefecture validation. Predicts susceptibility (pre-event), post-disturbance (change detection); preprocessing detailed, but normalization/seasonality assumed stable.

In conclusion, this advances hazard mapping via ResUNet multimodal fusion (SAR/optical/DEM/LULC, BCE+Dice, uncertainty viz), achieving F1=0.81/AUC=0.89 on Hiroshima data, enabling dynamic updates vs. surveys. Scientific value: quantifies input contributions, operational potential for weather-hit areas. As for the limitations, region-specific (no spatial generalization); random split risks leakage; lacks ML baselines, modern nets (Attention U-Net, DeepLabV3+), CIs; speculative robustness sans scenario tests.

FUTURE RESEARCH WORKS

We plan to develop a framework to generate land cover and DEM products from satellite imagery. Although static DEM and land cover maps are currently used, integrating a mechanism that directly derives these geospatial datasets from up-to-date satellite imagery would enable more dynamic extraction of sediment disaster hazard zones by immediately reflecting changes. In addition, we also plan to incorporate new inputs, such as geological conditions within the target area and real-time rainfall distribution, which are expected to advance the model toward more reliable extraction. Furthermore, we plan to use higher-resolution optical and SAR data to achieve further improvements in accuracy. Namely, spatially disjoint validation/cross-events have to be investigated with baselines (RF/XGBoost, DeepLab); dynamic DEM/LULC generation; geology/rain inputs; higher-res data.

ACKNOWLEDGMENT

The author would like to thank Prof. Dr. Osamu Fukuda of Saga University for his valuable comments and suggestions.

GENERATIVE AI USAGE STATEMENT

The authors declare that the generative AI tool Gemini 3.0 was used solely for language support and proofreading purposes, including improving grammar, wording, and overall clarity of the manuscript.

The development of scientific ideas, methodologies, analyses, results, and conclusions was carried out entirely by the human authors. All AI-assisted text was carefully reviewed and edited by the authors, who take full responsibility for the accuracy, originality, and integrity of the manuscript.

REFERENCES

[1] Ministry of Land, Infrastructure, Transport and Tourism, Japan. Landslide disasters in 2024 exceeded the long-term average (since statistics began)—Number of landslide disasters in 2024 released (Press Release, 27 January 2025). Available: https://www.mlit.go.jp/report/press/sabo02_hh_000156.html (accessed on January 14, 2026).

[2] Japan Meteorological Agency. Historical changes in extreme events such as heavy rainfall and extremely hot days (Climate Change Portal). Available: https://www.data.jma.go.jp/cpdinfo/extreme/extreme_p.html (accessed on January 14, 2026).

[3] T. Nishikawa, "Characteristics of sediment runoff occurring at locations not designated as sediment disaster warning areas during the July 2018 heavy rain," Suiri Kagaku (Water Science), vol. 64, no. 2, pp. 43–59, 2020.

[4] H. Saito, W. Murakami, H. Daimaru, and T. Oguchi, "Effect of forest clear-cutting on landslide occurrences: Analysis of rainfall thresholds at Mt. Ichifusa, Japan," Geomorphology, vol. 276, pp. 1–7, 2017.

[5] M. Scaioni, L. Longoni, V. Melillo, and M. Papini, "Remote sensing for landslide investigations: An overview of recent achievements and perspectives," Remote Sensing, vol. 6, no. 10, pp. 9600–9652, 2014.

[6] K. Kanai, S. Kubo, T. Yamane, and P.-J. Chun, “A method for automatic detection of landslide areas from aerial photographs using Mask R-CNN,” Artificial Intelligence and Data Science, vol. 2, pp. 223–231, 2021.

[7] W. Qi, M. Wei, W. Yang, C. Xu, and C. Ma, “Automatic mapping of landslides by the ResU-Net,” Remote Sensing, vol. 12, no. 15, p. 2487, 2020.

[8] N. Prakash, A. Manconi, and S. Loew, “A new strategy to map landslides with a generalized convolutional neural network,” Scientific Reports, vol. 11, p. 9722, 2021.

[9] X. Liu, L. Xu, and J. Zhang, “Landslide detection with Mask R-CNN using complex background enhancement based on multi-scale samples,” Geomatics, Natural Hazards and Risk, vol. 15, no. 1, p. 2300823, 2024.

[10] X. Liu et al., “Feature-fusion segmentation network for landslide detection using high-resolution remote sensing images and digital elevation model data,” IEEE Transactions on Geoscience and Remote Sensing, vol. 61, p. 4500314, 2023.

[11] K. T. Yoshida, D. Nakayama, and H. Matsuyama, "Construction of landslide detection model using optical satellite images and DEMs by machine learning with multi-instance decision tree model," Journal of the Remote Sensing Society of Japan, vol. 44, no. 4, pp. 317–334, 2025.

[12] M. Ado et al., “Landslide susceptibility mapping using machine learning: A literature survey,” Remote Sensing, vol. 14, no. 13, p. 3029, 2022.

[13] J. M. Habumugisha et al., “Landslide susceptibility mapping with deep learning algorithms,” Sustainability, vol. 14, no. 3, p. 1734, 2022.

[14] Y. Chen et al., “Mapping post-earthquake landslide susceptibility: A U-Net like approach,” Remote Sensing, vol. 12, no. 17, p. 2767, 2020.

[15] A. C. Mondini et al., “Landslide failure detection and mapping using synthetic aperture radar: Past, present and future,” Earth-Science Reviews, vol. 216, p. 103574, 2021.

[16] A. C. Mondini et al., “Sentinel-1 SAR amplitude imagery for rapid landslide detection,” Remote Sensing, vol. 11, no. 7, p. 760, 2019.

[17] R. Liang et al., “Landslide mapping from post-event single-temporal polarimetric SAR image by a deep learning method exploiting a morphological model,” Remote Sensing of Environment, vol. 328, p. 114904, 2025.

[18] K. Arai, Y. Nakaoka, and H. Okumura, “Method for landslide area detection with RVI data which indicates base soil areas changed from vegetated areas,” Remote Sensing, vol. 17, p. 628, 2025.

[19] L. Nava, K. Bhuyan, S. R. Meena, O. Monserrat, and F. Catani, “Rapid mapping of landslides on SAR data by attention U-Net,” Remote Sensing, vol. 14, p. 1449, 2022.

[20] N. Li et al., “A deep-learning-based algorithm for landslide detection over wide areas using InSAR images considering topographic features,” Sensors, vol. 24, no. 14, p. 4583, 2024.

[21] R. Yuan and J. Chen, “A hybrid deep learning method for landslide susceptibility analysis with the application of InSAR data,” Natural Hazards, vol. 114, no. 2, pp. 1393–1426, 2022.

[22] X. Yang, D. Chen, Y. Dong, Y. Xue, and K. Qin, “Identification of potential landslide in Jianzha county based on InSAR and deep learning,” Scientific Reports, vol. 14, p. 21346, 2024.

[23] Ministry of Land, Infrastructure, Transport and Tourism (MLIT), Water and Disaster Management Bureau, Sabo Department. Sediment Disaster Hazard Locations by Prefecture. Available: https://www.mlit.go.jp/river/sabo/link20.htm (accessed on January 14, 2026).

[24] K. Araiba, T. Nozaki, B.-P. Jeong, and Y. Fukumoto, “Distribution of landslides and their geological characteristics in Japan—Analysis report based on the data of nation-wide designated landslide-threatened areas—,” J. Japan Landslide Soc., vol. 44, no. 5, pp. 318–323, 2008.

[25] T. Watakabe and Y. Matsushi, “Lithological controls on hydrological processes that trigger shallow landslides: Observations from granite and hornfels hillslopes in Hiroshima, Japan,” Catena, vol. 180, pp. 55–68, 2019.

[26] S. Doshida, “The feature of the landslides by the Hiroshima heavy rain on August 20, 2014,” J. Jpn. Soc. Saf. Eng., vol. 57, no. 1, pp. 22–27, 2018.

[27] S. J. Riley, S. D. DeGloria, and R. Elliot, “A terrain ruggedness index that quantifies topographic heterogeneity,” Intermountain J. Sciences, vol. 5, no. 1–4, pp. 23–27, 1999.

[28] D. Jha et al., “ResUNet++: An advanced architecture for medical image segmentation,” in Proc. 2019 IEEE Int. Symp. Multimedia (ISM), pp. 225–230.

[29] C. F. Brown et al., “Dynamic World, near real-time global 10 m land use land cover mapping,” Sci. Data, vol. 9, p. 251, 2022.

[30] J. Gao, J. Liu, and S. Ji, “A general deep learning based framework for 3D reconstruction from multi-view stereo satellite images,” ISPRS J. Photogramm. Remote Sens., vol. 195, pp. 446–461, 2023.

AUTHOR’S PROFILE

Kohei Arai, He received BS, MS and PhD degrees in 1972, 1974 and 1982, respectively. He was with The Institute for Industrial Science and Technology of the University of Tokyo from April 1974 to December 1978, and was also with the National Space Development Agency of Japan from January 1979 to March 1990. From 1985 to 1987, he was with the Canada Centre for Remote Sensing as a Post-Doctoral Fellow of the National Science and Engineering Research Council of Canada. He was a professor of Saga University from 1990 to 2017. Also, he was a councilor for the Aeronautics and Space related to the Technology Committee of the Ministry of Science and Technology from 1998 to 2000 and was a councilor of Saga University for 2002 and 2003. He also was an executive councilor for the Remote Sensing Society of Japan for 2003 to 2005. He was the Vice Chairman of the Scientific Committee was an Award Committee member of ICSU/COSPAR during. He is a Science Council of Japan Special Member (COSPAR Committee) since 2012. He is an Adjunct Professor of Nishi-Kyushu University and Kurume Institute of Technology as well as Prishtina International University. He wrote 134 books and published 759 journal papers as well as 591 conference papers. He received 66 of awards, including the ICSU/COSPAR Vikram Sarabhai Medal in 2016, and the Science award of the Ministry of Mister of Education of Japan in 2015. He is now Editor-in-Chief of IJACSA. http://teagis.ip.is.saga-u.ac.jp/index.html