HR-GLDD: A globally distributed dataset using generalized DL for rapid landslide mapping on HR satellite imagery

Sansar Raj Meena [1], Lorenzo Nava [1], Kushanav Bhuyan [1], Silvia Puliero [1], Lucas Pedrosa Soares [2], Helen Cristina Dias [2], Mario Floris [1], Filippo Catani [1*]

1. Machine Intelligence and Slope Stability Laboratory, Department of Geosciences, University of Padova, 35129 Padua, Italy
2. Institute of Energy and Environment, University of São Paulo, São Paulo 05508-010, Brazil

* Correspondence: sansarraj.meena@unipd.it

Abstract:

Multiple landslide events occur often across the world which have the potential to cause significant harm to both human life and property. Although a substantial amount of research has been conducted to address mapping of landslides using Earth Observation (EO) data, several gaps and uncertainties remain when developing models to be operational at the global scale. The lack of a high resolution globally distributed and event-diverse dataset for landslide segmentation poses a challenge in developing machine learning models that can accurately and robustly detect landslides in various regions, as the limited representation of landslide and background classes can result in poor generalization performance of the models. To address this issue, we present the high-resolution global landslide detector database (HR-GLDD), a high resolution (HR) dataset for landslide mapping composed of landslide instances from ten different physiographical regions globally: South and South-East Asia, East Asia, South America, and Central America. The dataset contains five rainfall triggered and five earthquake-triggered multiple landslide events that occurred in varying geomorphological and topographical regions. HR-GLDD is one of the first dataset for landslide detection generated by high resolution satellite imagery which can be useful for applications in artificial intelligence for landslide segmentation and detection studies. Five state of the art deep learning models were used to test the transferability and robustness of the HR-GLDD. Moreover, two recent landslide events were used for testing the performance and usability of the dataset to comment on the detection of newly occurring significant landslide events. The deep learning models showed similar results for testing the HR-GLDD in individual test sites thereby indicating the robustness of the dataset for such purposes. The HR-GLDD can be accessed open access and it has the potential to calibrate and develop models to produce reliable inventories using high resolution satellite imagery after the occurrence of new significant landslide events. The HR-GLDD will be updated regularly by integrating data from new landslide events.

1. Introduction

With the increasing impacts of climate change, increased urbanization, and anthropogenic pressure in recent years, the risk from hazards to population, infrastructure, and essential life services has exacerbated. Landslides are quite ubiquitous and account for approximately 4.9% of all the natural disasters and 1.3% of the fatalities in the world (EM-DAT, 2018)~~(EM-DAT, 2018)~~. Induced by natural (earthquakes, volcanic eruptions, meteorological events) and anthropogenic triggers (slope modifications, mining, landscape engineering), the increase in the stress of slope materials causes landslides, which can harm numerous elements at risk. Landslides occur heterogeneously in many parts of the world including the Central and South Americas, the Caribbean islands, Asia, Turkey, European Alps, and East Africa (Froude &

Petley, 2018)~~(Froude & Petley, 2018)~~. In the past 15 years, we have seen a high number of events that have inadvertently led to the failure of thousands of slopes and causing damage to essential linear infrastructures and population. Some recent examples are Wenchuan, China (2008), Kedarnath, India (2013), Kaikoura, New Zealand (2016), Jiuzhaigou, China (2017), Dominica (2017), Porgera, Papua New Guinea (2018), Hokkaido, Japan (2018), Belluno, Italy (2018), Haiti (2021), Sumatra, Indonesia (2022).

These examples indicate that landslide occurrences will probably continue to increase in the short and medium term; therefore, an effective capability of rapid mapping is required to map future event-based landslides~~ and reduce losses~~. In recent years, state-of-the-art research has been conducted to better understand the impact of natural hazards such as landslides and the cascading effects on the elements-at-risk. A critical understanding of these complex processes begins with the onset of mapping slope failures. This information about the failed slopes is attributed as records and is documented in a “landslide inventory”. Landslide inventories include information on the spatial location and extent of the landslides and, if available, also crucial information about 1) the time of occurrence, 2) the triggering event that led slopes to fail, 3) the typology of the landslides based on the accepted standard classifications like (Cruden & Varnes, 1996)~~(Cruden & Varnes, 1996)~~ and (Hungr et al., 2014)~~(Hungr et al., 2014)~~, and 4) the volume of the failure. However, regarding rapid mapping of recently occurred landslides, information about the spatial location, distribution, and intersection with affected elements-at-risk are important. , and 4) the volume of the failure. However, regarding rapid mapping of recently occurred landslides, information about the spatial location, distribution, and intersection with affected elements-at-risk are important.

When it comes to detecting and mapping landslides over remotely sensed images, it is safe to say that a lot of the current literature in the past couple of years has devised and spent time employing artificial intelligence (AI) models to map landslides automatically, arguably, with good results. These AI models can classify remote sensing images to denote where the landslides are present in the analysed images. However, the core prerequisite for employing AI models is a reliable dataset to be used for training. Recent studies have only focused on mapping landslides with AI but at scales that are small or regional while also claiming that the proposed models can cater towards rapid mapping of landslides at any given time, location and scale (Liu et al., 2022; Meena et al., 2022a; Nava, Bhuyan, et al., 2022; Nava, Monserrat, et al., 2022; Soares et al., 2022a; Tang et al., 2022; Yang et al., 2022; Yang & Xu, 2022)~~(Liu et al., 2022; Meena et al., 2022a; Nava, Bhuyan, et al., 2022; Nava, Monserrat, et al., 2022; Soares et al., 2022a; Tang et al., 2022; Yang et al., 2022; Yang & Xu, 2022)~~. However, seldom has been the case where truly an approach has been taken to map landslides outside the regions where the models are initially trained on, and also towards actually applying the proposed models in capturing and mapping event-based landslides that has recently occurred. Some ~~recent ~~other works at collectively detecting and mapping landslides of different countries have been attempted by (Prakash et al., 2021)~~(Prakash et al., 2021)~~ and (Ghorbanzadeh et al., 2022)~~(Ghorbanzadeh et al., 2022)~~, which showcases the power of employing AI at mapping landslides. Recently, Bhuyan et al. (2023) made some strides at mapping landslides at larger spatiotemporal scales to provide multi-temporal inventories of some famous events but more experiments in to explore other geographical contexts are required. ~~However~~T~~, t~~he core of the~~se~~ mentioned studies also heavily relies on the availability of quantity and quality data for training an AI model. The accessibility of such data can 1) allow a model to identify landslides that were caused by different types of triggers (logically leading to the detection of different types of landslides), 2) to map landslides in different parts of the world that vary geomorphologically, and 3) the applicability of the model at mapping newly occurring landslides triggered by events in recent times. The contemporary works of the current literature brings about a critical discussion about the availability and accessibility of comprehensive and adequate data to effectively train models to detect landslides. Both (Prakash et al., 2021)~~(Prakash et al., 2021)~~ and (Ghorbanzadeh et al., 2022)~~(Ghorbanzadeh et al., 2022)~~ have used open-source Sentinel-2 imageries for multi-site landslide detection however, considering the fact that the spatial resolution is 10 metres, a lot of small landslides are missed out or not accurately captured (Meena et al., 2022b)~~(Meena et al., 2022b)~~. The latter ~~is created by~~ sampl~~ing~~ed data from 4 different areas/events Sentinel-2 imagery~~attempted to design a benchmark data set for landslide model training using moderate resolution sentinel-2 data~~ (four bands at 10 meters spatial resolution, six at 20, and three at 60) and combined it with DEM derived data from ALOS-PALSAR. The dataset we propose, instead, is sampled from 10 different areas/events and uses 3 meters spatial resolution imagery. Sampling from more areas can provide a more diverse representation of both landslide and background classes, which can improve the robustness of the model when applied to different regions. Moreover, a dataset with more diversity is likely to generalize better to new unseen data than one with limited diversity, making it more suitable for real-world deployment. Sampling from 10 areas also provides better coverage of the geographical region, reducing the risk of missing important features or patterns. Higher spatial resolution imagery captures more detail, allowing for more accurate identification and segmentation of landslide features. It also allows ~~to~~ obtaining a more detailed view, which can be useful to identify small landslides or details that may be difficult to see in lower resolution imagery. Moreover, it can provide more context for the location, helping to better understand the environment and the relationships between different objects and features. Therefore, the increased detail can result in improved accuracy when classifying features and objects, reducing the risk of misclassification.~~However, they only considered four study areas while training the model on 25% of area and testing the model performances in the remaining 75% area. Furthermore, they had varying non-uniform results for each of the models trained on the dataset. This showcases that a quality dataset is still not available where different models can give consistent results across the board.~~

To effectively and rapidly map landslides after an event, it is required first to determine the spatial extent of the affected areas. Collecting this data is frequently hazardous since it involves individuals on the ground investigating landslides first hand during or immediately after the event. With the increased availability of satellite imagery, this task has the potential to be completed not only remotely but also automatically through the use of powerful deep learning algorithms. Currently, adequate high-resolution satellite imagery of landslides is not widely available. To depict the complex and dynamic nature of the landslides, significant amounts of images must be provided. To ~~that~~ this purpose, we present high-resolution global landslide detector database (HR-GLDD)~~High resolution Global landslide dataset (HR-GLDD)~~, a large-scale satellite image dataset with ~~produced~~ assembled landslide inventories. The database currently houses 10 geographical areas and 3 recently transpired events (see Figure 1), and we plan to constantly update this database with newer events.

2. Study areas

The study areas were chosen based on the variety of triggering events that resulted in the occurrence of the landslides. Because of the availability of VHR archived Planet Scope imageries after 2016, the most significant landslide events were considered. The geomorphological diversity of the study sites results in a collection of complex landslide phenomenon. We selected the imageries based on the availability of cloud-free conditions in the areas and examined globally archived satellite remote sensing imageries from Planet Scope from the years between 2017 and 2022 (Table 1). We selected 8 study sites across the globe to assemble the database (see figure1). To further test the generalization capabilities of the models trained on the proposed dataset, we choose ~~two~~ three recently occurred events: co-seismic landslides in Haiti (August, ~~–~~2021) and rainfall-induced landslides in Indonesia (February, 2022) and Democratic Republic of Congo (April, 2020).

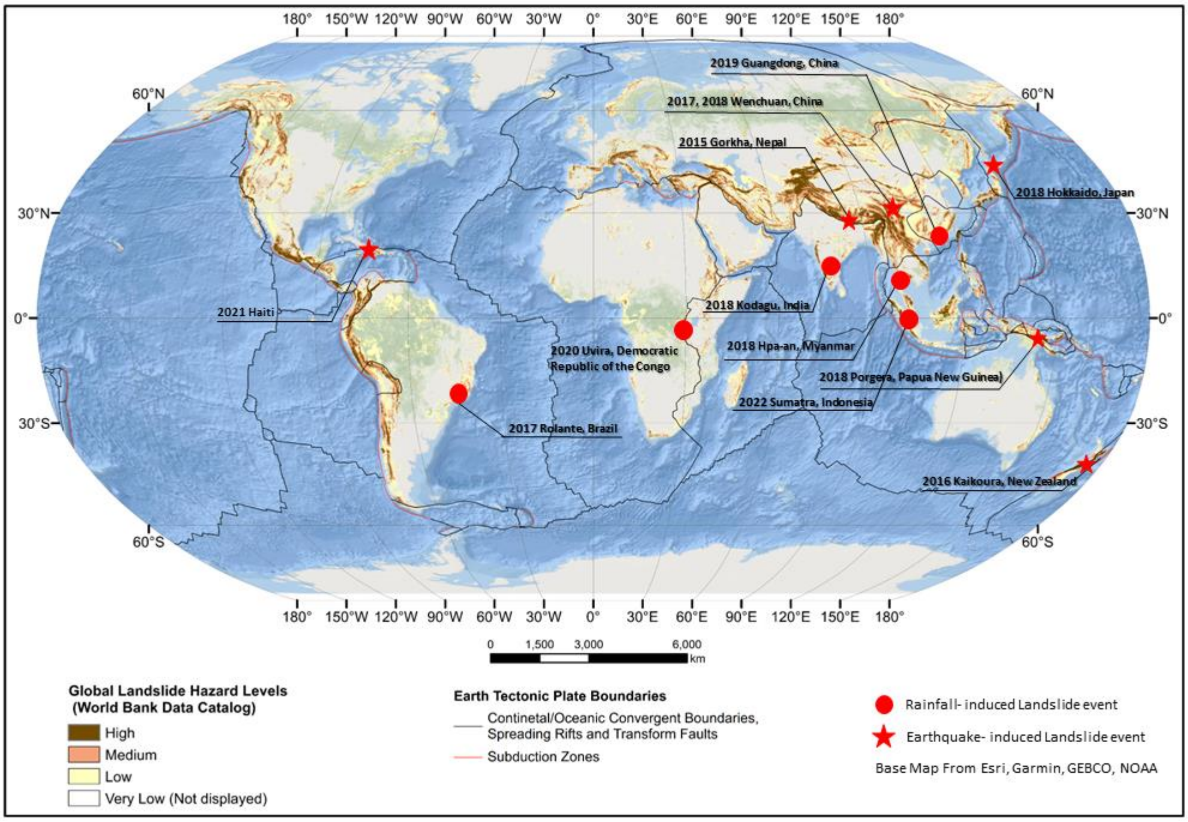


Figure 1: Collection of ~~R~~rainfall- and earthquake~~-~~-induced landslide events present in the HR-GLDD.

~~2.1.~~ ~~Study areas~~

~~2.1.1.~~2.1. Papua New Guinea

Papua New Guinea (PNG) is located on the Australian continent and is the eastern half sector of the New Guinea island. The region is characterized by active volcanos, earthquakes, elevations up to ~4.400 m.a.s.l., steep slopes and is part of the 'Ring of Fire' in the Pacific Ocean. Regarding the tectonic and geological elements, the island can be divided into four tectonic belts: Stable platform, Fold Belt, Mobile Belt, and Papuan Fold and Thrust Belt (Tanyaş et al. 2022). The east sector, where PNG is located has the presence of accreted Paleozoic structure of Tasman Orogen (Hill and Hall, 2003). Due to these conditions, the area is frequently affected by landslides associated with the occurrence of earthquakes (Tanyaş et al. 2022). On February 25, 2018, in the southern area of the Papuan Fold and Thrust belt (central highlands of PNG), a severe earthquake occurred, the magnitude hit Mw 7.5. The event was responsible for damage to buildings, and energy structures besides triggering a high number of landslides (Wang et al. 2020). Around 11,600 landslide scars were registered, and more than half had 50,000 m² (Tanyaş et al. 2022), according to Wang et al. 2020, the earthquake hit the highest magnitude in the region in the past 100 years.

~~2.1.2.~~2.2. Kodagu, India

Kodagu district is located in the Karnataka state, Western Ghats, India. The area is characterized by elevations approximately between 50 and 1.750 m a.s.l., metamorphic rocks (e.g., amphibolite, gneiss, and schist), steep slopes, high annual precipitation of about 4000 mm, and the presence of croplands (e.g., coffee, rice, and spices) (Jennifer and Saravan, 2020; Meena et al. 2021). In August 2018, a rainfall-induced high magnitude mass movement event occurred in Kodagu, the primary landslide type triggered was debris flow (Meena et al. 2021). A total of 343 landslides were recorded, including mudflows, rock falls, and debris flows (Meena et al. 2021). The event resulted in several damages to land resources, properties, and loss of human lives (Martha et al. 2018; Jennifer and Saravan, 2020).

~~2.1.3.~~2.3. Rolante, Brazil

The Rolante river catchment study area is located in the Rio Grande do Sul state, southern Brazil. The region being part of the Serra Geral geomorphological unit, has elevations up to ~1.000 m.a.s.l. (Uehara et al. 2020). Moreover, is characterized by the presence of basaltic rocks and sandstones, and annual precipitation thresholds between 1700 and 2000 mm (Uehara et al. 2020, Soares et al. 2022). On 5 January 2017, a high magnitude rainfall-induced mass movement event was triggered, and 308 landslides were registered (Gameiro et al. 2019; Quevedo et al. 2019), resulting in several damages to the Rolante municipality.

~~2.1.4.~~2.4. Tiburon Peninsula, Haiti

The Tiburon Peninsula study area is located in the western part of the Hispaniola island (Haiti) with elevation up to 2300 m. a.s.l. Tiburon Peninsula, mainly consists of volcanic rocks such as basalts and sedimentary rocks, namely limestones (Harp et al., 2016)~~(Harp et al., 2016)~~. The annual precipitation of the area is more than 1600 mm (Alpert, 1942; USAID, 2014)~~(Alpert, 1942; USAID, 2014)~~. On 14 August 2021, Tiburon Peninsula was struck by a Mw 7.2 earthquake, which was followed by several aftershocks. The strongest one (Mw 5.7) occurred on 15 August 2021. Two days after the mainshock the area was hit by the intense Tropical Cyclone Grace. The combination of the two events triggered thousands of landslides (Martinez et al., 2021)~~(Martinez et al., 2021)~~ in the Pic Macaya National Park located in western part of the peninsula.

~~2.1.5.~~2.5. Rasuwa, Nepal

The study area is located in the Rasuwa district (central Nepal) in the higher Himalayas with altitudes ranging from 904 to 3267 m. a.s.l and annual average precipitation of 1800-2000 mm (Karki et al., 2016)~~(Karki et al., 2016)~~,The geology includes Proterozoic metamorphic rocks such as amphibolite, gneiss, and schist (Tiwari et al., 2017)~~(Tiwari et al., 2017)~~. The area was struck by the Mw 7.8 Gorkha earthquake on 25 April 2015. The intense seismic sequence produced at least 25,000 landslides (Roback et al., 2018)~~(Roback et al., 2018)~~.

~~2.1.6.~~2.6. Hokkaido, Japan

The Hokkaido study area is in northern Japan and has a high presence of croplands. The area is characterized by elevations between 50 and 500 m a.s.l., the geology is composed of Neogene sedimentary rocks, formed by the accumulation of numerous layers formed by materials ejected by the Tarumai volcano from several events over the years (Yamagishi and Yamazaki, 2018; Zhao et al. 2020; Koi et al. 2022). A severe earthquake hit the Hokkaido Iburi-Tobu area in Japan on September 6th, 2018. The earthquake registered a magnitude of

6.7 according to the Japan Meteorological Agency (JMA) and its epicenter was at 42.72° North and 142.0° East (Yamagishi and Yamazaki, 2018), located along the southern frontier of Hokkaido. The event triggered thousands of landslides (~7059) in a concentrated area of 466 km² (Zhao et al. 2020) and was responsible for 36 deaths (Yamagishi and Yamazaki, 2018).

~~2.1.7.~~2.7. Wenchuan, China

The study area is in the Longmenshan region at the eastern margin of the Tibetan Plateau, China. The location is characterized by high elevations up to 7.500 m a.s.l., the geology consists of lithological units from the Mesozoic, Jurassic, Cretaceous, Paleozoic, Precambrian formations and three types of Quaternary sedimentary units (Qi et al. 2010; Gorum et al. 2011). The area is constantly affected by earthquake-induced landslides over the years (e.g., 2017, 2018, 2019, 2021). The 2008 Wenchuan event is one of the most destructive events of mass movements related to earthquakes in the region (Fan et al. 2018). The Wenchuan earthquake hit a magnitude of Mw 7.9. It was responsible for triggering nearly 200.000 landslides (Xu et al. 2014), besides missing, injured, and thousands of human fatalities in a total area of 31,686.12 km² (Qi et al. 2010).

~~2.1.8.~~2.8. Sumatra, Indonesia

The investigated area is Mount Talamau (2912 m) which is a compound volcano located in West Pasaman Regency, West Sumatra Province, Indonesia. Geologically, the volcano consists of andesite and basalt rocks belonging to Pleistocene-Holocene age (Fadhilah & Prabowo, 2020; Zulkarnain, 2016)~~(Fadhilah & Prabowo, 2020; Zulkarnain, 2016)~~. The climate of the area is humid and tropical and the mean annual precipitation in West Pasaman area is between 3500 and 4500 mm/year (Wilis, 2019)~~(Wilis, 2019)~~. The Mw 6.1 earthquake hit West Sumatra on 25 February 2022. This event triggered several landslides in an area of 6 km^2, along the eastern and north-eastern flank of Talamau volcano.

~~2.1.9.~~2.9. Longchuan, China

The study area is located in the vicinity of Mibei village in Longchuan County, Guangdong Province, China with elevations between 180 and 600 m. The area has a subtropical monsoon climate, affected by frequent typhoons and rainstorms from May to October. The average annual precipitation ranges from 1300 to 2500 mm (Bai et al., 2021)~~(Bai et al., 2021)~~. The area is composed of Paleozoic completely weathered granite and Quaternary granite residual soil (Bai et al., 2021)~~(Bai et al., 2021)~~. Between 10 and 13 June 2019, an intense rainfall event, which was characterized by cumulative rainfall of 270 mm, triggered 327 shallow landslides between 300 and 400 m of altitudes and slopes ranging from 35 to 45° (Feng et al., 2022)~~(Feng et al., 2022)~~.

~~2.1.10.~~2.10. ~~-~~Hpa-An, Myanmar

The study area is located in Hpa-An district (central Kayin State, South Myanmar) in a tropical and monsoon area with a mean annual precipitation between 4500 and 5000 mm (Win Zin & Rutten, 2017)~~(Win Zin & Rutten, 2017)~~ and elevations up to 1300 meters. The area is part of the Shan Plateau sequence, which includes low grade metamorphosed Precambrian, Palezoic and Mesozoic sedimentary rocks (Jain & Banerjee, 2020)~~(Jain & Banerjee, 2020)~~. On 28–30 July 2018, Myanmar was hit by an extreme rainfall event which caused a flood along Bago river basin and triggered 992 landslides only in Kayin State (Amatya et al., 2022)~~(Amatya et al., 2022)~~.

2.11. Porgera, Papua New Guinea

The 2018 Papua New Guinea earthquake triggered over 200 landslides across the affected region, resulting in fatalities and severe infrastructure damage. The landslides were primarily caused by strong ground shaking and the steep topography of the region. Factors such as soil characteristics, rainfall, and vegetation cover also played a role. Understanding these factors can improve landslide hazard assessments and reduce future risk. Characteristics of the landslides included high relief, steep slopes, and weak lithology. Effective strategies for managing landslide hazards in high-risk areas should be developed. (Dang et al., 2020; Xu et al., 2020).

2.12. Kaikoura, New Zealand

The 2016 Kaikoura earthquake triggered more than 10,000 landslides in New Zealand, causing extensive damage and disrupting transportation routes. The landslides were complex and involved multiple failure mechanisms, including rockfalls, rock avalanches, and debris flows. The intense shaking and steep topography of the region contributed to the landslides. To reduce the potential impact of future earthquakes, it is crucial to improve understanding of landslide mechanisms and develop effective early warning systems (Goda et al., 2020; Massey et al., 2020; Wang et al., 2020).

2.13. Uvira, Democratic Republic of Congo

The city of Uvira in the Democratic Republic of Congo experienced devastating landslides in 2020 due to heavy rainfall, poor land management practices, and the steep topography of the region. These landslides caused significant damage to infrastructure and displaced thousands of people. Landslides are a recurring hazard in the DRC, with an average of 100 occurring annually, and climate change is expected to exacerbate the problem. Efforts to mitigate the risk of landslides can include improved land use practices, early warning systems, and infrastructure designed to withstand landslides. Taking a comprehensive approach is key to minimizing the impact of landslides and protecting at-risk communities. (Mwene-Mbeja et al., 2020; Kervyn et al., 2020; United Nations Office for Disaster Risk Reduction, 2020)

~~1.~~3. ~~escription~~ High-~~r~~Resolution ~~g~~Global ~~l~~l~~G~~andslide ~~d~~Detector ~~d~~Database (HR-GLDD)

3.1. Data set description:

The dataset created in this study consists of images acquired from the PlanetScope satellites (see table 1) and landslide inventories collected from the literature. For ~~some~~ all the events, landslides were manually delineated due to unavailability of existing inventories at high reolution. PlanetScope is a constellation of approximately 130 satellites that acquire images of the Earth daily with 3 meters of spatial resolution. The sensors acquire the images with 8 spectral bands: coastal blue (431 - 552 nm), blue (465 - 515 nm), green (547 - 583 nm), yellow (600 - 620 nm), red (650 - 680 nm), red-edge (697 - 713 nm) and NIR (845 - 885 nm) (Planet Team, 2019). ~~(Planet Team, 2019)(Planet Team, 2019)(Planet Team, 2019)(Planet Team, 2019)(Planet Team, 2019)(Planet Team, 2019)(Planet Team, 2019)(Planet Team, 2019).~~ PlanetScope imagery consists of surface reflectance values and 16 bits images. The images from both sensors are orthorectified and radiometrically corrected by the providers.

The dataset was prepared using only the red, green, blue, and NIR bands. The pre-processing phase was based on three steps: generation of binary masks, data sampling, and tiles patching. First, the interpreted landslides polygons from each area were rasterized using the Rasterio Python library into a binary mask, where "1" represents the landslides and "0" the background. The satellite imagery, along with the mask was then sampled and patched into a regular grid that yields patches of dimension 128 x 128 pixels, which correspond to 14.7 km^2 per patch. Since the imbalance between background area and landslides is strong, the images that did not have any landslides pixel labelled were removed. The proportions for the positive samples of landslides against the non-landslides are 9.96% and 90.04%, respectively. Table 1 shows the number of tiles created for each area.

Table 1 - Number of tiles, satellite information and landslide statistics for each study area.

Study Area	Satellite	Number of tiles	Study Area in km^2	Number of landslides	Minimum Landslide area (m^2)	Maximum Landslide area (m^2)	Total Landslide area (km^2)
Kodagu (India) 2018	PlanetScope	530	4033.62	343	276.23	581342.19	5.67
Rolante (Brazil) 2017	PlanetScope	33	24.62	113	381.76	81277.53	0.67
Tiburon Peninsula, (Haiti) 2021	PlanetScope	461	130.85	1394	200.74	473696	8.24
Rasuwa (Nepal) 2017	PlanetScope	222	114.68	184	676.85	115567.96	2.45
Hokkaido (Japan) 2018	PlanetScope	159	50.17	715	237.76	48524.72	5.29
Wenchuan (China) 2017	PlanetScope	284	58.25	1415	23.78	98467.96	3.19
Wenchuan (China) 2018	PlanetScope	263	58.25	546	110.18	1289210.19	5.54
Sumatra, (Indonesia) 2022	PlanetScope	403	22.56	584	302.26	6206089.32	9.73
Longchuan, (China) 2019	PlanetScope	110~~6~~	32.22	228	235.21	61163.17	0.73
Hpa-An, (Myanmar) 2018	PlanetScope	101~~8~~	28.38	540	101.23	88044.20	0.97
Papua New Guinea	PlanetScope	~~56~~725	304.94	491	262.65	259392.71	5.48

New Zealand	PlanetScope	287	150.75	246	676.67	165943.82	3.50
Democratic Republic of the Congo	PlanetScope	247	38.64	394	500.25	106094.52	1.61

~~1.1.~~3.2. Design of ~~High-Resolution Global Landslide Detector Database (~~HR-GLDD ~~)~~ ~~design~~

The performance evaluation of the study sites was carried out using metrics and trained using five state-of-the-art U-~~n~~Net like models, showcasing the capability and applicability of the High-Resolution Global Landslide Detector Database (HR-GLDD). We used a total of ten geographically distinct study sites distributed globally, where landslide events were chosen including different triggering mechanisms such as five earthquake~~s-~~induced and five rainfall-landslides-, we separately divide the patches into 60% for training, 20% for validation, and 20% for testing the model capabilities. All the sets are then mixed to create a unique dataset composed of~~by~~ equal percentages of patches.

We designed three scenarios to train, predict, and evaluate model performances in order to assess the robustness and applicability of the HR-GLDD. Primarily, we evaluate the model performances on the individual test sets. Secondly, we evaluate the performances of the models on the HR-GLDD test set. Moreover, finally, we test on two completely unseen recently occurred landslide events in Haiti 2021 and Indonesia 2022 (see figure 2).

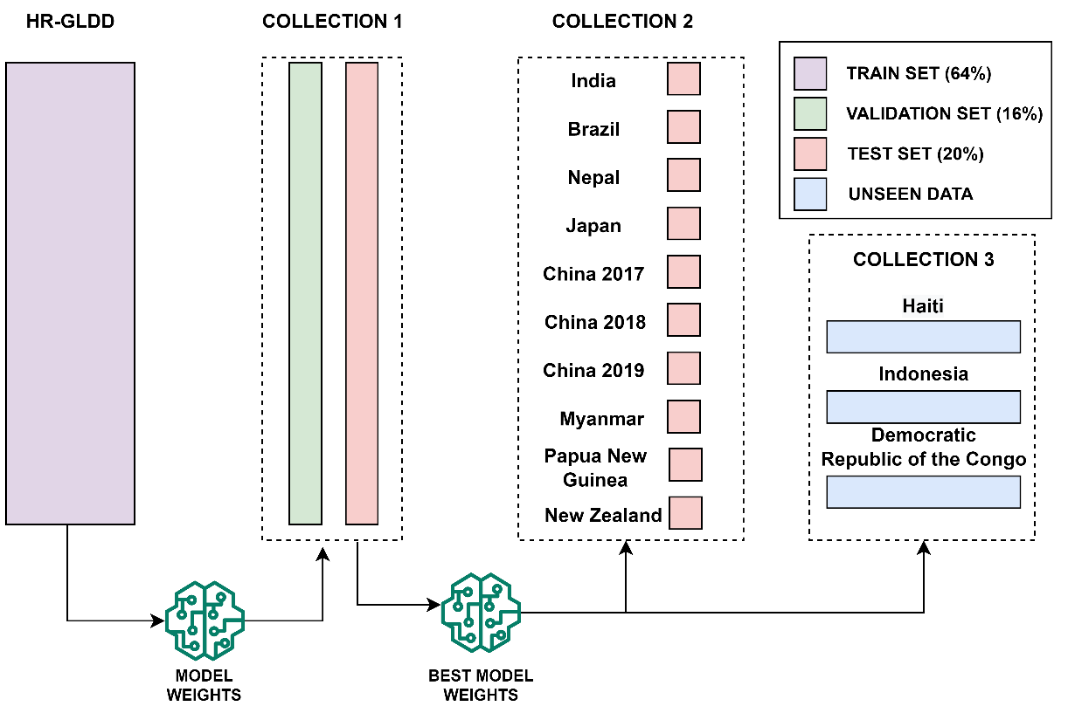


Figure 2: Schematic representation of the division of different components of HR-GLDD. Collection 1 refers to the test and validation data separated from the HR-GLDD. Collection 2 refers to the test dataset of individual sites. Collection 3 refers to the data from three~~wo~~ recent ~~data~~ events ~~set~~ for testing purposes.

~~1.~~4. Methodology

~~1.2.~~4.1. Model Architectures

The proposed dataset is evaluated through several state-of-the-art U-Net like Deep Learning segmentation models. In the past years, the U-Net (Abderrahim et al., 2020) has been used in several landslide detection applications which yield generally the most reliable results (Bhuyan et al., 2022; Meena et al., 2022c; Nava, Bhuyan, et al., 2022). Therefore, we decided to use it as a benchmark model when training on the proposed dataset. Moreover, several improved versions of the same are evaluated. We systematically trained the model using a variety of combinations of the hyper-parameters batch size (8, 16, 32, 64), learning rate (5e-4, 10e-4, 5e-5, 10e-5) and the number of filters of the first convolutional layer (8, 16, 32, 64). A description of the employed architectures is given in this section.

U-Net: This architecture has been utilized in various semantic segmentation applications, yielding generally outstanding results (Abderrahim et al., 2020). U-Net was employed initially in biomedical picture segmentation (Ronneberger et al., 2015). Low-level representations are captured by a contracting path (encoder), whereas a decoding path captures high-level representations. The encoding path consists of successive convolution blocks and is equivalent to a traditional CNN structure. Two convolutional layers with a 3 x 3 kernel size and a 2 x 2 max-pooling layer are present within every convolutional block. The rectified linear unit (ReLU) activation function is used to activate each convolutional layer (Agarap, 2018). A 2 x 2 max-pooling layer is added to the convolutional block's end in the encoder route to conduct non-linear downsampling, whereas, in the decoder path, a 2 x 2 upsampling layer takes its place. The upsampling layer is positioned right after a 3x3 convolutional layer (see figure S1). We refer to this combination as learnable upconvolution~~.S1)~~. We refer to this combination as learnable upconvolution.

Residual U-Net (Res U-Net): Res U-Net (Diakogiannis et al., 2020) follows the same U shape as U-Net, whereas here the above-explained convolutional blocks are replaced by residual blocks. This architecture's goal is to improve the learning capacities of the conventional U-Net as well as mitigate the gradient vanishing effect, especially when dealing with deep neural networks (such as U-Net) (see figure S2).

Attention U-Net and Attention Res U-Net: In the conventional U-Net as well as in the Res U-Net, cascading convolutions have been shown to provide false alerts for tiny objects with high form variability (Oktay et al., 2018). To select pertinent spatial information from low-level maps and therefore alleviate the problem, soft attention gates (AGs) are added (see figure S3, S4~~-~~). The attention gates are built on skip connections, which actively inhibit activations in unnecessary areas, lowering the number of duplicated features (Abraham & Mefraz Khan, 2018).

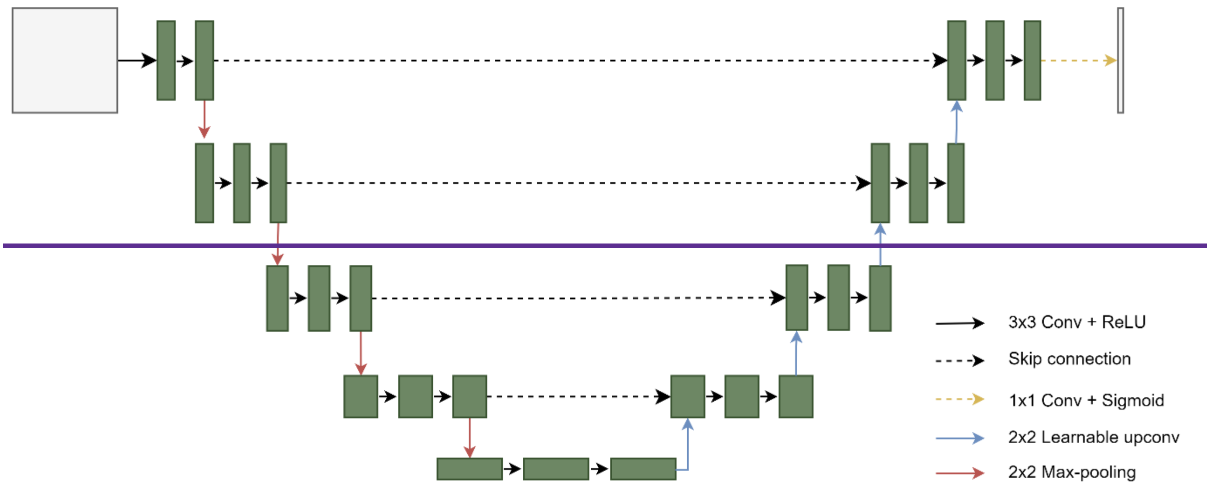
~~2.2.~~ Attention Deep Supervision Multi-Scale (ADSMS) U-Net: This architecture adopts the Attention U-Net structure, while, in addition, multi-scale image pyramid inputs are fed to the model, and a deep supervision strategy is applied (Abraham & Mefraz Khan, 2018). In practice, multi-scale inputs enable the model to gather that class data, which is more readily available at various sizes. This holds true for both background features and landslides. Lastly, where training data are few and networks are relatively shallow, deep supervision conducts a potent "regularization". More details about the deep supervision strategy used are available in the following section (see figure S5).~~The proposed dataset is evaluated through several state-of-the-art U-Net like Deep Learning segmentation models. In the past years, the U-Net (Abderrahim et al., 2020)(Abderrahim et al., 2020) has been used in several landslide detection~~

~~applications which yield generally the most reliable results (Bhuyan et al., 2022; Meena et al., 2022c; Nava, Bhuyan, et al., 2022)(Bhuyan et al., 2022; Meena et al., 2022c; Nava, Bhuyan, et al., 2022). Therefore, we decided to use it as a benchmark model when training on the proposed dataset. Moreover, several improved versions of the same are evaluated. We systematically trained the model using a variety of combinations of the hyper-parameters batch size (8, 16, 32, 64), learning rate (5e-4, 10e-4, 5e-5, 10e-5) and the number of filters of the first convolutional layer (8, 16, 32, 64). A description of the employed architectures is given in this section.~~

~~2.2.1. U-Net:~~

~~This architecture~~

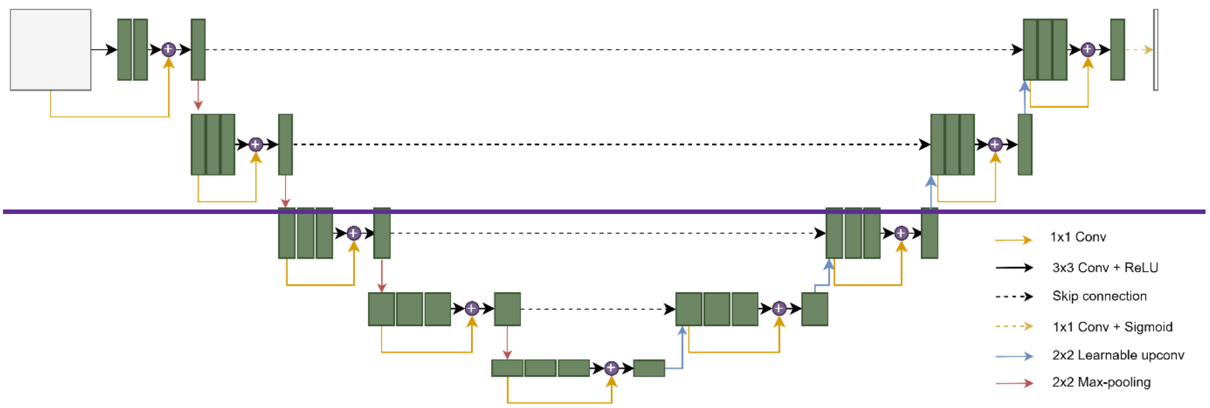
~~U-Net has been utilized in various semantic segmentation applications, yielding generally outstanding results (Abderrahim et al., 2020)(Abderrahim et al., 2020). U-Net was employed initially in biomedical picture segmentation (Ronneberger et al., 2015)(Ronneberger et al., 2015). Low-level representations are captured by a contracting path (encoder), whereas a decoding path captures high-level representations. The encoding path consists of successive convolution blocks and is equivalent to a traditional CNN structure. Two convolutional layers with a 3 x 3 kernel size and a 2 x 2 max-pooling layer are present within every convolutional block. The rectified linear unit (ReLU) activation function is used to activate each convolutional layer (Agarap, 2018)(Fred Agarap, n.d.). A 2 x 2 max-pooling layer is added to the convolutional block's end in the encoder route to conduct non-linear downsampling, whereas, in the decoder path, a 2 x 2 upsampling layer takes its place. The upsampling layer is positioned right after a 3x3 convolutional layer (see figure 3). We refer to this combination as learnable upconvolution.3). We refer to this combination as learnable upconvolution.~~



~~Figure 3: Model architecture of U-Net.~~

~~2.2.2. Residual U-Net (Res U-Net):~~

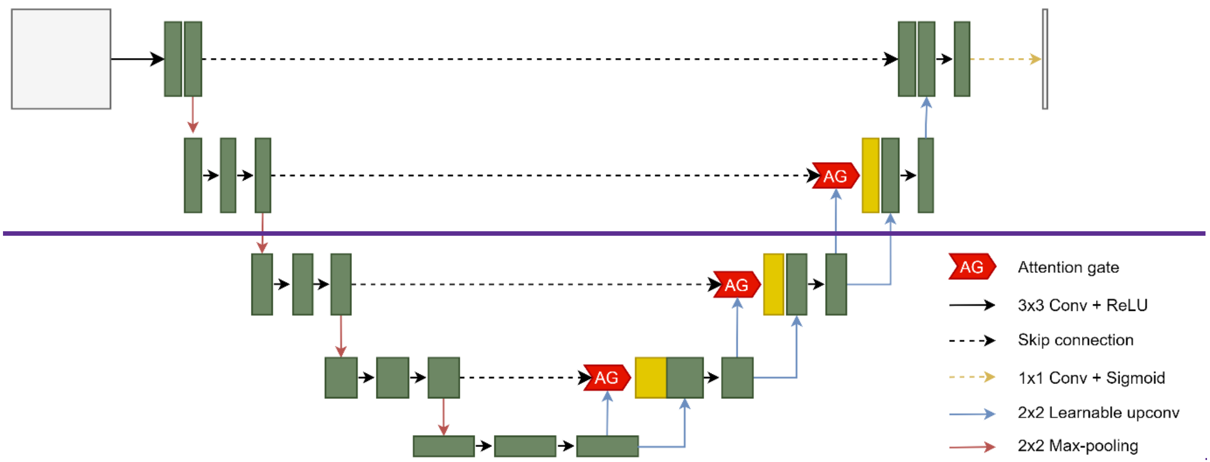
~~Res U-Net (Diakogiannis et al., 2020)(Diakogiannis et al., 2020) follows the same U shape as U-Net, whereas here the above-explained convolutional blocks are replaced by residual blocks. This architecture's goal is to improve the learning capacities of the conventional U-Net as well as mitigate the gradient vanishing effect, especially when dealing with deep neural networks (such as U-Net) (see figure 4).~~

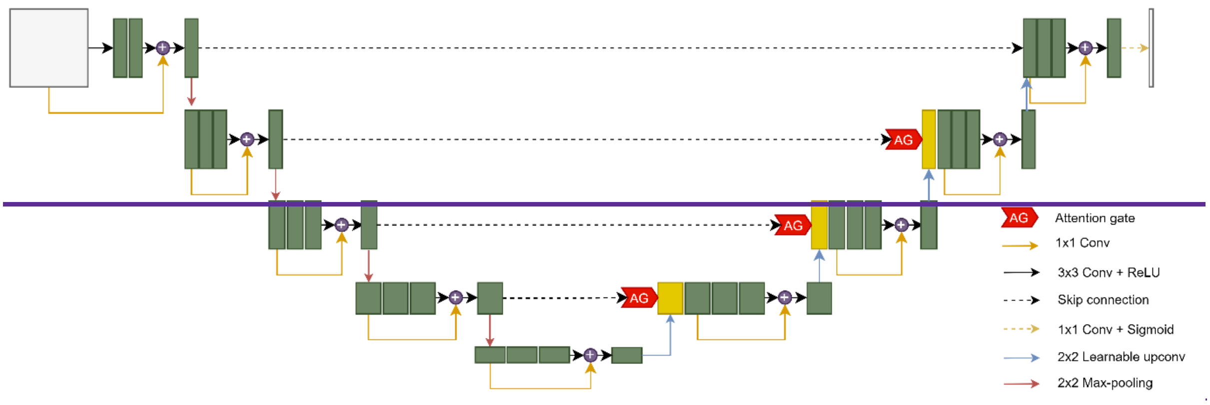


~~Figure 4: Model architecture of the Res U-Net.~~

~~2.2.3. Attention U-Net and Attention Res U-Net:~~

~~In the conventional U-Net as well as in the Res U-Net, cascading convolutions have been shown to provide false alerts for tiny objects with high form variability (Oktay et al., 2018) (Oktay et al., n.d.). To select pertinent spatial information from low-level maps and therefore alleviate the problem, soft attention gates (AGs) are added (see figure 5). The attention gates are built on skip connections, which actively inhibit activations in unnecessary areas, lowering the number of duplicated features (Abraham & Mefraz Khan, 2018)(Abraham & Mefraz Khan, 2018).~~

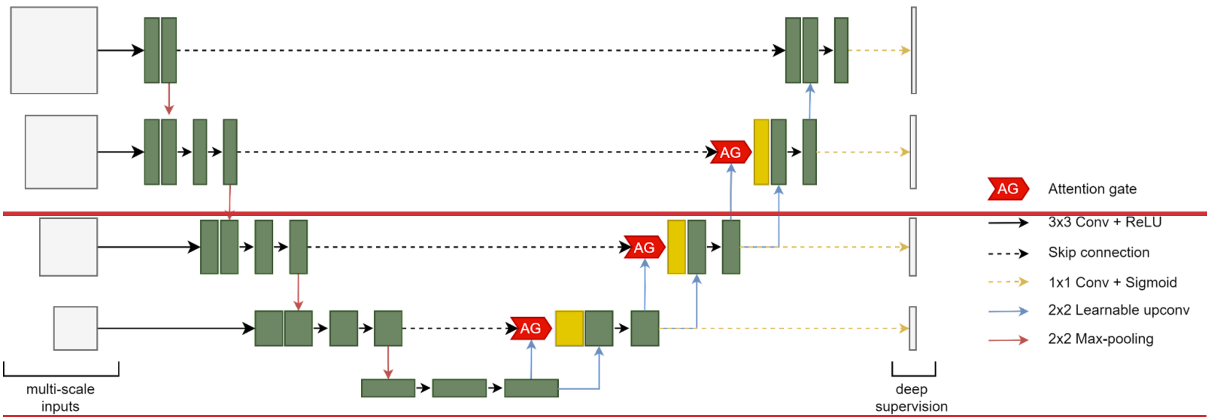




~~Figure 5: Model architecture of the (a) Attention U-Net and (b) Attention Res U-Net.~~

~~2.2.4. Attention Deep Supervision Multi-Scale (ADSMS) U-Net:~~

~~This architecture adopts the Attention U-Net structure, while, in addition, multi-scale image pyramid inputs are fed to the model, and a deep supervision strategy is applied (Abraham & Mefraz Khan, 2018)(Abraham & Mefraz Khan, 2018). In practice, multi-scale inputs enable the model to gather that class data, which is more readily available at various sizes. This holds true for both background features and landslides. Lastly, where training data are few and networks are relatively shallow, deep supervision conducts a potent "regularization". More details about the deep supervision strategy used are available in the following section (see figure 6).~~



~~Figure 6: Model architecture of the Attention Deep Supervision Multi-Scale U-Net.~~

~~1.3.~~4.2. Model~~s~~ training

To train the DL models, we utilized Dice Loss (Milletari et al., 2016)~~(Milletari et al., n.d.)~~ as the loss function:

Dice Loss =c1Ni=1picgic+Ni=1pic+gic

Equation (1) illustrates a two-class Dice score coefficient (DSC) variation for class c, where gic [0, 1] and pic [0, 1] are the ground truth and predicted labels, respectively. Furthermore, the numerical stability is assured by avoiding division by zero, while N specifies the total number of picture pixels. As an exception, in the ADSMS U-Net model, every high-dimensional feature representation is regulated by Focal Tversky Loss to avoid loss over-suppression, while the final output is controlled by the conventional Tversky Loss. This deep supervision strategy, described in Lee et al., (2015)~~(Lee et al., n.d.)~~, requires intermediate layers to be semantically discriminative at all scales. Furthermore, it contributes to ensuring that the attention unit has the power to change responses to a wide variety of visual foreground material. This strategy is adopted from (Abraham & Mefraz Khan, 2018)~~(Abraham & Mefraz Khan, 2018)~~, who propose it along with the ADSMS U-Net architecture. As the loss function optimizer, for all the models, we used a stochastic gradient descent strategy based on an adaptive estimate of first- and second-order moments (Adam), which is useful in problems with uncertain data and sparse gradients (Kingma & Ba, 2015)~~(Kingma & Lei Ba, n.d.)~~. The precision, recall, F1-score, and Intersection Over Union (IOU) score, the most common accuracy evaluation measures for segmentation models, all of which have been utilized in several landslide detection studies, were used to measure how well the applied DL models performed in detecting landslides. The appropriate combinations of hyper-parameters must be used while training such DL models in order to optimize the model and, therefore, output the best results.

~~2.~~5. Results

~~2.1.~~5.1. HR-GLDD evaluation results

The robustness and applicability of the HR-GLDD was tested using the best model weight. We train and calibrate the models using the HR-GLDD. The best weighs for each model are selected based on the model performances on the mixed test set of the HR-GLDD dataset. After running the models on test dataset, batch size of 16~~4~~ and Adam optimiser with learning rate 5.00E-04 ~~1e-3~~ resulted in best model weight. To further evaluate the efficiency and generalization capabilities of the models, we use the model on ~~two~~ three unseen datasets to map landslides in the two different geomorphological areas that were recently affected by multiple landslide events. We chose the most recent events one occurred after Uvira, Democratic Republic of Congo (DRC) heavy rainfall event of April 2020. Haiti earthquake in August 2021, ~~and another~~ one in Sumatra, Indonesia after a heavy rainfall event of February 2022. A total of 247, 461 and 403 unseen image patches were evaluated for DRC, Haiti and Indonesia, respectively.

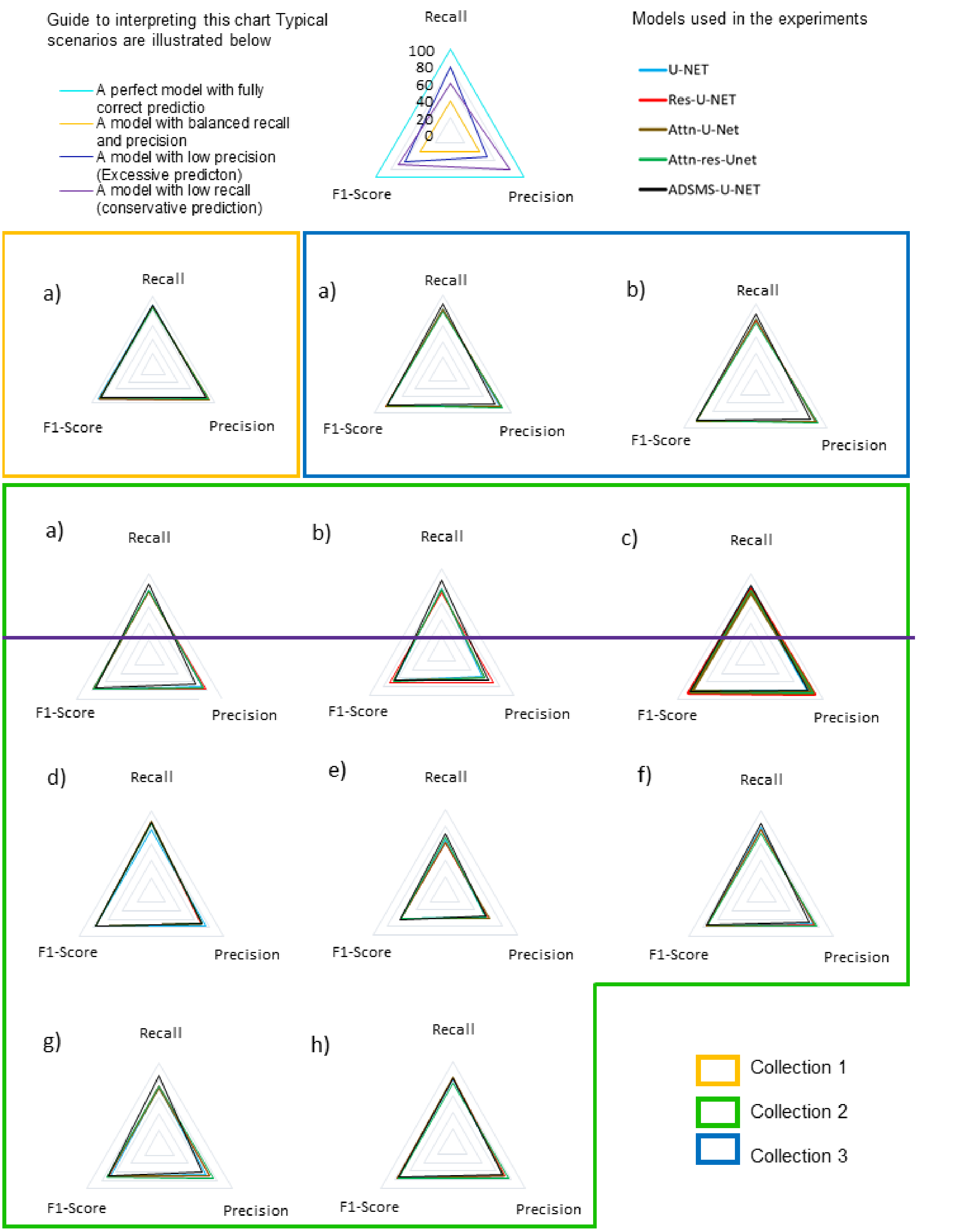
Experimental results for landslide detection by utilising the HR-GLDD are presented in Table 2. Overall, all the models performed consistently in collections 2 and 3. The F1-score evaluation results for each test case of all the models demonstrate the applicability of the HR-GLDD training dataset for landslide detection results~~, especially with employing only the optical bands~~. The average F1-score for HR-GLDD test dataset (collection 1) across all the models was around ~~72%~~0.7045~~, which is relatively uniform~~. Furthermore, the same was observed in the individual test sites in collection 2. We also notice that the ~~Recall and~~ Precision and ~~are~~Recall are ~~pretty~~ well balanced ranging between 0.~~72.15~~6346~~%~~-0.76~~.~~61~~%~~ and 0.6~~8.13~~672~~%~~-0.~~75.47~~8121~~%~~, respectively, indicating stable model predictions ~~(see figure 7figure 3)~~. In collection 3, the metrics reveal positive outcomes in terms of mapping the landslides following the respective events, with an average F1-score of 0.5562 ~~80%~~ for DRC, 0.7947 for Haiti and 0.8603~~86%~~ for Indonesia. The recall values are higher than precision values for all the models ~~which have a difference of about 3.32%~~ resulting in average F1-score of ~~72.54%~~0.7045 (see table 2)~~(see figure 7figure 3)~~. Higher values of recall in all models means that the models were able to identify landslide labelled pixels. However due to the use of only the optical bands, the spectral signatures of other similar features (such as riverbeds and flat barren areas) were labelled as landslides which result in false predictions, thereby, accounting for lower precision.

In ~~figure 8~~figure 4~~3~~ we chose a single image patch to showcase the predictions of the various models with~~in~~ respect to the referenced ground truth. Despite the differences in the spectral fingerprints of the satellite images for each study site and the events initiated by an earthquake or rainfall, the models were still capable of recognizing landslide features (see ~~figure 9~~figure 5~~4~~, 5 ~~-~~and ~~10~~6). Particularly, we were able to map the recent events in DRC (2020), Haiti (2021) and ~~,~~Indonesia (2022).~~, and DRC (2020).~~

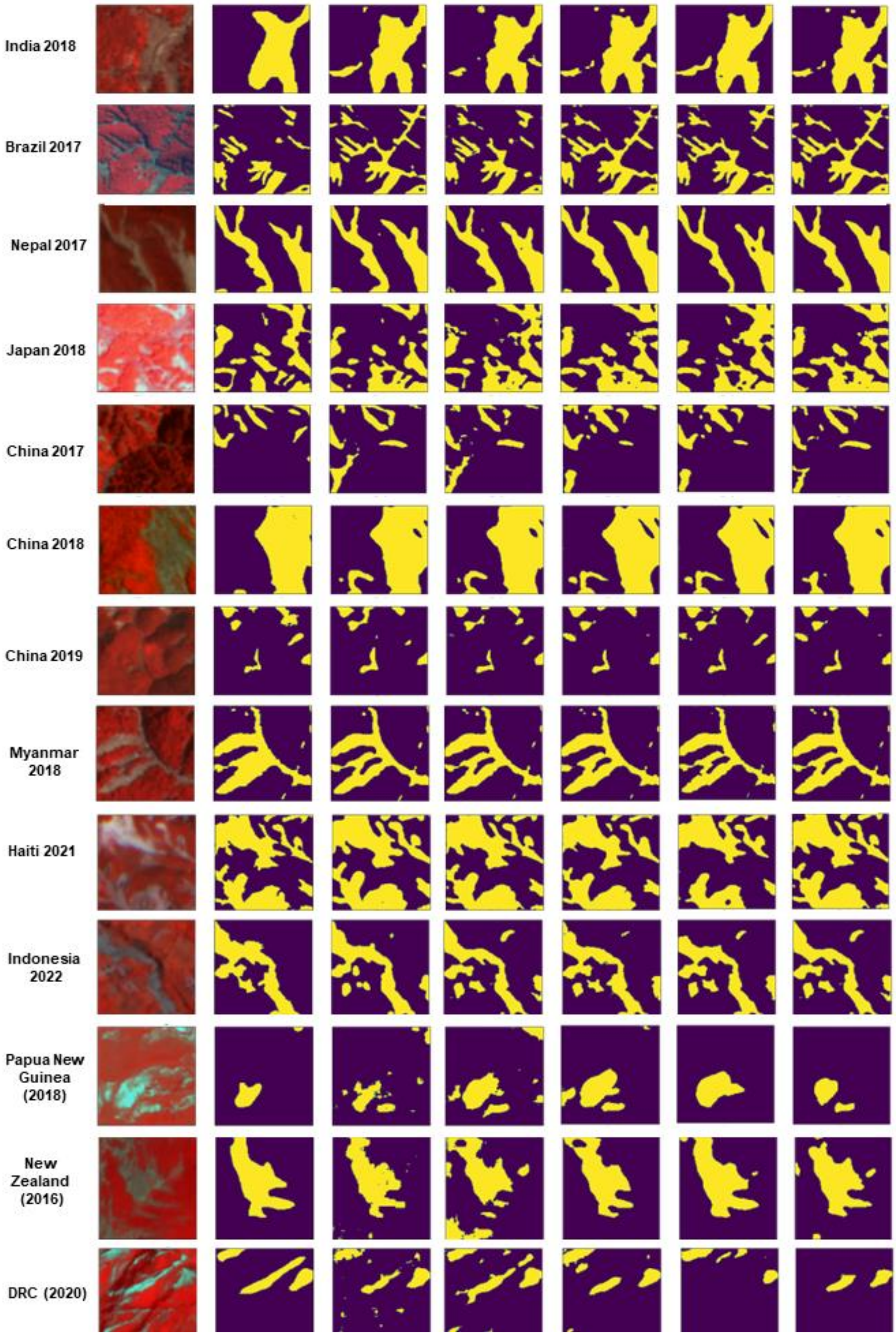
Table 2: F1 scores of different DL models across sites and HR-GLDD test dataset along with ~~two~~ three unseen test sites.

	~~Study sites~~	~~U-NET~~	~~Res-U-NET~~		~~Attn-U-NET~~	~~Attn-res-Unet~~
	~~Collection 1 (HR-GLDD Test)~~	**~~70.52~~**	**~~72.54~~**	**~~72.33~~**	**~~72.52~~**	**~~72.1~~**
Collection 2	~~India~~	~~75.89~~	~~77.77~~	~~76.62~~	~~77.CIA O15~~	~~74.2~~
	~~Brazil~~	~~64.88~~	~~71.73~~		~~66.19~~	~~67.1~~
	~~Nepal~~	~~82.65~~	~~84.56~~	~~81.99~~	~~83.15~~	~~81.7~~
	~~Japan~~	~~76.19~~	~~76.78~~		~~77.5~~	~~76.7~~
	~~China2017~~	~~60.46~~	~~60.13~~	~~61.04~~	~~60.98~~	~~62.3~~
	~~China2018~~	~~75.04~~	~~75.33~~		~~75.97~~	~~74.4~~
	~~China2019~~	~~67.9~~	~~70.62~~	~~69.93~~	~~73.17~~	~~70.2~~
	~~Myanmar~~	~~74.49~~	~~76.67~~		~~75.96~~	~~75.7~~
	~~Collection 3~~					
	~~Indonesia~~	~~88.4~~	~~87.86~~		~~87.75~~	~~85.9~~
	~~Haiti~~	~~78.55~~	~~82.86~~		~~80.28~~	~~81.6~~

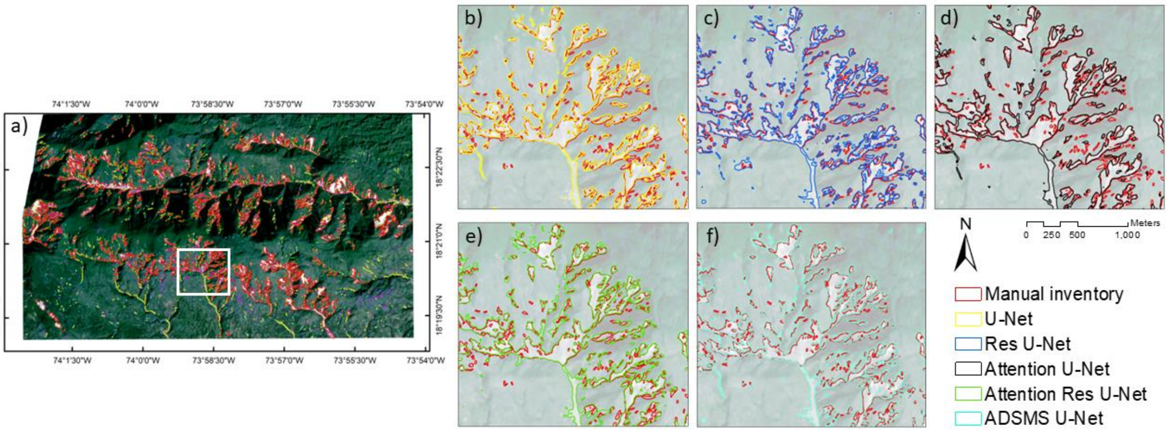
	Study sites	U-NET	Res-U-NET	Attn-U-NET	Attn-res-Unet	ADSMS-U-NET
	Collection 1 (HR-GLDD Test)	**0.7904**	**0.6825**	**0.7446**	**0.6477**	**0.6576**
Collection 2	India	0.7674	0.6980	0.7628	0.6664	0.6796
	Brazil	0.7739	0.6913	0.6539	0.6830	0.6726
	Nepal	0.8972	0.8149	0.8419	0.7695	0.7976
	Japan	0.8159	0.7479	0.8124	0.7317	0.7552
	Wenchuan (China2017)	0.7781	0.6507	0.6981	0.6162	0.6739
	Wenchuan (China2018)	0.8077	0.6886	0.7295	0.6704	0.6557
	Longchuan (China2019)	0.6842	0.5076	0.5422	0.4829	0.4398
	Myanmar	0.8415	0.7861	0.7826	0.7405	0.7709
	Papua New Guinea	0.7515	0.6150	0.7568	0.6572	0.6261
	New Zealand	0.7496	0.5456	0.7335	0.4922	0.6494
	Collection 3					
	Indonesia	0.8832	0.8810	0.8232	0.8534	0.8608
	Haiti	0.8357	0.8055	0.7869	0.7648	0.7808
	Democratic Republic of the Congo	0.5937	0.5366	0.5682	0.5008	0.5819



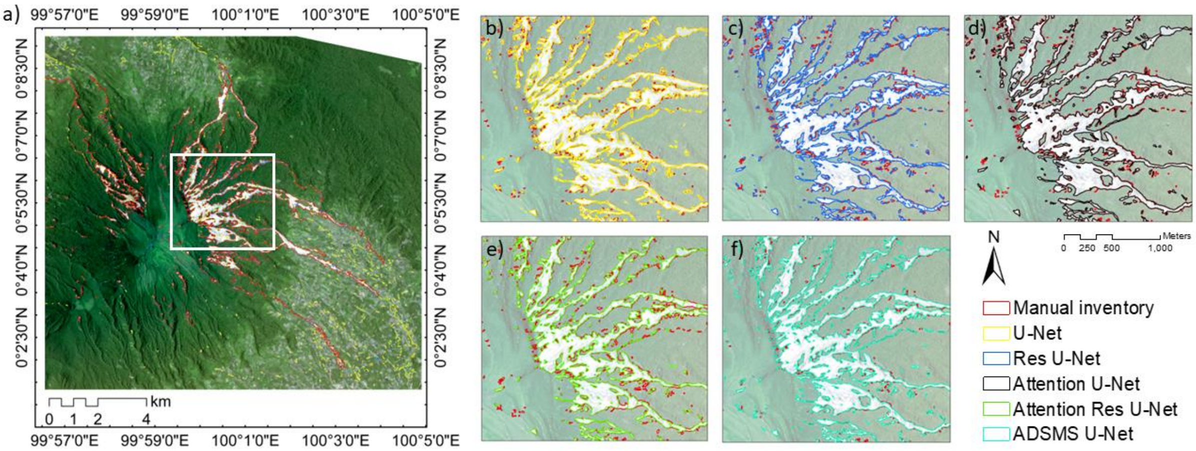
~~Figure 7: Radar charts visualising the quantitative performance metrics for the DL models on the three test set collections.~~



~~Figure 8~~Figure 34: Landslide predictions made by the different DL models against the ground truth. The base image is shown as a false colour composite (FCC) to better visualize the scars of the landslides.



~~Figure 9~~Figure ~~5~~4: Comparison of ground truth landslides with predictions from the DL models for the unseen dataset of Haiti.



~~Figure 10~~Figure ~~6~~5: Comparison of ground truth landslides with predictions from the DL models for the unseen dataset of Indonesia.

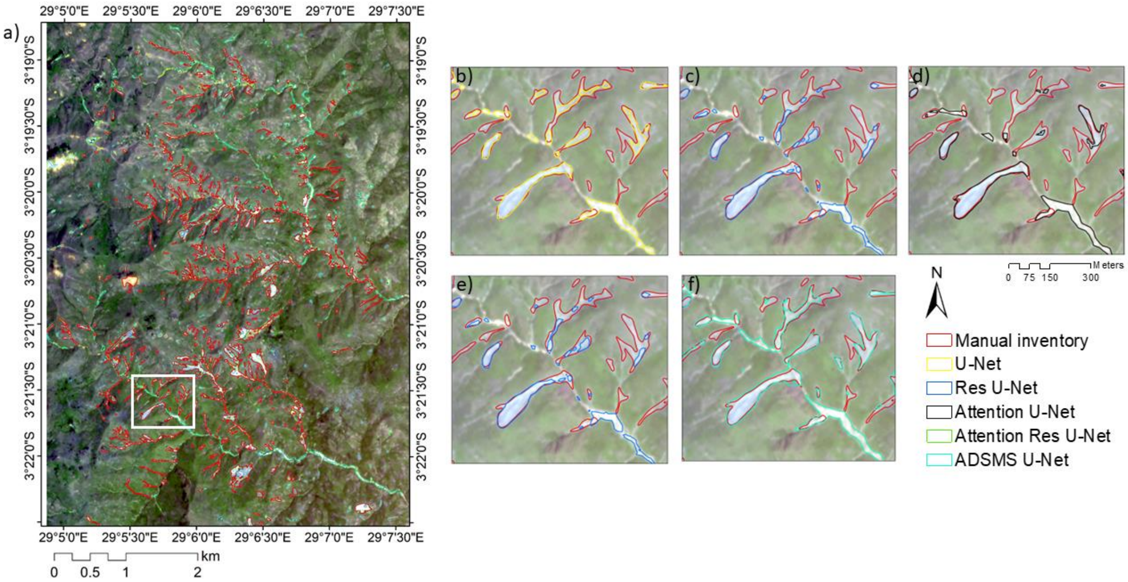


Figure ~~7~~6: Comparison of ground truth landslides with predictions from the DL models for the unseen dataset of DRC.

~~3.~~6. Discussions

~~3.1.~~6.1. Advantages of using HR images

The spatial resolution of Planet Scope imagery enables the detection of small size landslides that open access satellite missions like Sentinel and Landsat frequently miss due to their spatial and temporal resolution (Meena et al., 2021~~b~~)~~(Meena et al., 2021b)~~. Moreover, even though Sentinel-2 has additional spectral bands, the lack of improved spatial resolution inhibits precise boundary delineation and landslide localisation (Meena et al., 2022~~d~~)~~(Meena et al., 2022d)~~. The most prominent features of Planet Scope imagery, in addition to its competitive spatial resolution, are its daily temporal resolution and global coverage. Since the satellites have identical sensors, the imageries~~y~~ are orthorectified and image pre-processing are simplified and more accurate. Because Planet imagery provide global coverage, we may extend our study sites to new locations for generating more quality datasets that allow for a better model generalization.

~~3.2.~~6.2. Quality of HR-GLDD

The quality of any ML/DL model depends on the data that it is trained on, and the GLDD aims to meet this fundamental requirement. To our knowledge, no other quality data sets exist that can accommodate the wide range of landslide-triggering events and topographical diversity needed for efficient model training. As the GLDD is a strong collection of various landslide events caused both by rainfall and earthquakes. The GLDD is designed to calibrate models able to map new events that will occur in the future. The models investigated in our study gave promising and consistent results for two unseen datasets generated by completely different events, indicating a well-prepared, dependable, and resilient dataset. However, there are clear limitations with the GLDD that must be considered. These problems primarily stem from issues with manually delineated polygons and various uncertainties caused by satellite imagery. A number of different variables, including the mapping scale, the date, and the quality of the satellite imagery, affect how accurate an inventory is. The radiometric resolution and cloud coverage are additional variables that affect the generation of manual inventories. Additionally, haze effect caused by instrument errors hinders model performances. Subjectivity in the landslide polygon boundaries results from the amalgamation problem, which is caused by elements like the investigators' level of experience and the goal of the study.

~~3.3.~~6.3. Significance of the HR-GLDD

A thorough hazard and risk framework is made possible by quality landslide inventories however, the generating such inventories at large scales takes ample amount of time and resources. This is where such automatic pipelines can truly shine at creating inventories which can be used for the successive phases of a hazard and risk. Local, regional, and national stakeholders may include such inventories into their risk reduction efforts thanks to the availability of inventories produced automatically. Furthermore, this information may serve as the foundation for a legal framework that implements landslide risk. A landslide risk reduction plan is now more crucial than ever given the anticipated rise in worldwide landslide activity brought on by climate change. Higher landslide activity is expected in the future due to a number of factors, including an increase in the frequency and intensity of seismic events, anthropogenic events, heavy precipitation events, rising ground water levels, storm surges, and a general rise in relative sea level. Therefore, it is essential to comprehend the underlying mechanisms of landslides better and create practical risk reduction techniques to save people's lives and property.

~~3.4.~~6.4. Automated pipeline for HR-GLDD

At the moment, automated techniques are the only viable solution for mapping vast regions with accuracy appropriate for operational objectives. Nonetheless, reliable, reproducible, and accurate processes for automating landslide detection across huge data stacks are still absent. As a result, many landslide-affected regions remain unmapped because 1) they are challenging to map using standard methods, and 2) using high-resolution imagery is costly and labour-intensive, with a substantial part of the mapping process dependent on human judgment. By overcoming these challenges, automated pipelines that address these issues can considerably reduce the requirement for human involvement and pave the way for the development of reliable real-time mapping and monitoring of natural hazards at the continental and global scales. Based on the quality of GLDD, reliability of automated pipelines and rapidly growing availability of HR satellite imagery, we can realistically envision mapping of landslide instances and contribute towards generating and updating landslide inventories at large-scales, spatially and potentially, also temporally (Bhuyan et al., 2023)~~(Bhuyan et al., 2023)~~.

Providing an expert-based, high-quality, and scientifically validated landslide inventory to scientific communities is essential for frameworks of modelling, landslide prediction, machine learning, and deep learning research. The GLDD dataset has been verified, which increases the availability of much-needed training datasets for automated mapping algorithms. The consistently long time taken to compile landslide inventories manually contrasts with the rise in data accessible for landslide mapping. The development of technologies to successfully automate the procedure is the future direction in landslide inventory mapping. The precedence of quality dataset is noted in where they commented that the need for quality datasets will provide a valuable resource for training and developing algorithms.

The current dataset is an excellent resource for training and developing future algorithms for this purpose. Automated mapping methods, particularly when combined with publicly available elevation models, can potentially improve our results in future investigations.

~~4.~~7. Conclusions

Mapping landslides through space is a challenging endeavour. Automated efforts for the same have been explored to some extent, but a transferrable method based on a robust dataset has not yet been investigated. In this paper, we propose a reliable dataset which can be employed by deep learning algorithms to detect new landslides accurately. The predictive capabilities demonstrate the usefulness and application of the dataset to map landslides at large scales. However, the model's predictability must be investigated further in order to identify particular problems to enhance the findings and predictive capabilities for more complicated landscapes. Overall, despite the limitations, the findings are promising, since it is the first time such a HR dataset has been created that caters to a transferable approach of mapping landslides at so many different geomorphological and geographical locations.

Data availability

The data, working codes and a document with metadata are freely available at https://github.com/kushanavbhuyan/HR-GLDD-A-Global-Landslide-Mapping-Data-Repository where data in the format of arrays and model configurations in the framework of TensorFlow as can be displayed and used for reproducibility of our results. We also submit the generated landslide inventories in the form of an Environmental Systems Research Institute (ESRI) shapefile. Modules for deep learning can be found at https://www.tensorflow.org/ and original satellite imageries can be found at https://www.planet.com/.

Code availability

Code used to produce data described in this manuscript, as well as to create figures and tables, can be accessed at https://github.com/kushanavbhuyan/HR-GLDD-A-Global-Landslide-Mapping-Data-Repository

Author contribution

All the authors contributed to equally to preparation of manuscript from data curation to review of final manuscript.

Competing interests

-The authors declare that they have no conflict of interest.

References

Abderrahim, N. Y. Q., Abderrahim, S., & Rida, A. (2020). Road Segmentation using U-Net architecture. 2020 IEEE International Conference of Moroccan Geomatics (Morgeo), 1–4. https://doi.org/10.1109/Morgeo49228.2020.9121887

Abraham, N., & Mefraz Khan, N. (2018). A NOVEL FOCAL TVERSKY LOSS FUNCTION WITH IMPROVED ATTENTION U-NET FOR LESION SEGMENTATION. https://github.com/nabsabraham/focal-tversky-unet

Alpert, L. (1942). Rainfall map of Hispaniola. Bulletin of the American Meteorological Society, 23, 423–431.

Amatya, P., Kirschbaum, D., & Stanley, T. (2022). Rainfall-induced landslide inventories for Lower Mekong based on Planet imagery and semi-automatic mapping method. Geoscience Data Journal, 00, 1–13. https://doi.org/10.1002/gdj3.145

Bai, H., Feng, W., Yi, X., Fang, H., Wu, Y., Deng, P., Dai, H., & Hu, R. (2021). Group-occurring landslides and debris flows caused by the continuous heavy rainfall in June 2019 in Mibei Village, Longchuan County, Guangdong Province, China. Natural Hazards, 108(3), 3181–3201. https://doi.org/10.1007/s11069-021-04819-1

Bhuyan, K., Tanyaş, H., Nava, L. *et al.* Generating multi-temporal landslide inventories through a general deep transfer learning strategy using HR EO data. *Sci Rep* **13**, 162 (2023). https://doi.org/10.1038/s41598-022-27352-y

Bhuyan, K., Meena, S. R., Nava, L., van Westen, C. J., Floris, M., & Catani, F. (2022). Mapping landslides through a temporal lens: An insight towards multi-temporal landslide mapping using the U-Net deep learning model. Earth Arxiv. https://doi.org/https://doi.org/10.31223/X5DM0B

Cruden, D. M., & Varnes, D. (1996). LANDSLIDE TYPES AND PROCESSES. In National Research Council, Transportation Research Board.

Dang, K. T., Wang, G., Su, Y., Xu, Q., & Chen, W. (2020). Landslides triggered by the 2018 M7.5 earthquake in Papua New Guinea. Natural Hazards and Earth System Sciences, 20(6), 1647-1660.

Diakogiannis, F. I., Waldner, F., Caccetta, P., & Wu, C. (2020). ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data. ISPRS Journal of

Photogrammetry and Remote Sensing, 162, 94–114. https://doi.org/10.1016/j.isprsjprs.2020.01.013

EM-DAT. (2018). The emergency events database—Universit´ e catholique de Louvain (UCL)—CRED, D Guha-Sapir.

Fadhilah, & Prabowo, H. (2020). Genes and physical properties of iron sand from Kinali Pasaman. Journal of Physics: Conference Series, 1481(1). https://doi.org/10.1088/1742-6596/1481/1/012015

Fan, X., Scaringi, G., Domènech, G., Yang, F., Guo, X., Dai, L., He, C., Xu, Q., & Huang, R. (2019). Two multi-temporal datasets that track the enhanced landsliding after the 2008 Wenchuan earthquake. Earth System Science Data, 11(1), 35–55. https://doi.org/10.5194/essd-11-35-2019

Feng, W., Bai, H., Lan, B., Wu, Y., Wu, Z., Yan, L., & Ma, X. (2022). Spatial–temporal distribution and failure mechanism of group-occurring landslides in Mibei village, Longchuan County, Guangdong, China. Landslides, May. https://doi.org/10.1007/s10346-022-01904-9

Fred Agarap, A. M. (n.d.). Deep Learning using Rectified Linear Units (ReLU). Retrieved August 15, 2022, from https://github.com/AFAgarap/relu-classifier.

Froude, M. J., & Petley, D. N. (2018). Global fatal landslide occurrence from 2004 to 2016. Natural Hazards and Earth System Sciences, 18(8), 2161–2181. https://doi.org/10.5194/nhess-18-2161-2018

Ghorbanzadeh, O., Xu, Y., Ghamisi, P., Kopp, M., & Kreil, D. (2022). Landslide4Sense: Reference Benchmark Data and Deep Learning Models for Landslide Detection. https://doi.org/10.48550/arxiv.2206.00515

Goda, K., Ren, J., & Anderson, N. (2020). Landslides induced by the 2016 Kaikoura earthquake, New Zealand: characteristics, distribution, and correlations. Landslides, 17(4), 825-839.

Harp, E. L., Jibson, R. W., & Schmitt, R. G. (2016). Map of landslides triggered by the January 12, 2010, Haiti earthquake. https://doi.org/10.3133/sim3353

Hungr, O., Leroueil, S., & Picarelli, L. (2014). The Varnes classification of landslide types, an update. In Landslides (Vol. 11, Issue 2, pp. 167–194). Springer Verlag. https://doi.org/10.1007/s10346-013-0436-y

Jain, A. K., & Banerjee, D. M. (2020). The Indian Subcontinent: Its tectonics. In Proceedings of the Indian National Science Academy (Vol. 86, Issue 1). https://doi.org/10.16943/ptinsa/2020/49834

Karki, R., Talchabhadel, R., Aalto, J., & Baidya, S. K. (2016). New climatic classification of Nepal. Theoretical and Applied Climatology, 125(3–4), 799–808. https://doi.org/10.1007/s00704-015-1549-0

Kervyn, M., Jacobs, L., Sumbwe, J., Maki Mateso, J. C., Kervyn, F., Sebagenzi, S. N., & Havenith, H. B. (2020). Landslide hazard and risk assessment in the Democratic Republic of Congo. Landslides, 17(8), 1901-1916.

Kingma, D. P., & Lei Ba, J. (n.d.). ADAM: A METHOD FOR STOCHASTIC OPTIMIZATION.

Lee, C.-Y., Xie, S., Gallagher, P., Zhang, Z., & Tu, Z. (n.d.). Deeply-Supervised Nets.

Liu, Y. ;, Yao, X. ;, Gu, Z. ;, Zhou, Z. ;, Liu, X. ;, Chen, X. ;, Wei, S., Liu, Y., Yao, X., Gu, Z., Zhou, Z., Liu, X., Chen, X., & Wei, S. (2022). Study of the Automatic Recognition of Landslides by Using InSAR Images and the Improved Mask R-CNN Model in the Eastern Tibet Plateau. Remote Sensing 2022, Vol. 14, Page 3362, 14(14), 3362. https://doi.org/10.3390/RS14143362

Martinez, S. N., Allstadt, K. E., Slaughter, S. L., Schmitt, R. G., Collins, E., Schaefer, L. N., & Ellison, S. (2021). Landslides triggered by the August 14, 2021, magnitude 7.2 Nippes, Haiti, earthquake. Open-File Report.

Massey, C. I., Hancox, G. T., & Van Dissen, R. J. (2020). The 2016 Kaikoura, New Zealand, Earthquake: A Complex Multihazard and Multirisk Event. Annual Review of Earth and Planetary Sciences, 48, 235-259.

Meena, S. R., Ghorbanzadeh, O., van Westen, C. J., Nachappa, T. G., Blaschke, T., Singh, R. P., & Sarkar, R. (2021). Rapid mapping of landslides in the Western Ghats (India) triggered by 2018 extreme monsoon rainfall using a deep learning approach. Landslides, 18(5). https://doi.org/10.1007/s10346-020-01602-4

Meena, S. R., Soares, L. P., Grohmann, C. H., van Westen, C., Bhuyan, K., Singh, R. P., Floris, M., & Catani, F. (2022). Landslide detection in the Himalayas using machine learning algorithms and U-Net. Landslides, 19(5), 1209–1229. https://doi.org/10.1007/s10346-022-01861-3

Milletari, F., Navab, N., and Ahmadi, S.-A. (2016). V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. In Proceedings of the Fourth International Conference on 3D Vision (3DV), 565-571. doi: 10.1109/3DV.2016.79.

Mwene-Mbeja, M., Mugaruka, J. P., Bisimwa, B., & Tchicaya, L. (2020). Landslide hazard and risk assessment in Uvira city, Democratic Republic of Congo. Environmental Hazards, 19(3), 263-278.

Nava, L., Bhuyan, K., Meena, S. R., Monserrat, O., & Catani, F. (2022). Rapid Mapping of Landslides on SAR Data by Attention U-Net. Remote Sensing 2022, Vol. 14, Page 1449, 14(6), 1449. https://doi.org/10.3390/RS14061449

Nava, L., Monserrat, O., & Catani, F. (2022). Improving Landslide Detection on SAR Data Through Deep Learning. IEEE Geoscience and Remote Sensing Letters, 19. https://doi.org/10.1109/LGRS.2021.3127073

Oktay, O., Schlemper, J., le Folgoc, L., Lee, M., Heinrich, M., Misawa, K., Mori, K., Mcdonagh, S., Hammerla, N. Y., Kainz, B., Glocker, B., & Rueckert, D. (n.d.). Attention U-Net: Learning Where to Look for the Pancreas.

Planet Team. (2019). Planet Imagery Product Specifications August 2019. 97.

Prakash, N., Manconi, A., & Loew, S. (2021). A new strategy to map landslides with a generalized convolutional neural network. Scientific Reports, 11(1), 1–15. https://doi.org/10.1038/s41598-021-89015-8

Roback, K., Clark, M. K., West, A. J., Zekkos, D., Li, G., Gallen, S. F., Chamlagain, D., & Godt, J. W. (2018). The size, distribution, and mobility of landslides caused by the 2015 Mw7.8 Gorkha earthquake, Nepal. Geomorphology, 301, 121–138. https://doi.org/10.1016/j.geomorph.2017.01.030

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation (pp. 234–241). https://doi.org/10.1007/978-3-319-24574-4_28

Soares, L. P., Dias, H. C., Garcia, G. P. B., & Grohmann, C. H. (2022a). Landslide Segmentation with Deep Learning: Evaluating Model Generalization in Rainfall-Induced Landslides in Brazil. Remote Sensing, 14(9), 2237. https://doi.org/10.3390/rs14092237

Tang, X., Tu, Z., Wang, Y., Liu, M., Li, D., & Fan, X. (2022). Automatic Detection of Coseismic Landslides Using a New Transformer Method. Remote Sensing, 14(12), 2884. https://doi.org/10.3390/rs14122884

Tiwari, B., Ajmera, B., & Dhital, S. (2017). Characteristics of moderate- to large-scale landslides triggered by the Mw 7.8 2015 Gorkha earthquake and its aftershocks. Landslides, 14(4), 1297–1318. https://doi.org/10.1007/s10346-016-0789-0

United Nations Office for Disaster Risk Reduction. (2020). Landslide risk reduction in the Great Lakes Region. Retrieved from https://www.undrr.org/publication/landslide-risk-reduction-great-lakes-region

USAID. (2014). Haiti Securite Alimentaire en Bref.

Wang, F., Fan, X., Yunus, A. P., Siva Subramanian, S., Alonso-Rodriguez, A., Dai, L., Xu, Q., & Huang, R. (2019). Coseismic landslides triggered by the 2018 Hokkaido, Japan (Mw 6.6), earthquake: spatial distribution, controlling factors, and possible failure mechanism. Landslides, 16(8), 1551–1566. https://doi.org/10.1007/s10346-019-01187-7

Wang, T., Liu, C., Saito, H., Nishimura, Y., & Wang, G. (2020). Landslide susceptibility mapping following the 2016 Kaikoura earthquake, New Zealand. Landslides, 17(11), 2617-2630.

Wilis, R. (2019). Impact Variability of Rainfall Intensity to Horticulture Productivity at West Pasaman Regency , Province of West Sumatera , Indonesia. Core.Ac.Uk, 7(10), 138–145.

Win Zin, W., & Rutten, M. (2017). Long-term Changes in Annual Precipitation and Monsoon Seasonal Characteristics in Myanmar. Hydrology: Current Research, 08(02). https://doi.org/10.4172/2157-7587.1000271

Xu, Q., Dang, K. T., Su, Y., Chen, W., & Wang, G. (2020). Characteristics of landslides triggered by the 2018 M7.5 earthquake in Papua New Guinea. Landslides, 17(1), 145-157.

Yang, Z., & Xu, C. (2022). Efficient Detection of Earthquake−Triggered Landslides Based on U−Net++: An Example of the 2018 Hokkaido Eastern Iburi (Japan) Mw = 6.6 Earthquake. Remote Sensing, 14(12), 2826. https://doi.org/10.3390/rs14122826

Yang, Z., Xu, C., & Li, L. (2022). Landslide Detection Based on ResU-Net with Transformer and CBAM Embedded: Two Examples with Geologically Different Environments. Remote Sensing, 14(12), 2885. https://doi.org/10.3390/rs14122885

Zulkarnain, I. (2016). Sumatra is not a homogeneous segment of Gondwana derived continel blocks: A New Sight based on Geochemical Signatures of Pasaman Volcanic in West Sumatera. Jurnal RISET Geologi Dan Pertambangan, 26(1), 1. https://doi.org/10.14203/risetgeotam2016.v26.271

SUPPLEMENTARY MATERIALS

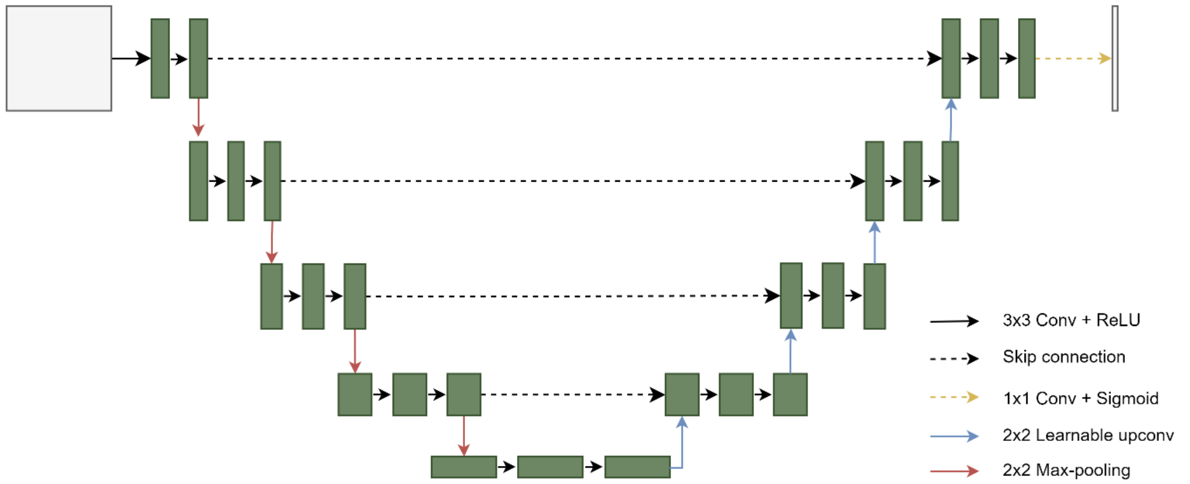


Figure S1~~3~~: Model architecture of the conventional U-Net.

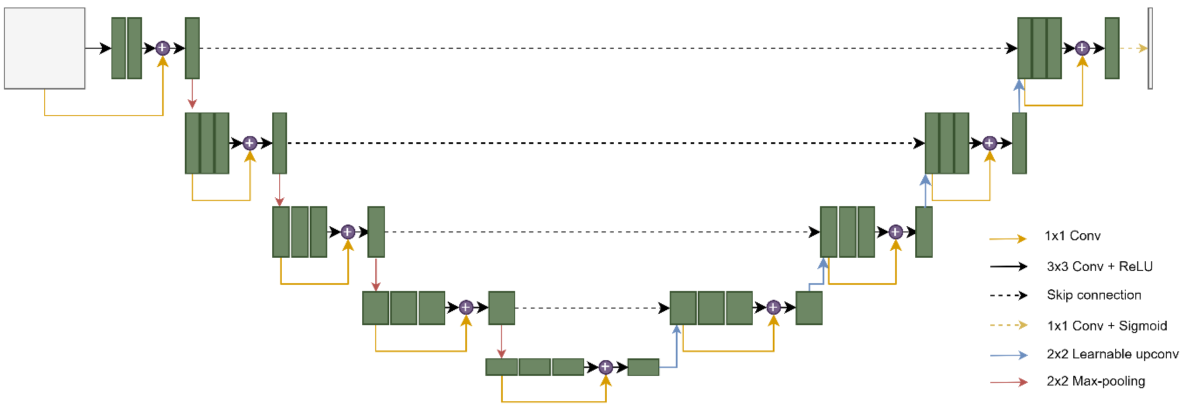


Figure S2~~4~~: Model architecture of the Res U-Net.

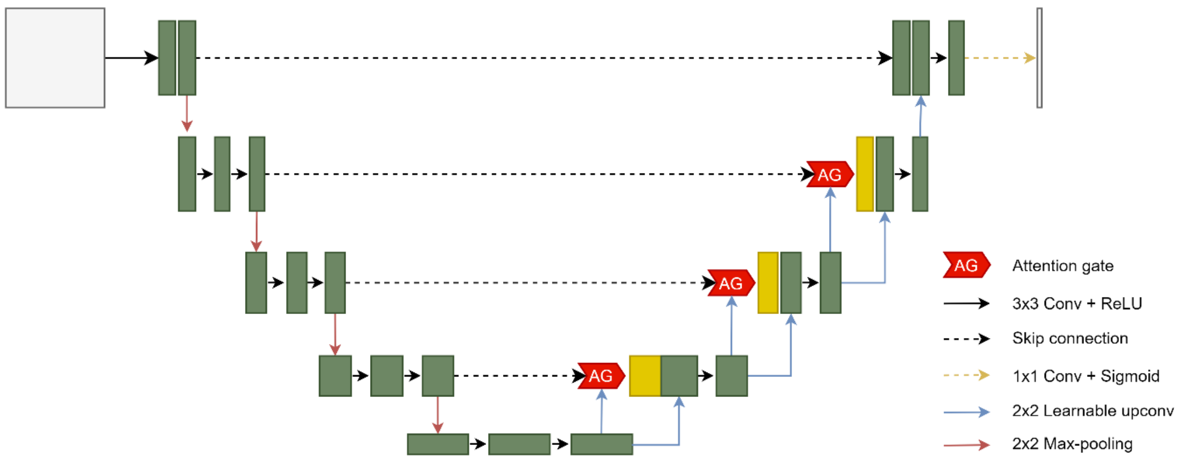


Figure S3: Model architecture of the Attention U-Net.

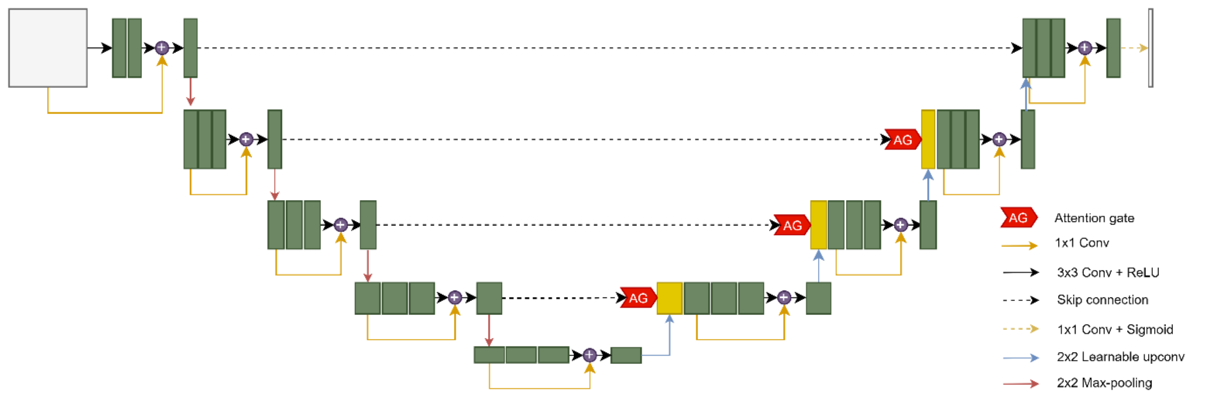


Figure S4~~5~~: Model architecture of the ~~(a) Attention U-Net and (b)~~ Attention Res U-Net.

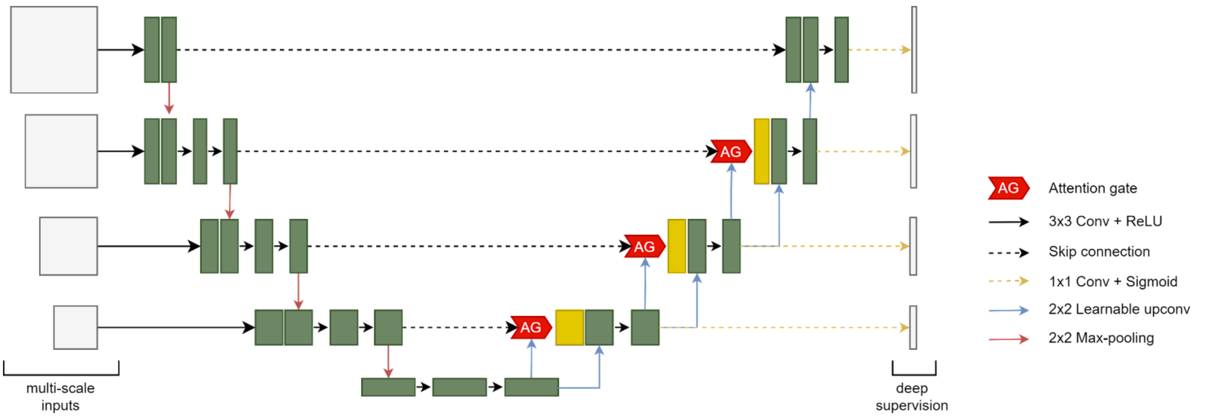


Figure S5~~6~~: Model architecture of the Attention Deep Supervision Multi-Scale U-Net.