Masters Program in Geospatial Technologies

Application of deep learning method in automatically detecting rainfall-induced shallow landslides in a data-sparse context

Roquia Salam

Dissertation submitted in partial fulfilment of the requirements
for the Degree of *Master of Science in Geospatial Technologies*

NOVA IMS Information Management School
UNIVERSITAT JAUME·I
Westfälische Wilhelms-Universität Münster

Supervised by:

Professor Filiberto Pla Bañón

Universitat Jaume I (UJI), Castellón de la Plana, Spain

Co-supervised by:

Dr Bayes Ahmed (Associate Professor)

Institute for Risk and Disaster Reduction,
University College London (IRDR UCL), UK

Professor Marco Painho

NOVA Information Management School
(NOVA IMS), Portugal

Declaration

I, **Roquia Salam** confirm that the work presented in this thesis is my own and has not previously been submitted elsewhere for any other degree or publication. Information or resources that have been collected from other sources are cited in this piece of work.

Roquia Salam

Universitat Jaume I (UJI), Castellón de la Plana, Spain

20 February 2024

Acknowledgments

I, Roquia Salam, am an Erasmus Mundus Scholarship holder funded by the European Commission. First, I would like to express my gratitude to my principal supervisor Professor Filiberto Pla Bañón (Universitat Jaume I, Spain). His right suggestions, helpful ideas, friendly spirit, intellectual guidance, and constructive criticism motivated me to complete the present task. He is always understanding and easy to work with, making my journey smoother. I am lucky to have him as my principal supervisor, and I truly believe that without his continuous guidance, it would not have been possible for me to complete this task. I am thankful to my co-supervisor, Dr Bayes Ahmed (Associate Professor at the University College London, UK), who provided an important dataset for this study. Moreover, his continuous guidance in completing my thesis work, and motivation which I truly appreciate has made my journey with this master's program easy. Without him I would not be here today; he has inspired me to apply to this master's program. Professor Marco Painho (NOVA Information Management School, Portugal), my other co-supervisor, has helped me improve the quality of my thesis by providing his insightful comments. His positive attitude brings positive energy during the tough situation. I am grateful to him for his assistance.

I am thankful to Professor Marco Painho (again), Professor Joaquín Huerta Guijarro, and Dr Christoph Brox for coordinating this joint master's program perfectly, making it a memorable experience academically, socially, and professionally.

Thank you to all the teachers from this program and my previous academic institutions for their inspiring lessons. Gloria Edo Piñana from Jaume I University was incredibly helpful with timely information, making life abroad easier.

My friend Mst Runa Khatun, along with Muhammad Ammad and Lukumon Olaitan Lateef from this master's program, provided me with mental support and courage. My other friends and classmates taught me valuable life lessons, making my journey adventurous and fulfilling.

I am deeply grateful to my parents for their love and encouragement, foundational to my success. My siblings have also been a constant source of motivation, pushing me to grow and improve daily.

Finally, I appreciate the open-source community for enabling access to data and resources that significantly contributed to my thesis.

Abstract

Detecting rainfall-induced shallow landslides in data-sparse contexts has become an environmental concern in recent decades and is crucial for a comprehensive landslide disaster management plan (CLDMP). Most of the previous works have contributed to the development of automated methods for detecting earthquake-triggered landslides. Despite the substantial contributions of researchers in this field, gaps and uncertainties still exist in developing a method for automatically detecting rainfall-induced shallow landslides. To address this gap, the present study has utilized the deep learning (DL) based U-net model for automatically detecting rainfall-induced shallow landslides from multi-temporal, very high-resolution (VHR) PlanetScope, medium resolution (MR) Sentinel-2 imagery, and ALOS PALSAR-provided digital elevation model (DEM), collected from the years 2018, 2019, 2022, and 2023. Four different data sets have been prepared for this study: Dataset A, comprising red, green, blue (RGB), and near-infrared (NIR) bands of PlanetScope imagery; Dataset B, comprising RGB and NIR bands of PlanetScope imagery with the inclusion of the normalized difference vegetation index (NDVI) calculated from the red and NIR bands, elevation, and slope derived from DEM; Dataset C, comprising RGB and NIR bands of Sentinel-2 imagery; and Dataset D, comprising RGB and NIR bands of Sentinel-2 imagery with the inclusion of NDVI, elevation, and slope. As a case study, the Chittagong Hill Tracts (CHT) of Bangladesh have been selected. For training the U-net model with ground truth data, 181 landslide polygons have been created from Google Earth Pro, which is a small set of ground truth data. So, the horizontal flip technique has been applied to augment the dataset, effectively doubling the entire dataset. Each dataset (A, B, C, and D) has been experimented with in 4 different trials utilizing the repeated stratified hold-out validation method so that all data is used as test data, to avoid biased results. Comparatively, Trials 1 and 2 contain a larger set of landslide training samples than Trials 3 and 4. Thus, 16 different experiments have been conducted in the present study. The performance of the U-net model is evaluated by precision, recall, F1 score, loss, and accuracy metrics. It is explored from the experiment that Datasets A and B perform the best; however, the integration of the DEM data does not enhance the accuracy of the model. The datasets comprised of Sentinel-2 imagery (Datasets C and D) exhibited very poor performance in all trials (4) in detecting rainfall-induced shallow landslides. Among the four Trials, utilizing Dataset A and B, Trials 1 and 2 outperformed, indicating the necessity of using larger training

samples for DL model implementation. The mean precision, recall, F1 score, loss, and accuracy based on Trials 1 and 2 are 1, 0.625, 0.625, 0.380, and 0.999, respectively (same results found in both Datasets A and B). Overall, the performance of the model indicates that the U-net model can be used to detect rainfall-induced shallow landslides across similar geographic regions and temporal contexts around the world.

Keywords: Rainfall-induced shallow landslides, PlanetScope imagery, Sentinel-2 imagery, ALOS PALSAR digital elevation model, Deep learning, U-net, Repeated stratified hold-out validation, Data-sparse context, Chittagong Hill Tracts, Bangladesh.

Table of Contents

List of Figures

List of Tables

List of Acronyms

ADSMS	Attention Deep Supervision Multi-Scale
AI	Artificial Intelligence
ALOS PALSAR	Advanced Land Observing Satellite Phased Array L-band Synthetic Aperture Radar
ASF	Alaska Satellite Facility
BBS	Bangladesh Bureau of Statistics
BC	Bayesian Classifiers
CDMP	Comprehensive Disaster Management Programme
CHD	Chittagong Hill Districts
CHT	Chittagong Hill Tracts
CLDMP	Comprehensive Landslide Disaster Management Plan
CNN	Convolutional Neural Network
CUN	Convolutional U-net
DCNN	Deep Convolution Neural Network
DEM	Digital Elevation Model.
DL	Deep Learning
ESA	European Space Agency
FN	False Negatives
FP	False Positives
ICA	Independent Component Analysis
InSAR	Interferometric Synthetic Aperture Radar
KNN	K-Nearest Neighbor
LR	Logistic Regression
LR	Low Resolution
MCC	Mathews Correlation Coefficient
ML	Machine Learning
MR	Medium Resolution
NASA	National Aeronautics and Space Administration
NDVI	Normalize Difference Vegetation Index

NIR	Near-Infrared
PCA	Principal Component Analysis
R-CNN	Region-based Convolutional Neural Networks
ReLU	Rectified Linear Unit
RF	Random Forest
RGB	Red, Green, and Blue
RS	Remote Sensing
SAR	Synthetic Aperture Radar
SRTM	Shuttle Radar Topography Mission
SVM	Support Vector Machines
SWIR	Shortwave Infrared
TN	True Negative
TP	True Positives
VHR	Very High Resolution
YOLO	You Only Look Once

1 Introduction

1.1 Problem Statement, Background, and Motivation

Landslides are a geological phenomenon that occurs when the hills slopes become unstable which leads to the downward movement of rock, soil, and debris under the force of gravity (Varnes, 1958). There are four main categories of landslides which are debris flow, rockfalls, rotational slides, and transitional slides. Different underlying factors, for example, heavy rainfall, volcanic activity, earthquake, human intervention act as the trigger of landslides (Paul, 2020). Landslides alter the terrain immediately, and also results in devastating negative impacts upon communities, societies, infrastructure, and environment (Ye et al., 2019).

All over the world, a significant number of extreme landslides have been occurred in rainy season in the areas of mountains (Kumar et al., 2017). These rainfall-induced landslides stand as recurrent and catastrophic disasters, with long-lasting impacts on people's lives, livelihoods, critical infrastructure, and sustainable development (Mondini et al., 2023; Amatya et al., 2022). These dynamic events are impacted and influenced by several natural factors like topography, geology, soil composition, and climate. Moreover, climate change has a profound negative impact, resulting in more frequent and severe rainfall-induced landslides nowadays (Das and Wegmann, 2022) as it (climate change) has contributed to the change of the world's hydrological cycle leading to remarkably heavy rainfall (Tabari et al., 2020). Therefore, it is essential to properly understand the mechanisms, that drives landslides, for reducing the risks associated with this disaster (Casagli et al., 2023). For minimizing the landslides disaster risks, it is very crucial to have a comprehensive disaster management plan (Thirugnanam et al., 2020).

Landslide inventories is the main resource for the formulation of a comprehensive landslide disaster management plan (CLDMP). Without landslide inventories, it is impossible to landslides risk and hazard assessment, landslides susceptibility mapping, land use planning, and sustainable infrastructure development which are the components of CLDMP. However, the mechanism of creating landslide inventories is not same all over the world. There are some countries, like Italy, USA, Poland, Austria, Norway, and so on, in the world where there are strong and dedicated teams or groups from specific departments under recognized institutions who only work for creating or collecting landslide inventories by using different means (Karimi et al., 2019; Herrera et al., 2018). Moreover, the characteristics of landslides are different in these areas (mostly earthquake-

triggered, dry, larger size deep seated landslides, and so on) that can be even identified, from the satellite image or field visits, long time after the events occurrence. Normally, these are developed countries having no resource shortage issue to create the inventories. From these dedicated departments, it is possible to collect the landslide inventories data, and implement the CLDMP. On the other hand, there are some countries or places in the world where there is no or proper institutional mechanism present to collect or create landslide inventories, although they are recognized as landslide prone countries like Bangladesh, Myanmar, India, New Zealand, Rwanda, Ethiopia, Uganda, Democratic Republic of the Congo, Burundi, and so on (Meena et al., 2023; Rabby & Li, 2019; Monsieurs et al., 2018). Landslides in these countries are mostly rainfall-triggered wet landslides that can be sometimes difficult to identify, from the satellite image or field visits, because of the presence of heavy rainfall associated with high percentage of cloud cover, revegetation issue, shallow seated landslides, and so on (Rabby & Li, 2019; Monsieurs et al., 2018). Moreover, these countries have not enough resources to create the landslide inventories from field visits. For these data-sparse rainfall-induced shallow landslides prone countries, there is a need to have an alternative option for creating landslide inventories. An automated landslides detection method by utilizing present available resources can be a solution for creating landslide inventories in these data-sparse countries to battle against the landslide disasters.

Landslides detection which is also known as landslide inventories has become easier due to the recent technological advancement of remote sensing (RS) (Lu et al., 2023). In recent years, RS is proven as the principal investigative tools for landslides detection promptly without field visits. In this case, there are several techniques existed for detecting landslides by utilizing RS which are visual interpretation, change detection-based approaches, knowledge-based methods, machine learning techniques, and deep learning strategies (Novellino et al., 2024). Although, visual interpretation is an effective technique, it requires substantial professional expertise as well as it is resource-intensive (Zhang et al., 2019).

Amatya et al. (2022) have developed a semi-automated method by modifying three change detection methods, based on pre- and post-satellite images collected from two different sources, to detect landslides in Vietnam, Laos, Myanmar, and Thailand. They have utilized Principal Component Analysis (PCA), Normalize Difference Vegetation Index (NDVI), and Independent Component Analysis (ICA) for detecting the change and thus, to detect landslides. As change

detection-based methods require both pre-event and post-event satellite images, which can be challenging to acquire promptly, especially because of the weather conditions it is very challenging to obtain the necessary images just after the rainfall-induced landslides have occurred (Lu et al., 2023). Yi et al. (2023) have utilized the interferometric synthetic aperture radar (InSAR) method to develop a semi-automatic method to detect landslides in a part of the Qinghai–Tibet Plateau. Although, synthetic aperture radar (SAR) data is not affected by weather conditions, its applicability to detect landslides is not extensive because it usually produces lower accuracies (Deng et al., 2023). Many previous studies have been carried out by utilizing the knowledge-based methods, for instance, threshold segmentation, object-based segmentation, and image enhancement (Tehrani et al., 2022). However, these methods are restricted to some specific places because of their limited transferability.

Machine learning (ML) has emerged as a powerful automated classification approach in recent decades. There are several ML methods that scholars extensively use for different purposes including support vector machines (SVM), random forest (RF), logistic regression (LR), Bayesian classifiers, and so on. For instance, Meena et al. (2022) have applied four machine-learning models RF, SVM, Convolutional U-net (CUN), and K-Nearest Neighbor (KNN) to develop an automated landslide detecting method based on satellite images collected from two different satellites. However, the performance of ML methods depends on the clarity of input features and requires rigorous data preprocessing and feature engineering (Ma et al., 2021).

Deep learning (DL) is a cutting-edge technique that can provide reliable solutions for environmental issues by utilizing the RS (Abdollahi et al., 2022). The integration of DL methods with both object-based and pixel-based approaches addresses challenges associated with conventional approaches, like manual feature extraction and limited adaptability to dynamic environmental conditions (Chen et al., 2018). Both data-driven and automated solutions are possible by utilizing DL with RS, which is flexible to varied landscape and temporal scales. Landslides detection can be improved significantly by increasing accuracy and reliability with the utilization of DL approach together with the satellite imagery, thus offering a comprehensive perspective (Lu et al., 2020).

There are several object-based DL methods used in landslide detection, for instance, Region-based Convolutional Neural Networks (R-CNN) and You Only Look Once (YOLO), however, these

methods usually make use of the rectangular bounding boxes in the satellite imagery to label the position of the landslides (Lu et al., 2023). Therefore, these object-based methods cannot detect the exact boundaries of landslides due to using rectangular bounding boxes for detecting landslides. This is regarded as a limitation of this method as it is crucial to identify the precise delineation of landslide boundaries for a comprehensive understanding of the landslide-affected areas and their associated impacts on the affected areas (Keyport et al., 2018).

Conversely, pixel-based DL methods for landslides detection outperform object-based DL methods due to their characteristic, fine-grain spatial resolution and their applicability to use in varied landscape conditions (Wen et al., 2021). For instance, the pixel-based popular DL technique convolutional neural network (CNN) shows superior performance in capturing detailed variations in terrain automatically which is very crucial for identifying the early signs of landslides as well as accurate delineation of their extents (Lu et al., 2023). Feature engineering which is a manual intervention, labour intensive, and requires expert knowledge; now is replaced by CNNs as it can extract the relevant features automatically that are associated with landslides contributing to the advancement of RS and geospatial analysis in the field of the assessment and mitigation of natural hazards (Su, 2021). Thus, pixel-based methods contribute to the development of an accurate and robust landslide detection approach by utilizing their capability to extract, in detail, values, textures, and other spectral characteristics of each pixel which is a must to identify the pattern of the landslide's prone pixels (in other words landslides prone areas) (Meena et al., 2023; Ghorbanzadeh et al., 2020). Therefore, the continuous integration of pixel-based DL methods with RS contributes to improving precise landslides detection. For these above-mentioned reasons, recently, pixel-based DL methods have emerged as the most preferred, reliable, and accurate in detecting landslides, especially in diverse topographic regions and varied environmental conditions (Meena et al., 2023).

However, a few important things have influenced the performance of the DL models. First, an extensive dataset is required to detect an object using DL (Meena et al., 2023). Furthermore, spatial resolution is one of the most important characteristics of satellite imagery which is very crucial in detecting an object (Lu et al., 2023). Generally, very high resolution (VHR) imagery performs better than both medium resolution (MR) and low resolution (LR) imagery (Ghorbanzadeh et al., 2022; Lin et al., 2022; Meena et al., 2021). Additionally, certain topographical variables, for

instance, elevation, and its derivatives including slope, aspect, and curvature, have been investigated in conjunction with the RS data for the enhancement of landslide detection (Prakash et al., 2020). Moreover, a multi-temporal dataset can generate more reliable results in detecting landslides (Bhuyan et al., 2023). Consequently, this study conducts an experiment encompassing the four aforementioned aspects.

1.2 Research Gap

Among the data-sparse countries prone to rainfall-induced shallow landslides, Bangladesh is one of them. Either government or any private organization does not keep record of the historical landslide inventories in Bangladesh (Ahmed et al., 2014). Because of this creating landslide inventories is challenging in Bangladesh which is exacerbated by three main factors:

- First, the manual collection of landslide inventories is excessively costly and time-consuming (CDMP, 2012).
- Second, the extremely hilly terrain becomes inaccessible during the heavy monsoon season when landslides are most likely to occur (Abedin et al., 2020).
- Third, conflicts in the hilly areas of Bangladesh pose life-threatening risks, hindering traditional data collection methods (Ahmed et al., 2014).

In light of these unique circumstances, there is an urgent need for an automated landslide inventory method. Remarkably, despite the significance of this issue, no study has been conducted, previously, in Bangladesh to detect landslide inventories using an automatic method, highlighting a critical gap in environmental science research. Recognizing this gap, this study selects Bangladesh as a case study for developing a reliable automatic method for detecting rainfall-induced shallow landslides that incorporates advanced geotechnical and geospatial techniques to overcome the limitations of traditional approaches and contribute to CLDMP.

1.3 Objectives of this Study

This study aims to develop an automated DL based method for detecting rainfall-induced shallow landslides by leveraging the RS in Bangladesh. Therefore, to achieve this goal, the following intermediate objectives are defined:

a. Find the appropriate DL model for landslides detection in Bangladesh.

b. Select the data structure based on the previously conducted similar works in similar contexts.
c. Develop the DL based model for the development of the automated method for landslides detection by leveraging both VHR PlanetScope and MR Sentinel-2 imagery.
d. Experiment with both small and large number of landslides training samples to know the performance of the model.
e. Evaluate the performance of the model with both VHR and MR imagery in different experimental samples using several statistical metrics.
f. Find the best-performed dataset in detecting landslides in the study area.

1.4 Importance and Contribution of this Study

The development of this automated method employing the DL model for rainfall-induced shallow landslides detection is of paramount importance in the realm of landslide research and hazard mitigation. This automated approach ensures efficiency and rapid analysis of extensive datasets, facilitating the timely identification of shallow landslides triggered by rainfall in the study area. Detection of landslides is essential to create a landslide inventories. Landslide inventories are the main basis for risk and hazard assessment, landslides susceptibility mapping, land use planning, and so on. The data-sparse area where no proper institutional mechanism is present to create the landslide inventories needs an alternative option which consumes less time, less labour-intensive, cost-effective, based on the publicly available resources, and automated. Therefore, an automated method based on DL is a solution to this issue.

Thus, the developed automated DL based approach for rainfall-induced shallow landslides contributes to make a landslide inventories. Consequently, ease the whole process of CLDMP in the landslides affected area with minimal effort which, ultimately, reduces the overall impacts upon the affected community and environment created by landslides.

Additionally, the automated method holds promise for the development of early warning systems, enabling rapid evacuation and early mitigation measures to reduce potential societal and infrastructural impacts. Overall, the adoption of an automated DL model stands at the forefront of scientific advancements, in landslide detection, and in the broader domains of RS, ML, and geospatial analysis. By minimizing subjectivity and ensuring a standardized methodology, the

automated DL model contributes to consistent and reliable results across similar geographic regions and temporal contexts around the world.

1.5 Outline of this Work

Here, the research and development process has been described with a view to develop an automated DL based method for detecting rainfall-induced shallow landslides in Bangladesh. The present work can be divided into 3 broadly parts which are exploration, development and evaluation. Figure 1 presents the simple breakdown of the present work.

Section one is the introduction section which has divided into sub-sections namely problem statement, background and motivation; research gap; objectives; importance and contribution, and outline of the study.

Section two is the study area section describing the suitable area where this study has been conducted.

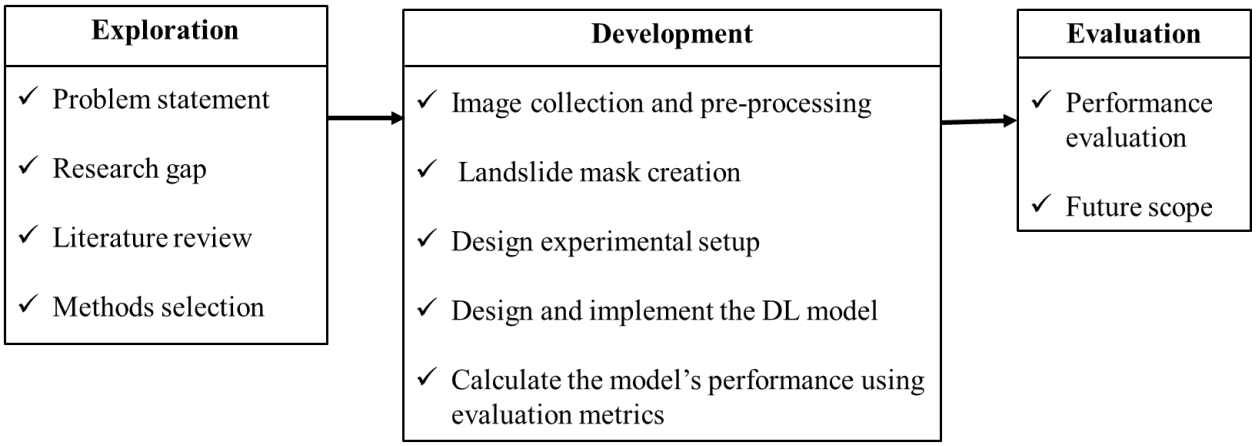

Figure 1: Breakdown of the present work.

Section three includes the methodology part that finds the suitable structure for this study from literature review and the main methodological design of the DL model, loss function, performance evaluation metrics, and data augmentation.

Section four describes the experimental setting of this study which includes the types of data used in this study and how the data is used in different experiments during the implementation of the

DL model. This section is divided into satellite data, ground truth data, repeated stratified hold-out validation, and experimental setup.

Section five is the results and discussion section that contains the experimental results, scholarly discussions for the selection of the best performed dataset, limitations of this study.

Section Six is the conclusion section that includes summary, and future research focus.

2 Study Area

The Chittagong Hill Districts (CHD) is the only extensive hilly region in Bangladesh prone to landslides, situated in the southeastern part of the country, known for its unique geographical and cultural features. This CHD is further divided into two parts based on the types of inhabitants residing there which are: i) urbanized hill districts (comprise Chittagong, and Cox's Bazar districts) and Chittagong Hill Tracts (CHT) (comprises Bandarban, Khagrachari, and Rangamati districts) (Figure 2) (Ahmed, 2021). Mostly, Bengali, and Rohingya refugees (who came from Myanmar because of the ongoing genocide and violence in their country) are residing in urbanized hill districts, and indigenous people are living in CHT (Ahmed, 2021). The area is surrounded by Bay of Bengal from its south and southwestern part; Myanmar by its southeastern part; and India by its north and northeast (Figure 2). The total area and population of the CHD are 19,888 km^2, and 13,835,544 (including both Bengali and indigenous), respectively (BBS, 2023). According to the recent report made by Bangladesh Bureau of Statistics (BBS, 2023), in this area, the average population density is 979 per km^2.

The whole Bangladesh including the CHD is characterized by a tropical monsoon climate with three distinct seasons, dry (November to March), pre-monsoon (April to May), and monsoon (June to October). This area has an average annual monsoon (June-October) rainfall ranging from approximately 2540 to 3810 mm. Undulating hills are the main feature of this area, ranging from the Semutang anticline with an average elevation of 80-160m to the Bandarban anticline with a mean elevation of 650-800m (Banglapedia, 2021). These divides the area into two distinct hilly ranges namely low hill (Dupi Tila and Dihing formation where the elevation is less than 300m), and high hill (Surma and Tipam formation where the elevation is more than 300m) ranges (Brammer, 2012). Moreover, this area is characterized by unconsolidated or semi-consolidated tertiary sedimentary rocks made of sandstone, shale, siltstone, limestones along with pebbles and cobbles (Alam et al., 2003; Brammer, 1986).

Additionally, dense forest and river valleys are other common features in this extreme hilly region. It is essential to note that various social and natural issues are prevalent in the CHD, including land disputes, displacement, and environmental issues and so on which should be the subjects of research and development for the well-being of the local communities (Ahmed, 2021). This unique geological, lithological, and geomorphological characteristics of weak lithology, steep slopes,

valleys together with the intense and prolonged monsoon rainfall contribute to the instability of the slopes that ultimately increase the risk of frequent and severe landslides (mostly debris flow and rock slides) in this area. Rapid deforestation and unplanned urbanization in this region have intensified the susceptibility of landslides in recent decades (Ahmed, 2021). For instance, on the 13th of June 2017, landslides induced by intense rainfall resulting in a minimum of 168 fatalities in Rangamati, Chittagong, and Bandarban districts (Ahmed et al., 2018).

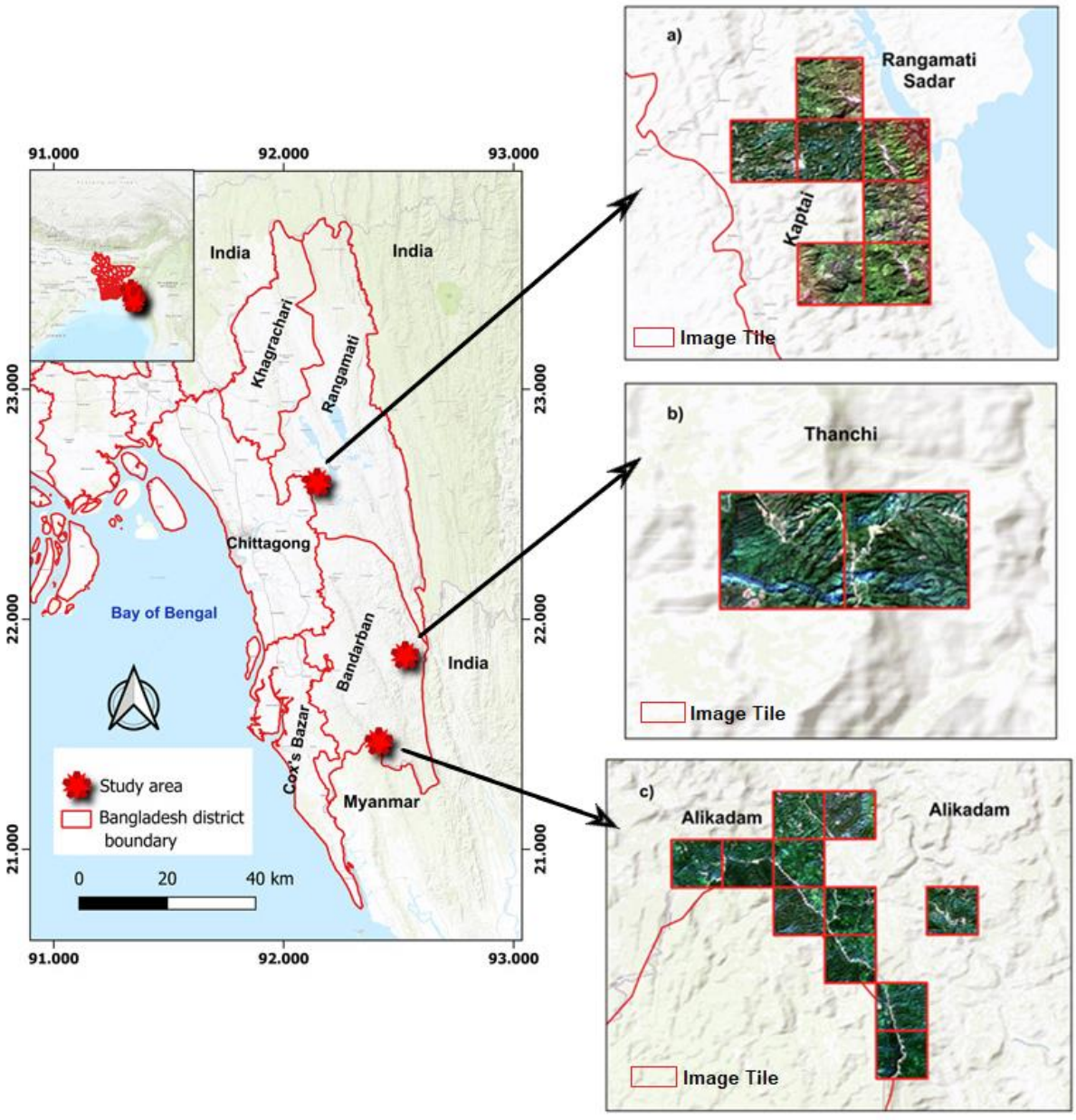

Figure 2: Map showing the geographical location of the study area.

A significant number of families had to take shelters in different shelter houses. Up to date, the landslides in 2017 is recognized as the most perilous landslide disaster in the history of Bangladesh (Ahmed et al., 2018). Recently, in August 2023, about 55 people have lost their lives just because of the landslides and floods disasters in Bandarban, Rangamati, Cox's Bazar, and Chittagong districts (The Hindu, 2023).

The consequences of these rainfall-induced landslides are not limited to human casualty; however, these also have significant negative effects on the environment, human settlements, infrastructure, and agricultural lands. Therefore, to address this challenge, a CLDMP is essential, encompassing sustainable land-use planning, afforestation initiatives to stabilize slopes, and the implementation of forecasting systems. Without having the landslide inventories, it is not possible to formulate any comprehensive plan. Unfortunately, there is no institutional mechanism in Bangladesh to create the landslides inventories either manually or automatically based on the available modern technologies and methods. Therefore, this study aims to contribute to the development of an automatic rainfall-induced landslides detection method.

Table 1: Spatiotemporal information of the satellite imagery used in this study.

Number of image tiles	**Acquisition date (PlanetScope Imagery)**	**Acquisition date (Sentinel 2 Imagery)**	**Upazila**
8	27 October 2023	20 October 2023	Alikadam
3	23 December 2022	24 December 2022	Alikadam
1	16 November 2018	15 November 2018	Kaptai
4	16 November 2018	15 November 2018	Rangamati Sadar
2	24 January 2019	11 January 2019	Rangamati Sadar
2	8 January 2022	8 January 2022	Thanchi

It is already mentioned that Chittagong, and Cox's Bazar are the urbanized hill districts where rapid urbanization by cutting down hills is common (Ahmed, 2021); so, when landslides occur in these areas, it is difficult to differentiate the small size shallow landslides, from the surrounding barren areas dedicated for the development purposes, from the satellite image with both medium

and high spatial resolution (Sofi, 2022). As, this study utilized MR imagery, these urbanized hill districts are excluded from the area considering as a study area. Therefore, only CHT is considered for this present work. However, the CHT is not a small area, it is an area of 11,476.26 sq. km (BBS, 2017). In this study, apart from MR Sentinel-2 imagery, VHR PlanetScope is also used. So, it is difficult to manage the cloud-free both MR and VHR imagery on the same or closer date for the entire the CHT. Additionally, landslides are not evenly distributed in all parts of CHT. Similarly, not all area is affected with the same extend. Furthermore, this study aims to use multi-temporal satellite data, and landslides does not occur in the same place in every year inside the CHT rather it occurs in different places in different years; so, data has been collected from different parts of CHT. Therefore, the area is selected in a tricky way based on the most recent landslides events occurred in the CHT.

In 2022 and 2023 the area adjacent to the roads in Alikadam upazila is severely affected (bdnews24, 2023; Nafiu, 2021); Rangamati Sadar, Kawkhali and Kaptai were the most affected upazilas by landslides in 2017 (Sifa et al., 2020); and Thanchi is one of the severe affected upazilas by landslides in every year (Sofi, 2022). Therefore, 4 different upazilas from Rangamati (Rangamati Sadar and Kaptai upazilas) and Bandarban (Alikadam, and Thanchi upazilas) districts have been selected as a case study. The number of image tile collected from each upazila is varied based on the available scars of landslides in those selected upazilas. There are 20 image tiles, each one covers an area of 4.002859 km^2, have been used in the present study (Table 1).

Table 1 demonstrates the acquisition date and the specific location from where the different image tiles have been extracted. In total, 11, 6, 2, and 1 image tiles have been extracted from Alikadam (Figure 2c), Rangamati Sadar (Figure 2a), Thanchi (Figure 2b), and Kaptai (Figure 2a) upazilas, respectively (Table 1). It is also presented in the Table 1 that the 20 image tiles have been extracted from 4 different years:

- 2018 (5 image tiles),
- 2019 (2 image tiles),
- 2022 (5 image tiles), and
- 2023 (8 image tiles).

3 Methodology

3.1 Literature Review

Scholars and researchers have been investigating to detect landslides by leveraging advanced geospatial techniques and RS imagery since the beginning of this decade. These days, scientists have emphasized on the pixel-based methods as pixel-based methods can delineate the exact boundaries of landslides (Lu et al., 2023; Su et al., 2021). Therefore, in this section, the studies which are conducted based on pixel-based DL models are discussed. In a simpler way, DL is a subset of ML, and ML is a subset of artificial intelligence (AI) (Figure 3). Thus, DL is an AI technique which simulates human brain's processes. While, ML requires human intervention to learn the features from the input data for making prediction, DL learns the features automatically from the input data by utilizing the multi-layer neural networks, and make the prediction which ease to execute the complex tasks (Bengio et al., 2017).

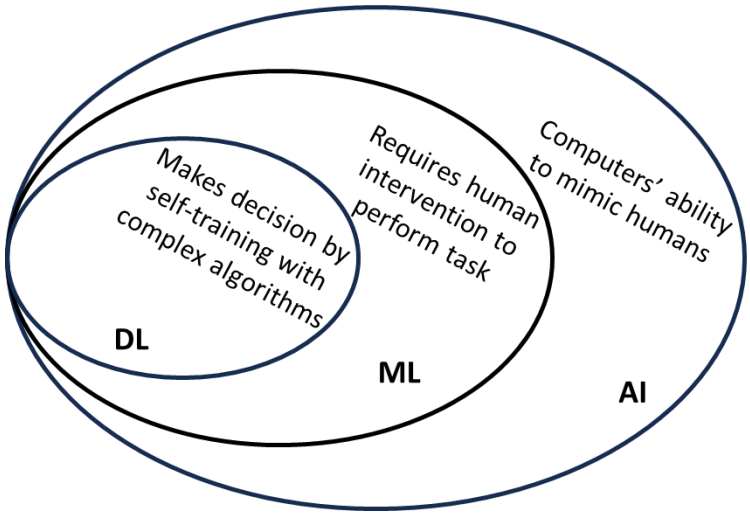

Figure 3: Relation among DL, ML, and AI.

This literature only covers those studies which are conducted in those areas dominated by tropical or sub-tropical monsoon climate; affected by mostly rainfall-induced landslides; and the lithology and geology of those areas are mostly influenced by the sedimentary rocks. A short review of the datasets, models used, lithology, geology, climate, and other factors, for the purpose of detecting landslides in different regions of the world, is presented in Table 2. Bhuyan et al. (2023) have proposed a DL based technique called Attention Deep Supervision Multi-Scale U-net model for

creating multi-temporal landslide inventories, they basically selected the earthquake triggered landslides. In this case, they have used VHR PlanetScope Imagery dated from 2009 to 2021. It is found in their study that the proposed model can produce reliable and very accurate multi-temporal landslide inventories. Meena et al. (2023) have created a global dataset of landslide inventories based on VHR PlanetScope imagery utilizing several DL based models like U-net, Res U-net, Attention U-Net, Attention Res U-Net, and Attention Deep Supervision Multi-Scale (ADSMS). They have experimented in 11 different countries with diverse geomorphology. In all countries, U-net has produced the higher F1 score over other models' indicating the most suitable DL based model applicable for any area of the world.

Chandra et al. (2023) utilized the VHR TripleSat imagery for landslides detection with U-net, ResNet-50, ResNet-101, DenseNet-121, and VGG-19. Based on the accuracy, precision, and F1 score, U-net has been explored as the best model for the detection of landslides. Lu et al. (2023) proposed a dual encoder U-net and compared the performance of this proposed model based on the F1 score to the general U-net, Seg-Net, and Attention U-net models for detecting landslides induced by severe earthquakes. Dual encoder U-net which incorporates the topographical data derived from the digital elevation model (DEM) through one encode and optical sentinel-2 imagery through another encode has been found as the best method for the specific area to detect landslides. Xia et al. (2022) have experimented with CNN, AlexNet, ResNet152V2, DenseNet201, InceptionV3, Xception, and InceptionResNetV2 with the integration of Sentinel-2 imagery. They have explored that among the 7 DL based models, the DenseNet201 has performed the best by producing lower error to detect the landslides, in this case, they have used the evaluation metrics F1 score, precision, and recall. Meena et al. (2022) have experimented with two different datasets which are VHR RapidEye imagery and RapidEye imagery with the inclusion of the topographical data derived from Advanced Land Observing Satellite Phased Array L-band Synthetic Aperture Radar (ALOS-PALSAR). In their study, they have used 3 ML models SVM, RF, and KNN along with DL based U-net model. It is revealed in their study that the U-net model outperforms the other algorithms, by assessing the accuracy based on the metrics named F1 score, precision, recall, and Mathews Correlation Coefficient (MCC).

Bragagnolo et al. (2021) also found U-net as the robust model for landslides detection in Nepal. Wang et al. (2021) applied CNN along with 4 other ML algorithms SVM, LR, RF, and boosting to

compare the performance, based on the metrics of recall, precision, F1 score, accuracy, MCC and specificity, of these models for landslides detection. They have found CNN as the outperformed model, although the rest have produced considerably reliable results. Chen et al. (2018a) utilized the change detection techniques of Deep Convolution Neural Network (DCNN) for detecting landslides using pre and post event satellite data. Their accuracy assessment based on the true positive, false positive, and false negative detection of landslides explored that the model can be competitive for landslides detection.

3.2 Selection of the DL model for this study

It is found from the literature review that in the last few years, a number of researchers have conducted researches for the development of an automatic method to detect landslides (Table 2). Those studies have been conducted in different regions of China, India, Myanmar, Indonesia, Haiti, Brazil, New Zealand, Japan, Nepal, Papua New Guinea, Congo, and Hong Kong dominated by tropical or subtropical monsoon climate; steep slopes; the land areas characterized by sedimentary, metamorphic, and volcanic rocks; and are affected by specially debris flow, rock falls, rock slides, mudflows, rock avalanches. From the literature review it can be concluded that although several DL models have been tested for landslides detection automatically, the **U-net model** outperformed other models in those areas characterized by tropical or subtropical monsoon climates; sedimentary rocks; and rainfall-induced both slides and debris flow landslides (these characteristics are similar to the present study area) (Table 2).

Most of those studies (which found U-net as the best model) utilized the VHR satellite imagery from PlanetScope, RapidEye, and TripleSat imagery. ALOS-PALSAR DEM is also used to test its influence on the accuracy of the U-net model. Considering these factors, the present study is also employed U-net model for landslides detection by leveraging both VHR PlanetScope, MR Sentinel-2, and ALOS-PALSAR DEM imagery.

Table 2: Short review of the recently published works for automatic landslides detection.

Authors	Primary data	Experimental model	Region	Lithology, and Geology	Type of landslides	Under tropical / sub-tropical and monsoon climate	Performance evaluation metrics	Best performed model
Bhuyan et al. (2023)	PlanetScope Imagery	Attention Deep Supervision Multi-Scale U-net	Rasuwa district of Nepal; Wenchuan County, Sichuan Province of China; western part of Papua New Guinea; and Kaikōura region of New Zealand	Metamorphic rocks namely schists, amphibolite, phyllites, metasandstones, gneiss, schists, and shales; a few numbers of sedimentary rocks namely sandstones and limestones; volcanic rocks including granites and diorites volcanic rocks; and intermediate to mafic volcanic rocks	Rock falls, rock avalanches, and bed rock slides	Yes	Precision, recall, and F1 score	Attention Deep Supervision Multi-Scale U-net
Meena et al. (2023)	PlanetScope Imagery	U-net, Res U-net, Attention U-Net, Attention Res U-Net, and ADSMS	Porgera, Papua New Guinea; Kodagu, India; Tiburon Peninsula, Haiti; Rasuwa, Nepal; Hokkaido, Japan; Wenchuan, China; Sumatra, Indonesia; Longchuan, China; Hpa-An, Myanmar; Kaikoura, New Zealand; Uvira, Democratic Republic of the Congo	Metamorphic rocks like amphibolite, gneiss, and schist; volcanic rocks namely andesite and basalts; and Mesozoic, Paleozoic and Neogene sedimentary rocks	Mudflows, rockfalls, rock avalanches, and debris flows	Yes	Dice loss, and F1 score	U-net
Chandra et al. (2023)	TripleSat imagery	U-net, ResNet-50, ResNet-101,	Bijie, Guizhou Province, City in China	Permian coal	Debris flow, rock slides, and rockfalls	Yes	Overall accuracy,	U-net

		DenseNet-121, and VGG-19					precision, and F1 score	
Lu et al. (2023)	Sentinel-2 imagery and DEM	Dual encoder U-net, U-net, Seg-Net, and Attention U-net	Wen-Du, Sichuan province, China; Iburi, Hokkaido, Japan	Igneous and clastic rocks	Rock slides	Yes	F1 score	Dual encoder U-net
Xia et al. (2022)	Sentinel-2	CNN, AlexNet, ResNet152V2, DenseNet201, InceptionV3, Xception and InceptionResNetV2	Wenchuan county, China	Luvisols and Cambisols	Rock falls, rock slides, rock avalanches, and debris flows	Yes	Precision, recall, and F1 score	DenseNet201
Meena et al. (2022)	RapidEye imagery; ALOS-PALSAR DEM with RapidEye imagery	SVM, RF, KNN, and U-net	Rasuwa district, Nepal	Metamorphic rocks namely amphibolite, gneiss, and schist	Debris flow	Yes	Precision, recall, F1 score, and MCC	U-net
Bragagnolo et al. (2021)	Landsat 8	U-net	Nepal	Sedimentary rocks namely shale, limestone, sandstone; and metamorphic rock;	Debris flow, rock slides, and rock falls	Yes	Precision, recall, and F1 score	U-net
Wang et al. (2021)	Rainfall, geology, and DEM	SVM, LR, RF, boosting, and CNN	Lantau island, Hong Kong	Sedimentary rocks namely sandstone, siltstone, and mudstone; and volcanic rocks namely tuff, lava, and granite	Debris flow, and rock slides	Yes	Precision, recall, F1 score, and MCC	CNN
Chen et al. (2018a)	Landsat 8, Formosat 2, and DEM	DCNN	Shenzhen, Zhouqu Country, and Beichuan County, China	Sedimentary and volcanic rocks	Landfill slides along with debris flow	Yes	True positive, false positive, and false negative rate	DCNN

3.3 U-net Model

Pixel-based DL models also include the so-called image segmentation models. This is because the image which is provided to the model is divided into multiple segments or regions. Each segment is made of a number of pixels and each pixel is associated with a specific object or class. Image segmentation is crucial to the field where the precise size, shape, and location of an object in an image is important. Such fields include environmental monitoring with satellite imagery, vehicle monitoring and detection, medical imaging, and so on.

Generally, image segmentation is conducted by leveraging DL and CNNs are the most popular as well as widely used image segmentation model (Minaee et al., 2021). CNNs are a deep neural network, initially proposed by Fukushima et al. (1983). CNNs have specific three types of layers which are:

i) *Convolutional* layers for extracting features from the input image,
ii) *Pooling* layers which minimize the spatial resolution of the input feature for balancing the computational load (also known as downsampling), and
iii) *Deconvolutional* layers which actually retrieve the information that has been lost during downsampling through skip connection function (also known as upsampling).

There exist several CNNs models like U-net, SegNet, AlexNet, ResNet, and VGG. In the present study, landslides have been detected utilizing the U-net model (Figure 4), a fully CNN proposed by Ronneberger et al. (2015) specifically for the segmentation of medical images. Although U-net is developed for biomedical image segmentation, recently it has been utilized for landslide detection which produces outstanding results (Meena et al., 2022; Nava et al., 2022).

For this reason, this study is carried out by utilizing the U-net model by adopting the Python code available in GitHub (Meena et al., 2023), a slight modification of the code has been done for this research. In this case, the Google Colab Pro platform is utilized to conduct the main analysis part.

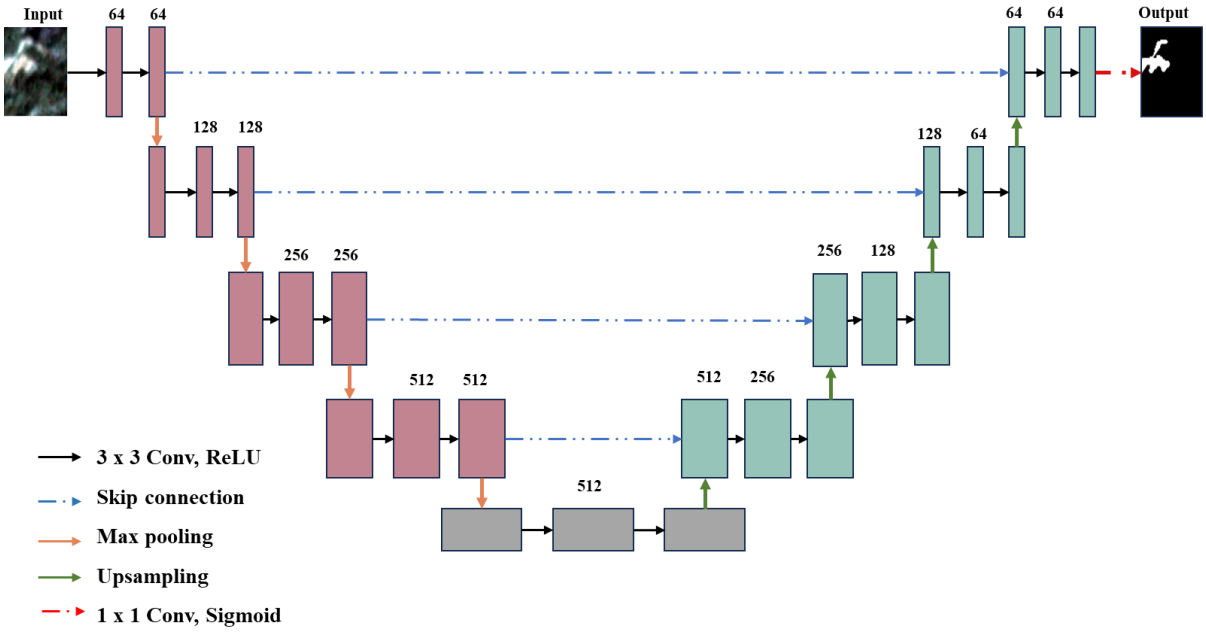

Figure 4: The U-net structure utilized in the present study.

The U-net model features an encoder-decoder architecture reminiscent of the letter "U" (Meena et al., 2022). The architectural configuration of the model used in this study incorporates a dual pathway system designed for the extraction of features (Figure 4). The encoder is specialized in capturing low-level representations and is similar to the organizational structure observed in the CNNs. Within the encoder path, individual convolution blocks are systematically arranged, each comprised of two convolutional layers employing a 3 × 3 kernel size, succeeded by a 2 × 2 max-pooling layer. Activation of these convolutional layers is executed by the rectified linear unit (ReLU) activation function. To introduce non-linear downsampling, a 2 × 2 max-pooling layer is appended after each convolutional block within the encoder route.

In contrast, the decoding path incorporates a 2 × 2 upsampling layer positioned after each 3 × 3 convolutional layers, strategically positioned to facilitate the reconstruction of high-level representations. In the final layer, a sigmoid function is employed to produce class predictions within a probability range of 0 to 1. As the main purpose of this study is to detect landslides and non-landslide areas (binary), the sigmoid function is used which is used only for the binary problem solution.

The model was meticulously trained using diverse combinations of hyperparameters values to find the adequate hyperparameters setting, including batch sizes of 4, 8, 16, and 32; learning rates of 0.01, 0.001, 0.0001, 0.005, and 0.0005; and varying number of filters in the first convolutional layer with 4, 8, 16, and 32.

3.4 Loss Function

An adequate loss function is an essential part of image segmentation to train DL models accurately and quickly, for learning the objects from the image. Scholars have used different types of loss functions which are suitable for the specific and respective tasks. There are four main categories of loss functions: distribution-based, region-based, boundary-based, and compound loss (Jadon, 2020). Among the above-mentioned four types of loss functions, the region-based loss functions outperform other loss models for image segmentation by minimizing the mismatch between the predicted and ground truth segmentations (Xu et al., 2023; Ma et al., 2021a). Moreover, the region-based loss function is more suitable and widely used in RS, more specifically in landslide detection (Meena et al., 2023; Li et al., 2022; Qin et al., 2021; Prakash et al., 2020).

Therefore, in this study, a region-based loss function is utilized to train the U-net model. There are five region-based loss functions which are Dice loss, Tversky loss, Focal Tversky loss, Sensitivity-Specificity loss, and Log-Cosh Dice loss (Jadon, 2020). However, all these five functions are not suitable for all problems. The input data which is used in this study is not balanced (will be discussed in the sections 4.3 and 4.4), meaning that between the two objects (landslides and non-landslides), the non-landslides object is dominant in the image. As well as, in some input images landslide objects are very minority, with a very small landslide area. For this reason, the input data is imbalanced. In this case, the **Dice loss** function performs better to deal with the imbalance dataset (Li et al., 2019). This loss function is used for binary classification in conjunction with the U-net model (Xu et al., 2023). Therefore, Dice loss is utilized in the present work for training the U-net model.

The Dice loss is derived from the dice coefficient, a statistical metric assessing the likeness between two sets of data (Xu et al., 2023). In the context of image segmentation, this coefficient quantifies the similarity between a predicted mask and its corresponding ground truth mask. Mathematically, the dice coefficient is defined in the following equation number 1 and graphically represented in Figure 5:

$$DSC(A,B) = \frac{2(A \cap B)}{A \cup B} \quad \text{.....1}$$

Where, DSC denotes dice coefficient of two different dataset, A is the predicted dataset, B represents the ground truth data, $\cap$ is the symbol of intersection, and $\cup$ is the symbol of union.

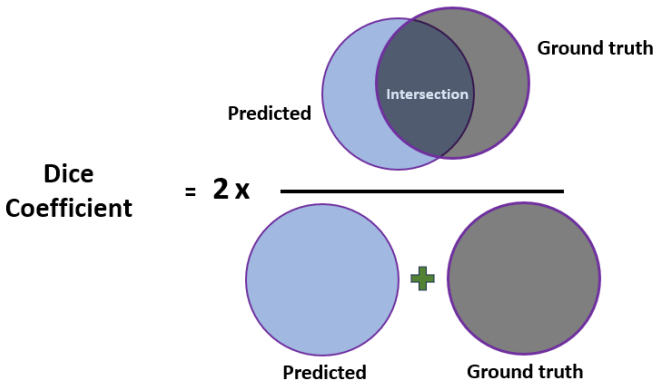

Figure 5: Pictorial representation of the dice coefficient.

“Intersection” means the number of overlapping pixels between the predicted and ground truth masks, and "Union" denotes the total number of pixels counting from both predicted and ground truth masks. The dice coefficient ranges from 0 to 1, with 1 indicating a perfect match.

The Dice loss is derived from the dice coefficient which is defined by the following equation:

$$Dice\ Loss = 1 - Dice\ Coefficient \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 2$$

The main objective of using the dice loss during the training of U-net model is to minimize the overall loss, so that the small-scale landslides can be detected more accurately.

3.5 Performance Evaluation Metrics

A confusion matrix is a table that is used to evaluate the performance of a classification model. It provides a comprehensive view of the model's predictions by comparing them to the actual class labels (Maxwell et al., 2021). The matrix is particularly useful when dealing with binary or multiclass classification problems. For a binary classification task, the confusion matrix typically has four entries: True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN) (Figure 6).

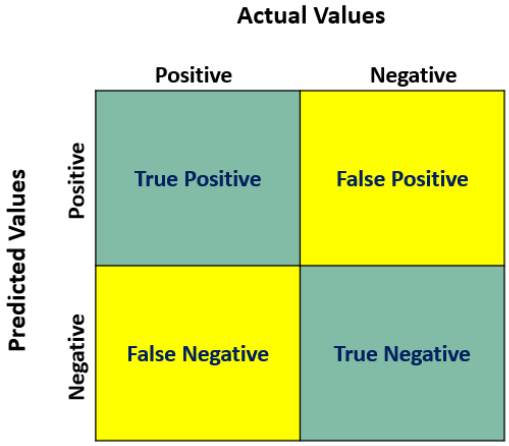

Figure 6: The confusion matrix defines the evaluation metrics used in this study.

From the confusion matrix, several performance metrics can be calculated, including precision, recall, F1-score, and accuracy. Accuracy alone cannot be used for a comprehensive performance evaluation of any model, therefore other metrics like precision, recall, and F1 score are used to deal with the imbalanced datasets as well as when there is a need to minimize the false positives or false negatives pixels. Therefore, in this study, those metrics have been used to evaluate the overall performance of the model to detect landslides following the equations used by Xu et al. (2024). The values of all the above-mentioned metrics range from 0 to 1 and are defined as,

$$Precision = \frac{TP}{TP+FP} \quad \text{..... 3}$$

$$Recall = \frac{TP}{TP+FN} \quad \text{.....4}$$

$$F1\ Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad \text{.....5}$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad \text{.....6}$$

Where, TP means true positives (indicates the pixels identified as landslides pixels which are actually landslides pixels), TN means true negative (indicates the pixels identified as non-landslides pixels which are actually non-landslides pixels), FN denotes the false negatives (indicates the pixels identified as non-landslides pixels which are actually landslides pixels),

and FP refers to false positives (indicates the pixels identified as landslides pixels which are actually non-landslides pixels).

For all the four above-mentioned metrics, the closer value to 1 the better, meaning that the model performs well detecting the landslides pixels. Conversely, values closer to 0 indicate the poor performance of the model to detect landslides pixels.

3.6 Data Augmentation

Data augmentation is the generation of artificial data based on existing observations. It serves as a DL strategy to enhance model accuracy and generalization, as well as address overfitting issues (Hao et al., 2023). The primary aim of employing data augmentation is to augment the dataset, and mitigating representational gaps (De et al., 2017). This strategy strengthens the model's capabilities, and also safeguards against biased learning across the entire system. In the context of image classification, particularly in satellite image classification, prevalent data augmentation techniques include clipping, rotating, flipping, and shifting, contributing to the effectiveness of various enhancement approaches (Taylor & Nitschke, 2018).

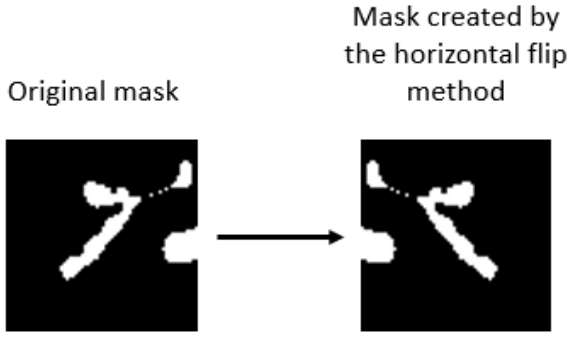

Figure 7: Horizontal flip data augmentation technique used in this work.

In this work, DL based U-net model is used and for the implementation of any DL model, it is very important to use as many data as possible to get precise results. As there is no prior landslide masks dataset prepared by anyone for the current study area, and there is a little information regarding the past landslide spots, a small set of landslide masks are created for the study based on the available information. Thus, the present study has employed the horizontal flip technique to augment the training dataset. Therefore, in each experiment, the number of training samples (original) appeared as double (for instance, if during training, 380 original patch numbers are fed to the model as input, during the implementation of the model this will be 760).

In the context of horizontal flip, each image in the dataset is horizontally mirrored, creating a new version of the image with reversed left and right orientations (Figure 7). This is achieved by flipping the pixels of the image along a vertical axis. By presenting the model with both the original and horizontally flipped versions of an image during training, the model becomes more robust and capable of recognizing patterns and features from different perspectives.

4 Experimental Setting

4.1 Satellite Data

The dataset used for the current study comprises imagery obtained from three distinct earth observation satellites, namely PlanetScope, Sentinel-2, and ALOS-PALSAR. Cloud-free 20 image tiles of every satellite were extracted from 4 different upazilas (described in the study area section) have been used in the present study.

Table 3: Information about different types of satellite data.

Name of the satellite image	Spatial resolution	Temporal resolution	Spectral bands (nm)	Collected source
PlanetScope	3m	Daily	8 bands: Coastal blue (431–552), blue (465–515), green (547–583), yellow (600–620), red (650–680), red-edge (697–713), and NIR (845–885) (Meena et al., 2023)	Planet Lab (https://www.planet.com/)
Sentinel-2	10m, 20m, and 60m	5 days	13 bands: Coastal aerosol (433-453), Blue (458-523), green (543-578), red (650-680), red-edge 1(698-713), red-edge 2 (733-748), red-edge (773-793), NIR (785-900), NIR narrow (855-875), water vapor (935-955), shortwave infrared (SWIR) (1360-1390), SWIR1 (1565-1655), and SWIR2 (2100-2280) (Hawryło & Wężyk, 2018)	Copernicus Data Space Ecosystem (https://dataspace.copernicus.eu/)
ALOS-PALSAR DEM	12.5m	46 days	L band (1.27 GHz) (Not related to spectral bands as it is radar data) (Khal et al., 2020)	Alaska Satellite Facility (https://search.asf.alaska.edu/)

The spatial and temporal resolution, spectral bands, and the source from where the images have been collected are presented in Table 3. It is already mentioned in the study area section that each image tile covers an area of 4.002859 km^2; therefore, from 20 image tiles, total an area of 80.05718 km^2 is used in this study. VHR imagery has smaller size pixels and contains more detailed information. Conversely, MR imagery has comparatively bigger size pixels and contains less information. Thus, considering the same area, VHR imagery has more pixels and MR imagery has fewer pixels.

For this reason, in this study, to cover the specifically 4.002859 km^2 area in each image tile, more pixels are contained in each VHR image tile than in the MR image tile which is discussed in the next two sub-sections. It is important to note that the present study area is characterized by different types of vegetation that usually create shade in those areas. Although, it is good enough to use the red, green, and blue (RGB) bands of a satellite image for landslide detection, it is difficult to distinguish landslides from the vegetation which are occurred in the shaded area (Meena et al., 2021). To mitigate this limitation, in this study near-infrared (NIR) band is also used, so that the NIR band can provide more detailed information like moisture content of the surface. Thus, these combined four bands RGB, and NIR can identify more accurately the landslides that are looking like vegetation because of the shade.

4.1.1 Very High Resolution (VHR) Imagery

The PlanetScope imagery is VHR imagery with 3m of spatial resolution, encompassing a total of 8 spectral bands (Table 3). However, in this research, only 4 spectral bands, namely RGB, and NIR, are utilized. Images have been collected freely from the official website of PlanetScope Lab under the Education and Research Standard Category. All image tiles maintain uniform dimensions, measuring 667 x 667 pixels. Notably, these images have undergone both orthorectification and radiometric correction as provided by the data providers. In the present study, the dataset comprises harmonized images, thereby ensuring a standardized and coherent foundation for subsequent analyses.

4.1.2 Medium Resolution (MR) Imagery

The Sentinel-2 imagery encompasses a total of 13 spectral bands (Table 3). However, in this study, only 4 spectral bands, namely RGB, and NIR, are utilized. Images have been collected freely from the Copernicus Data Space Ecosystem website. For this study, only MR imagery with 10m spatial resolution has been used. All image tiles maintain uniform dimensions, measuring 200 x 200 pixels. Particularly, these images have undergone both orthorectification and radiometric correction as provided by the data providers.

However, the 5 image tiles collected from the year 2018 on 15 November from Kaptai and Rangamati Sadar upazilas (Table 1) did not include atmospheric correction. Before employing these 5 image tiles for landslide mapping, atmospheric correction has been completed for these image tiles by employing the Sen2Cor plugin, accessible within the SNAP software provided by the European Space Agency (ESA). As all the spectral bands were downloaded separately, the band composition technique has been applied to merge the 4 bands into a single image

through the open-source software QGIS 3.34.1. This band composition has been applied for each image tile separately.

4.1.3 ALOS-PALSAR DEM

The ALOS-PALSAR satellite, through its ALOS-PALSAR, generates DEM data, offering insights into Earth's surface elevation and topography. The Alaska Satellite Facility (ASF) website facilitates direct access to the ALOS-PALSAR DEM, characterized by a spatial resolution of 12.5 m. In this study, the DEM sourced from the ASF website is utilized. Given the incorporation of images at spatial resolutions of 3m (PlanetScope) and 10m (Sentinel-2) in this study, the DEM undergoes resampling (using QGIS 3.34.1.) to align with both images of 3m and 10m spatial resolutions. Finally, elevation and slope have been extracted from the DEM which are also used in this study.

Table 4: Information related to the landslide polygons (masks) used in this study.

Number of image tiles (N = 20)	**Number of polygons created as landslide masks (N = 181)**	**Total area drawn as landslide polygons (m^2) (Total area 314,523 m^2)**	**Year**	**Type of the landslides**	**Upazila**
8	151	225,882	2023	Debris flow	Alikadam
3	11	11,538	2022	Debris flow	Alikadam
1	3	8,136	2018	Debris flow	Kaptai
4	5	5,256	2018	Debris flow and rock slides	Rangamati Sadar
2	4	49,527	2019	Rock slides	Rangamati Sadar
2	7	14,184	2022	Debris flow	Thanchi

4.2 Creation of the Ground Truth Data (Landslide Masks)

To accurately detect the landslides from satellite imagery, the model requires the incorporation of landslide masks (ground truth data) during its training and validation phases. It is important to note that among 4 major types of landslides (falls, topples, slides, and flows), slides and debris flow are the dominant types of landslides in CHD (Rabby & Li, 2019a). Therefore, in the present study, mostly debris flow and a very few numbers of slides are considered for the creating of the landslide masks (Figure 8). It is already mentioned earlier that no institutional mechanism is present to create the landslide masks for Bangladesh (Rabby & Li, 2019).

Therefore, for this study, landslide masks have been created manually. Fortunately, the latitude and longitude of several past landslide events have been collected from several field visits by a team of Conflict and Disaster Research Hub that is led by Dr Bayes Ahmed, Associate Professor at the Institute for Risk and Disaster Reduction, University College London (IRDR UCL), UK. These points are used as reference points of landslides and by leveraging Google Earth Pro, these points are cross-checked to know whether there were landslide events in those areas during the respective periods.

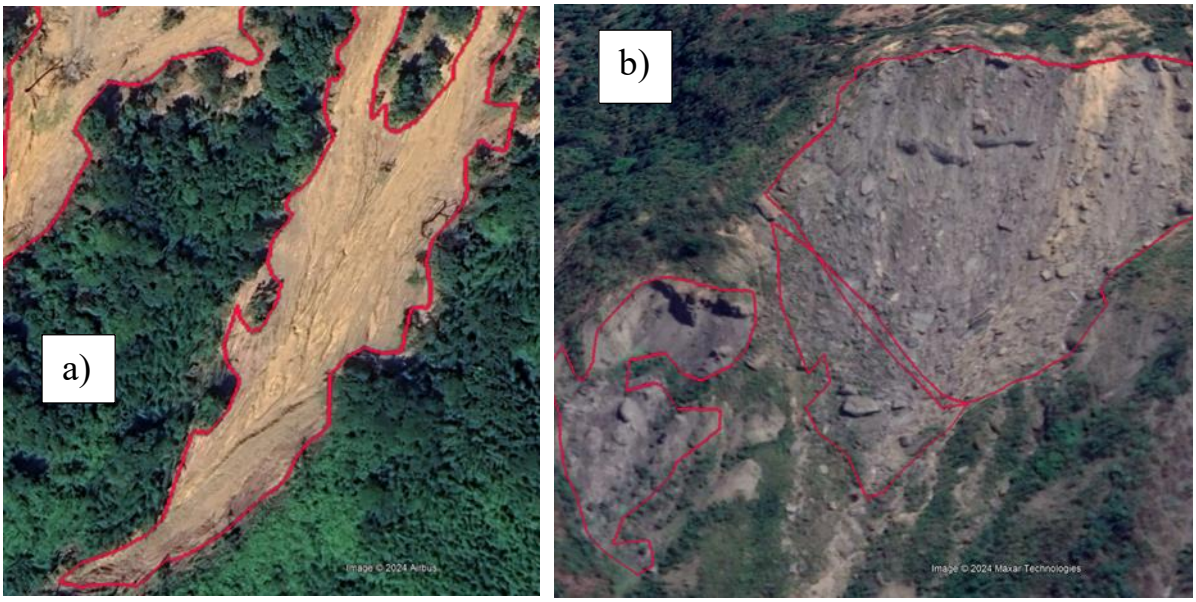

Figure 8: Creation of the landslide masks a) debris flow (2023 October 20 in Alikadam upazila), and b) rock slides (2019 January 11 in Rangamati Sadar upazila) from Google Earth Pro.

After carefully scrutinizing the points, the points where landslides are visible (examining in Google Earth Pro) by aligning with the same or closer date of the collection of the respective points, have been considered for the creation of landslides masks for this study. In each landslide point, the polygon has been created based on the area visible as landslides on the image in Google Earth Pro (Figure 8). With a view to the meticulous creation of polygons for each tile, the temporal context is systematically set up through the use of the Google Earth Pro timer, aligning with the date of the collection of the respective image tile.

Table 4 demonstrates the number of polygons created as landslide masks, their time, place, types, and total area (m^2). In total, 181 landslide polygons (total area 314,523 m^2) have been created. The majority of the landslide polygons (151 polygons; total area 225,882 m^2) created from Alikadam upazila and the landslides have been occurred in 2023. Except 4 polygons (rock

slides in Rangamati Sadar upazila; Figure 8b), all the polygons represent the debris flow (Figure 8a) landslides. After creating all the polygons, the polygons are assigned to the respective image tile aligning with their location, and then, the polygons of each image tile are rasterized employing QGIS 3.34.1, where the landslide areas are assigned a value of 1, while non-landslide areas are designated with a value of 0. The rasterization process is executed for each image tile at 3m and 10m spatial resolutions.

4.3 Repeated stratified hold-out validation

In the present study, the entire dataset has been divided by following the hold-out method. More specifically, the entire dataset has been divided into training (50% of the dataset), validation (25% of the dataset), and testing (25% of the dataset) parts. It is called as hold-out because the entire dataset splits into 3 parts and each part is used for specific purpose. The training set is used for training the U-net model to learn the objects from the images. Although it is known as validation set, is also a part of training phase. The validation set is used for tuning the hyperparameters of the model with different learning rate, batch size, and filters (discussed in section 3.3); in this stage, the U-net model does not learn the objects from the image from the validation set. Validation can be done from the training dataset or from a separate set of data. There is a chance to leak the information, if validation has been done from the same set of training data (Tennenholtz et al., 2018). Therefore, having a separate set of validation dataset is better to achieve the optimal performance of the model (Tennenholtz et al., 2018). For this reason, this study used a separate set of validation data. Finally, the test dataset is used for evaluating the final performance of the model.

Table 5: Formation of the folds used for the repeated stratified hold-out validation.

Fold number	Number of image tiles	Upazila (from where the image tile collected)	Year	Number of polygons (landslide masks)
Fold 1	5	Alikadam	2023	114
Fold 2	5	Alikadam	2023 and 2022	43
Fold 3	5	Alikadam, Kaptai, and Rangamati Sadar	2022 and 2018	12
Fold 4	5	Rangamati Sadar, and Thanchi	2018, 2019 and 2022	12

The dataset used in the present study is not a very large dataset. Therefore, it can be possible to get a biased results from a specific hold-out of data used, including training, validation and

test sets. To overcome this issue, this study has adopted the stratified technique, indicating the entire dataset is organized in a way so that all data is used as test set in such a way that the datasets are approximately balanced with respect to the number of landslide samples present among the training, validation and test sets (see Table 6). This comprehensive approach enhances the model evaluation process.

Figure 9 represents how the entire dataset is utilized to perform the experiments. It is already noted that this study incorporates a dataset comprising 20 distinct image tiles. To apply the stratified technique in this dataset, the dataset is partitioned into 4 folds, each constituted by five image tiles where each image tile is used in a specific fold (Table 5). Table 5 presents how many landslide polygons, from which year and upazila, are present in each fold. It is demonstrated that the fold 1 contains the larger number (114 out of 181) of landslide polygons (Table 5). However, both folds 3 and 4 contain only 12 landslide polygons indicating an imbalance set of data. After the creation of the 4 folds, 4 trials are designed where in each trial the fold used as test dataset has been changed, thus covering the whole dataset (Figure 9).

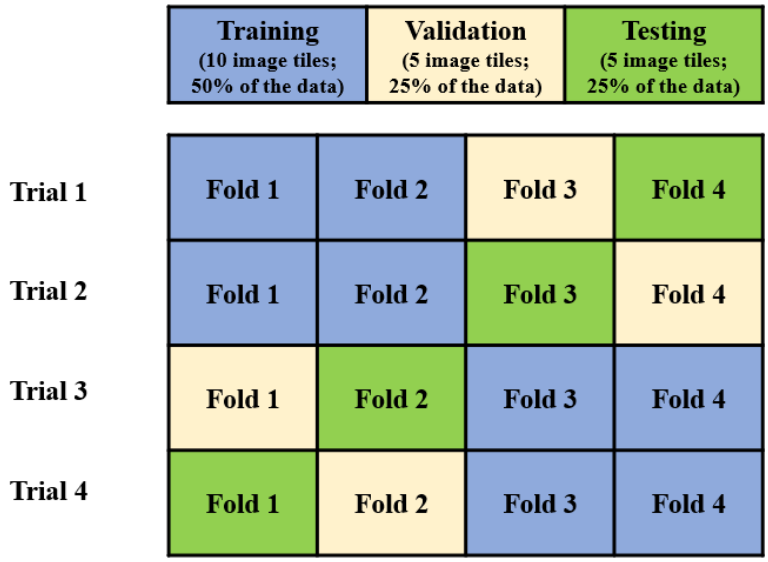

Figure 9: Repeated stratified hold-out validation method used in this study.

In Trial 1, folds 1 and 2 constitute the training set, fold 3 is allocated for validation, and fold 4 is designated for testing. Subsequently, in Trial 2, folds 1 and 2 serve as the training set, fold 4 is assigned to validation, and fold 3 is utilized for testing. Following this, Trial 3 employs folds 3 and 4 for training, fold 1 for validation, and fold 2 for testing. Finally, Trial 4 involves training with folds 3 and 4, using fold 2 for validation, and employing fold 1 for testing. The final

evaluation has been done by averaging the values of evaluation metrics produced from the all 4 trials. Thus, the hold-out method is applied in all the four trials. For this reason, the whole process is called as the repeated stratified hold-out validation.

4.4 Experimental Setup

This study incorporates four distinct datasets to determine their relative performance, as illustrated in Figure 10:

- Dataset A exclusively comprises the RGB and NIR bands derived from VHR PlanetScope Imagery.
- Dataset B extends its composition to include RGB and NIR bands derived from VHR PlanetScope Imagery, along with additional layers encompassing Elevation, Slope, and the NDVI (computed from the NIR and red bands).
- Dataset C encompasses the RGB and NIR bands derived from MR Sentinel-2 Imagery.
- Dataset D integrates RGB and NIR bands from MR Sentinel-2 Imagery, along with Elevation, Slope, and NDVI.

Figure 10 visually describes the comprehensive workflow of the proposed landslide detection approach employed in this study. Each dataset is experimented separately by maintaining the repeated stratified hold-out method described in the previous section. Therefore, in this study 16 (4 trials with each dataset) different experiments have been conducted. Every single experiment is fed into the U-net model for the detection of landslides. Performance of each experimental dataset is evaluated based on the precision, recall, F1 score, loss, and accuracy metrics values. Then, the results are visualized by graphs, maps, and table. Finally, the best performed experimental dataset is selected.

From each image tile, multiple patches are generated. For Datasets A and B, characterized by a spatial resolution of 3m (VHR Imagery), 100 patches are derived from each image tile, with a standardized patch size of 64 x 64. Consequently, during the training phase, a total of 1,000 patches are generated from 10 image tiles, 500 patches for validation from 5 image tiles, and 500 patches for testing from 5 image tiles (Tables 5 and 6). In contrast, for Datasets C and D, featuring a spatial resolution of 10m (MR Imagery), 36 patches are generated from each image tile, maintaining a patch size of 32 x 32 results in the creation of 360 training patches from 10 image tiles, 180 validation patches from 5 image tiles, and 180 testing patches from 5 image tiles (Tables 5 and 6).

Table 6: Total number of pixels used as landslide samples.

Trial Number	Datasets A and B			Datasets C and D		
	Training (1,000 patches)	Validation (500 patches)	Testing (500 patches)	Training (380 patches)	Validation (180 patches)	Testing (180 patches)
1	25,935	7,161	1,851	2,309	635	161
2	25,935	1,851	7,161	2,309	161	635
3	9,012	10,467	15,468	796	939	1,370
4	9,012	15,468	10,467	796	1,370	939

It is important to note that in creating the least number of patches, it was necessary to keep the patch size 32 x 32. Otherwise, if it would be 64 x 64, the number of total patches would be 18 from each image tile which is not a compatible number compared to the VHR imagery as well as the number of samples (patches) is unsuitable for the DL model. Given the application of data augmentation during the training phase in this study, the overall number of training patches in each trial amounts to 2,000 for datasets A and B, and 760 for datasets C and D.

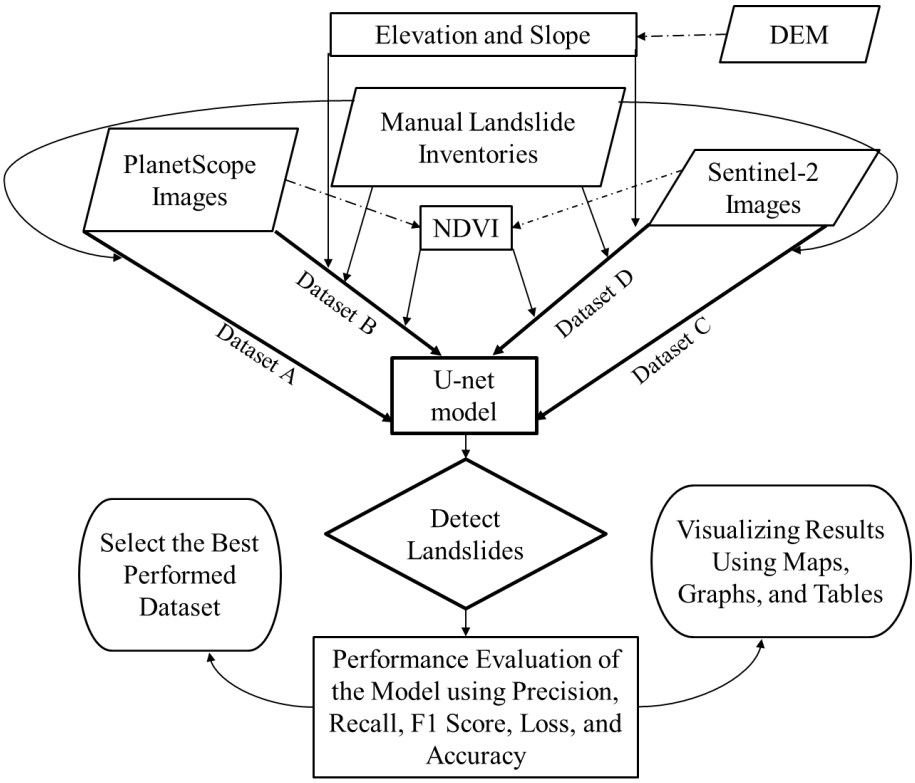

Figure 10: The flow chart shows the full process of the proposed landslide detection approach.

It is important to highlight that the aggregate number of pixels employed to indicate landslides varies across the trials designated for the training, validation, and testing sets. Moreover, the total number of pixels extracted from landslide masks utilized in conjunction with the Datasets A and B differs from that derived from landslide masks utilized in conjunction with the Datasets C and D. Table 6 shows the cumulative number of landslide pixels are same during the training phase in Trials 1 and 2 amounts to 25,935 for Datasets A and B; and 2,309 for Datasets C and D. Conversely, the total number of landslide pixels during the training phase in Trials 3 and 4 is only 9,012 for Datasets A and B; and 796 for Datasets C and D.

It is evident from a number of works conducted with DL models that a higher number of target pixels (in this case, landslides pixels) during the training phase tends to enhance the model's performance (Meena et al., 2022). The current investigation systematically explores both the use of the highest (Trials 1 and 2) and lowest (Trials 3 and 4) landslide samples during the training phase, in contrast to the remaining phases, to assess the model's performance.

5 Results and Discussion

A total of 16 different experiments have been conducted by utilizing the U-net model, and each experiment involves different sets of hyperparameter configurations, to detect landslides more precisely. The best weights model was determined for each experiment based on the outcomes derived from the testing phase. The hyperparameter configuration characterized by a filter size of 8, a batch size of 4, a learning rate of 0.01, and the utilization of the Adam optimizer emerged as, comparatively, the best model weight across all experiments.

The F1 scores across successive epochs (5 epochs), illustrating the performance of the U-net model with the aforementioned hyperparameter set during both training and validation phases for all 16 experiments, are presented in Figures A1 and A2. It is observed in Figure A1 that the F1 score shows an upward curve in Trials 1 and 2 for both Datasets A and B. However, in trials 3 and 4, the F1 score curves are not good enough which indicates that the detection of landslides can be misleading in these Trials (3 and 4). Nevertheless, in Figure A2, the F1 score shows an upward curve in all trials in both Datasets C and D. Figure A3 displays the loss curve throughout epochs, outlining the optimization evolution of the model under the same set of hyperparameters across all experimental trials for the Datasets A and B. The loss curves are going downwards in Trials 1 and 2 indicating a good delineation of landslides detection using Datasets A and B. Furthermore, in Trials 1 and 2 using Dataset C, and trials 1, 2, and 3 using Dataset D, the model has produced proper downward loss curves (Figure A4).

The evaluation metrics of the experiments conducted for landslide detection utilizing the 16 experimental datasets are shown in Table 7. Table 7 demonstrates the precision, recall, F1 score, loss, and accuracy metrics for the 16 experimental datasets during their testing phase. Figures 11 and 12 visually represent the area identified as landslides based on the 16 experimental datasets, utilizing the U-net model. A noteworthy observation from the table is the consistent findings in Datasets A and B, comprising VHR imagery from PlanetScope, across all 4 trials.

For both Datasets A and B, the mean precision, recall, F1 score, loss, and accuracy metrics found from Trials 1, 2, 3, and 4 are 1, 0.422, 0.422, 0.589, and 0.996, respectively (Table 7). The standard deviation (std) calculated from Trials 1, 2, 3, and 4 for the evaluation metrics are same in Datasets A and B which are 0, 0.231, 0.231, 0.237, and 0.002, for precision, recall, F1 score, loss, and accuracy, respectively. Remarkably, Trials 1 and 2 show the best performance in landslides detection compared to Trials 3 and 4.

Table 7: Precision, Recall, F1 score, Loss, and Accuracy of 16 experimental test datasets.

Trial number	Precision	Recall	F1 score	Loss	Accuracy
Dataset A					
1	***1***	***0.625***	***0.625***	***0.381***	***0.999***
2	***1***	***0.625***	***0.625***	***0.383***	***0.997***
3	1	0.062	0.062	0.958	0.993
4	1	0.375	0.375	0.634	0.996
Mean	1	0.422	0.422	0.589	0.996
Std	0	0.231	0.231	0.237	0.002
Dataset B					
1	***1***	***0.625***	***0.625***	***0.380***	***0.999***
2	***1***	***0.625***	***0.625***	***0.383***	***0.997***
3	1	0.062	0.062	0.958	0.993
4	1	0.375	0.375	0.634	0.996
Mean	1	0.422	0.422	0.589	0.996
Std	0	0.231	0.231	0.237	0.002
Dataset C					
1	1	0.167	0.167	0.998	0.999
2	***1***	***0.333***	***0.333***	***0.700***	***0.997***
3	0.004	0.955	0.009	0.996	0.055
4	0.056	0.111	0.059	0.992	0.976
Mean	0.515	0.392	0.142	0.922	0.757
Std	0.485	0.335	0.124	0.128	0.405
Dataset D					
1	1	0.167	0.167	0.785	0.999
2	***1***	***0.333***	***0.333***	***0.717***	***0.997***
3	0.004	0.955	0.009	0.996	0.055
4	0.711	0.004	0.008	0.991	0.995
Mean	0.679	0.365	0.129	0.872	0.762
Std	0.407	0.360	0.134	0.124	0.408

Note: The best performed trial/trials in 4 datasets is/are highlighted in bold and italic format.

It should be noted here that the samples (landslides pixels) used in Trials 3 and 4 during the training phase to train the U-net model are significantly small compared to the training samples used in Trials 1 and 2- (Table 6). This is evident that the DL models perform better with more training samples (Ahmed et al., 2023). Trials 1 and 2, with a higher number of training samples than testing samples, demonstrate better performance; while Trials 3 and 4, with the opposite configuration, exhibit poorer performance in detecting landslides on test Datasets A and B (Figure 11). Meena et al. (2022) also revealed similar findings that the usage of large and small training samples influences the performance of the DL model in detecting landslides, and the DL model always outperforms with a larger set of training samples. Hence, it is concluded that DL models like U-net should be trained with a higher number of training samples to get reliable accuracy.

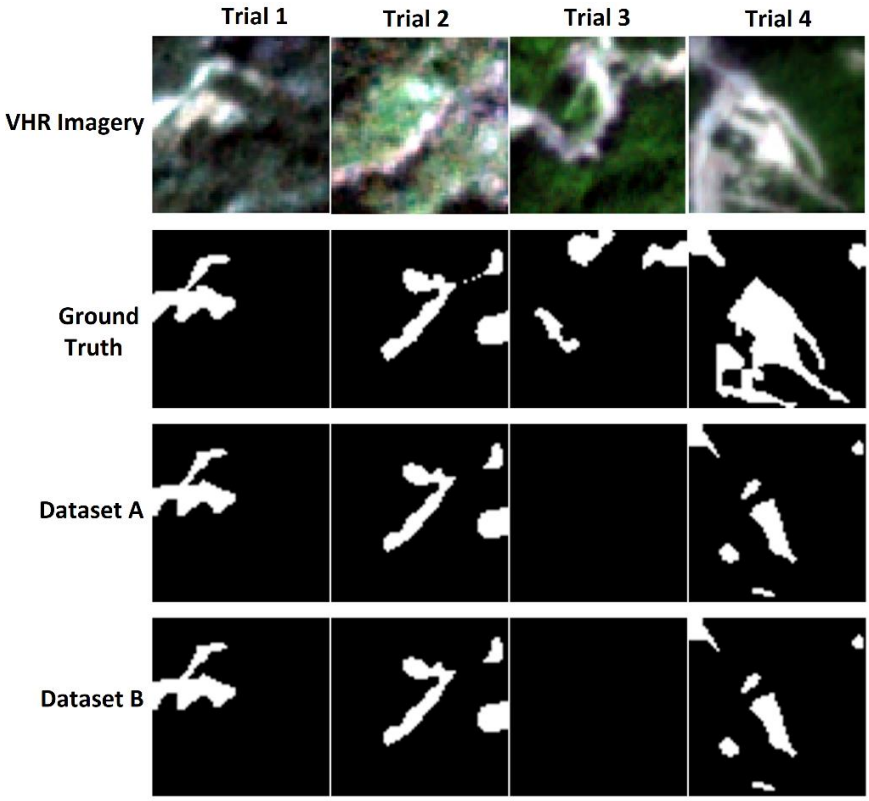

Figure 11: Landslide detection results getting from using Datasets A, and B in 4 trials.

Considering the above discussion on the number of training samples on U-net model's performance, Trials 1 and 2 from Datasets A and B are considered as acceptable for landslides detection in the current study area. Consequently, the mean precision, recall, F1 score, loss, and accuracy metrics for both Datasets A and B, based on Trials 1 and 2, are 1, 0.625, 0.625, 0.380, and 0.999, respectively. The std values for precision, recall, F1 score, loss, and accuracy in Datasets A and B based on Trials 1 and 2 are 0, 0, 0, 0.001, and 0.001, respectively, indicating the preciseness of these results. Figure 11 visually supports the close alignment between the model's landslides detection and the ground truth. Figure A5 shows the comparison between landslides detection by the U-net model, experimented with Dataset A in Trials 1 and 2, in a whole image tile and the ground truth.

Meena et al. (2023) utilized 5 U-net based models for the detection of both rainfall and earthquake-induced landslides in different parts of the world, utilizing the RGB and NIR bands of PlanetScope imagery and a larger set of landslides samples. The used 5 models U-net, Attention U-Net, Attention Res U-Net, and ADSMS models produced F1 score 0.7904, 0.6825, 0.7446, 0.6477, and 0.6576, respectively. The values of the F1 scores found by them are closer to the F1 score (0.625) produced in the present study by the utilization of the Dataset A and B.

Another study of Bhuyan et al. (2023) by utilizing only the optical (RGB) bands of PlanetScope imagery in Japan, Nepal, New Zealand, and Papua New Guinea showed that U-net based ADSMS model is a good model (F1 score 0.8) for detecting earthquake-triggered landslides. Furthermore, Meena et al. (2022) conducted a study in Rasuwa in Nepal by utilizing the VHR RapidEye imagery (5m of spatial resolution) along with the ALOS PALSAR DEM in detecting landslides. Two sets of data like the present study have been created that are:

i) only 5 spectral bands RGB, NIR, and red-edge of RapidEye imagery (dataset 1), and
ii) RGB, NIR, and red-edge bands of RapidEye imagery with elevation and slope derived from ALOS PALSAR DEM (dataset 2).

They used both DL (U-net) and ML (SVM, RF, and KNN), and found that U-net outperformed ML models in both sets of the data. Interestingly, in this work, all models produced better results with dataset 1 over dataset 2. The values, produced by the U-net for the dataset 1, of precision, recall, F1 score, and MCC are 0.765, 0.663, 0.711, and 0.710, respectively. Whereas, the values, produced by the U-net for the dataset 2, of precision, recall, F1 score, and MCC are 0.614, 0.797, 0.694, and 0.697, respectively.

The F1 scores (0.711 without elevation and slope data; and 0.694 with elevation and slope data) produced in the work of Meena et al. (2022) is compatible with the F1 score (0.625) produced in the present study. It also indicates that the inclusion of the additional topographic data elevation, and slope with the RGB and NIR bands of the satellite imagery did not improve the results. Meena et al. (2023) also revealed the similar outcomes that the spectral bands RGB, and NIR of PlanetScope imagery is good enough to detect landslides, and there is no need to integrate the additional topographic data to enhance the accuracy.

Datasets C and D yielded comparatively poor results in contrast to Datasets A and B, as illustrated in Figure 12 and Table 7. It is crucial to highlight here that Datasets C and D made of MR Imagery obtained from Sentinel-2. Like Datasets A and B, the model exhibited largely analogous outcomes for Datasets C and D, except Trial number 4 (Table 7). The mean precision, recall, F1 score, loss, and accuracy metrics, based on the Trials 1, 2, 3, and 4 for Dataset C, are 0.515, 0.392, 0.142, 0.922, and 0.757, respectively (Table 7). The std values for precision, recall, F1 score, loss, and accuracy metrics for Dataset C across the same trials are 0.485, 0.335, 0.124, 0.128, and 0.405, respectively (Table 7). In Trials 1 and 3, the model did not identify any landslide area utilizing both Datasets C and D (Figure 12). Conversely, in Trial

2, the U-net model identified a few landslide pixels in both datasets. The model consistently identified the same areas as landslides on Dataset C and D in Trials 1, 2, and 3 (Figure 12). However, in Trial 4, the model yielded conflicting results by detecting all landslide areas but erroneously classifying some non-landslide areas as landslides, an outcome considered unacceptable (Figure 12).

Lu et al. (2023) conducted study in Wu den, China and Iburi in Japan to detect landslides automatically by leveraging the MR Sentinel-2 imagery and the National Aeronautics and Space Administration (NASA) provided DEM. They have modified the original U-net model to dual encoder U-net model and revealed that their model is competitive enough (precision, recall, and F1 score are 0.783, 0.801, and 0.792, respectively) for detecting earthquake-induced landslides. Xia et al. (2022) also utilized Sentinel-2 imagery for detecting earthquake-induced landslides with CNN, AlexNet, ResNet152V2, DenseNet201, InceptionV3, Xception and InceptionResNetV2. In that work DenseNet201 was found as the best performed model by producing 0.934, 0.8448, 0.8872, and 0.9172 values, respectively, for precision, recall, F1 score, and accuracy. Most of the studies conducted previously based on Sentinel-2 detected earthquake-triggered landslides (Lu et al., 2023; Xia et al., 2022; Jelenek & Kopačková-Strnadová, 2021).

It is important to note that the landslides in the present study area are characterized as rainfall-induced shallow landslides, and the affected areas are not extensive (Kamal et al., 2022; Abedin et al., 2020). Consequently, detecting small-scale shallow landslides using Sentinel-2 imagery is challenging due to its limited spatial resolution. Satellite imagery, like Sentinel-2, which is accessible free of charge, often possesses a relatively coarse resolution or MR (10m spatial resolution). Because of this, MR imagery is composed of bigger size pixels which contain less information, and the information is not detailed (Lanaras et al., 2018). This may make this MR imagery less suitable for effectively detecting rainfall-induced shallow landslides (Das & Wegmann, 2022).

Furthermore, the temporal resolution, indicating the frequency of satellite revisits, significantly impacts the quality of landslide mapping, especially rainfall-induced shallow landslides (Novellino et al., 2024). Moreover, most of the time, it is not possible to extract the data from the closer data after the landslide events because of the presence of a high percentage of cloud cover (Zhang et al., 2020). Therefore, a prolonged revisiting time may lead to the disappearance of landslide indicators on the terrain due to factors including re-vegetation, mark degradation,

or human-induced activities, potentially diminishing the accuracy of landslide detection; which do not appear as challenges for detecting earthquake-triggered landslides.

In contrast, the superior spatial (VHR) and temporal (daily) resolution of PlanetScope imagery facilitate the detection of shallow landslides; as it has smaller size pixels with detailed information and more available data which facilitates the collection of appropriate data (Muetingn & Bookhagen, 2023). Additionally, despite Sentinel-2's inclusion of additional spectral bands, its inferior spatial resolution hampers precise boundary delineation and accurate localization of landslides (Meena et al., 2023). The noteworthy attributes of PlanetScope imagery, including its competitive spatial resolution, and daily temporal resolution, contribute to its effectiveness in landslide detection by delineating the precise boundary of landslides. Furthermore, the standardized sensors across PlanetScope satellites simplify and enhance the accuracy of image pre-processing and orthorectification which has eased the overall process of landslide detection (Meena et al., 2023).

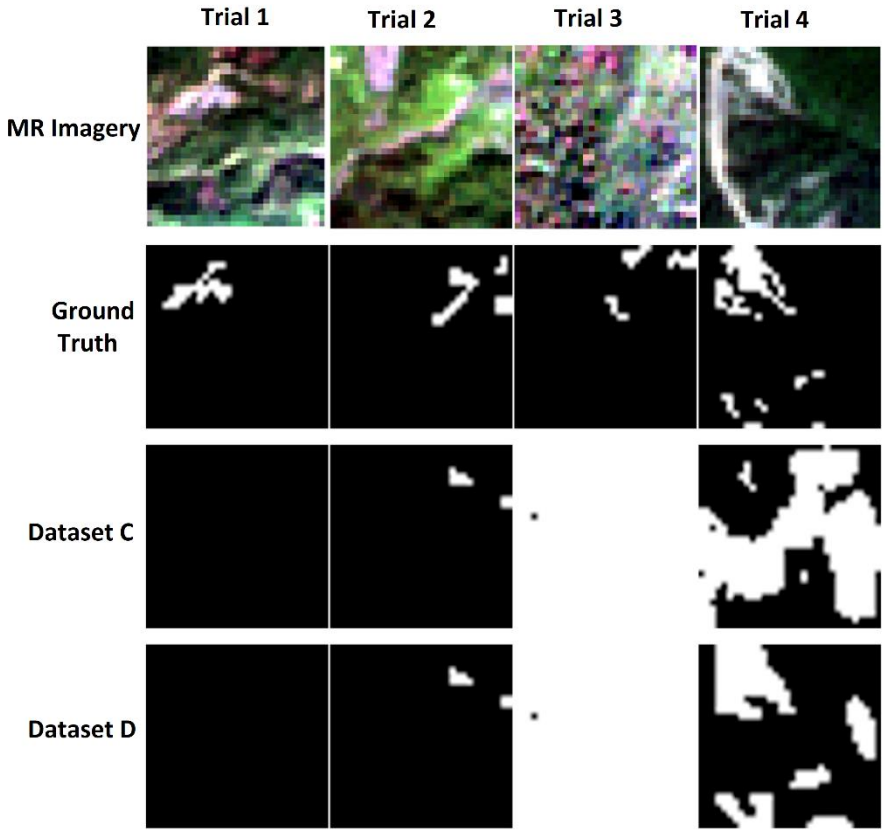

Figure 12: Landslide detection results getting from using Datasets C, and D in 4 trials.

Therefore, it can be concluded from the above results and discussion that PlanetScope images are optimal for detecting rainfall-induced shallow landslides. Moreover, for the effective implementation of a DL model like U-net, a higher number of training samples representing landslides is considered crucial to achieving acceptable and reliable results. As well as, it is not essential to the inclusion of elevation, slope, and NDVI to enhance the accuracy, rather the spectral bands RGB and NIR of VHR Imagery are enough to detect landslides. Thus, Datasets A and B are found as the best-performed datasets in Trials 1 and 2 for the detection of rainfall-induced shallow landslides in the study area.

5.1 Limitations of this Study

Several limitations in this study impact the accuracy of the experimental findings, which could be summarized as follows:

- The main challenge was manually delineating landslide polygons used as landslide masks (ground truth data), and uncertainties introduced by satellite imagery because of the presence of clouds which leads to choose a date far from the landslides occurrence when the image has been found as cloud free.
- The inability to conduct on-site field observations hindered the verification of precise landslides boundaries when creating the landslide masks by leveraging the Google Earth Pro, despite having latitude and longitude points of landslides.
- Various factors, for example, mapping scale, date, and satellite image quality, contribute to the precision of landslide detection. The radiometric resolution is an additional element affecting the creation of manual inventories. Furthermore, the model's performance is hampered by the haze effect caused by instrument errors.
- Additionally, DL requires a large dataset which has not been possible to use in this study, because there were a small number of available landslides masks that could use here.
- Moreover, even with the VHR PlanetScope imagery, sometimes shallow landslides cannot be detected. It requires more finer resolution imagery like Maxar or Pleiades.
- Besides, this study only used ALOS PALSAR DEM which has some limitations according to the scholars.

All these aforementioned obstacles have made it difficult to execute the current study in detecting rainfall-induced shallow landslides.

6 Conclusion

This study employed a systematic approach for training the U-net model in 16 different experiments. Aimed at identifying the optimal model weights, a diverse set of hyperparameters have been used. The hyperparameter set featuring a filter size of 8, a batch size of 4, a learning rate of 0.01, and the Adam optimizer is found as the most effective across all experiments.

The model is trained with the Dice loss function to learn the landslide features more precisely. The overall performance of the model in detecting landslides is evaluated by precision, recall, F1 score, and accuracy metrics. Consistent results are found in both Datasets A and B in all 4 Trials. However, Trials 1 and 2 show a superior performance in detecting landslides by producing less error over the rest of two Trials, indicating the better performance of DL model with larger training samples. As Trials 3 and 4, produce a very low accuracy which is not acceptable, Trials 1 and 2 from both Datasets A and B are considered as optimal for landslide detection in the study area. Therefore, the mean precision, recall, F1 score, loss, and accuracy based on Trials 1 and 2 are 1, 0.625, 0.625, 0.380, and 0.999, respectively (same results found in both Datasets A and B). On the contrary, unlike Datasets A and B, Datasets C and D produced poor results, indicating the importance of high spatial and temporal resolutions of satellite imagery in detecting rainfall-induced shallow landslides.

It can be summarized that the unique attributes of VHR PlanetScope imagery with its daily temporal and high spatial resolution make the process of automatic detection of rainfall-induced shallow landslides, comparatively, reliable, robust, and accurate. Additionally, it is also explored that the integration of the topographic data with PlanetScope imagery does not improve the accuracy; only PlanetScope imagery is good enough for landslide detection. Furthermore, it is found that for an effective and precise result, a higher number of training samples is needed to train the DL model.

To overcome the limitations of this study, future study should apply more data augmentation methods to increase the training samples synthetically. Moreover, future study should integrate both channel and spatial attention layers to the U-net model to focus on the important features and discard the unnecessary feature for enhancing the performance and accuracy of the model. Experiment can be done with Shuttle Radar Topography Mission (SRTM) provided DEM data as well. Lastly, other DL models can be utilized to reveal their performance in detecting landslides in the study area.

References

Abdollahi, A. (2022). Automatic updating and verification of road maps using high-resolution remote sensing images based on advanced machine learning methods (Doctoral dissertation).

Abedin, J., Rabby, Y. W., Hasan, I., & Akter, H. (2020). An investigation of the characteristics, causes, and consequences of June 13, 2017, landslides in Rangamati District Bangladesh. Geoenvironmental Disasters, 7(1), 1-19.

Ahmed, B. (2021). The root causes of landslide vulnerability in Bangladesh. Landslides, 18(5), 1707-1720.

Ahmed, B., Rahman, M. S., Islam, R., Sammonds, P., Zhou, C., Uddin, K., & Al-Hussaini, T. M. (2018). Developing a dynamic Web-GIS based landslide early warning system for the Chittagong Metropolitan Area, Bangladesh. ISPRS International Journal of Geo-Information, 7(12), 485.

Ahmed, B., Rahman, M. S., Rahman, S., Huq, F. F., & Ara, S. (2014). Landslide Inventory Report of Chittagong Metropolitan Area, Bangladesh.

Ahmed, S. F., Alam, M. S. B., Hassan, M., Rozbu, M. R., Ishtiak, T., Rafa, N., ... & Gandomi, A. H. (2023). Deep learning modelling techniques: current progress, applications, advantages, and challenges. Artificial Intelligence Review, 1-97.

Alam, M., Alam, M. M., Curray, J. R., Chowdhury, M. L. R., & Gani, M. R. (2003). An overview of the sedimentary geology of the Bengal Basin in relation to the regional tectonic framework and basin-fill history. Sedimentary geology, 155(3-4), 179-208.

Amatya, P., Kirschbaum, D., & Stanley, T. (2022). Rainfall-induced landslide inventories for Lower Mekong based on Planet imagery and a semi-automatic mapping method. Geoscience Data Journal, 9(2), 315-327.

Banglapedia (2021). Geological Structure. Encyclopedia of Bangladesh.

BBS (2017). Small Area Atlas Bangladesh. Bangladesh Bureau of Statistics (BBS), Statistics and Informatics Division (SID), Ministry of Planning, Dhaka, Bangladesh.

BBS. (2023). Population and Housing Census 2022, National Report (Volume I). Bangladesh Bureau of Statistics (BBS), Statistics and Informatics Division (SID), Ministry of Planning, Dhaka, Bangladesh.

Bdnews24. (2023, August 8). Government opens 200 shelters amid floods, power outage in Bandarban. https://bdnews24.com/bangladesh/5fkxz9sblm

Bengio, Y., Goodfellow, I., & Courville, A. (2017). Deep learning (Vol. 1). Cambridge, MA, USA: MIT press.

Bhuyan, K., Tanyaş, H., Nava, L., Puliero, S., Meena, S. R., Floris, M., ... & Catani, F. (2023). Generating multi-temporal landslide inventories through a general deep transfer learning strategy using HR EO data. Scientific reports, 13(1), 162.

Bragagnolo, L., Rezende, L. R., da Silva, R. V., & Grzybowski, J. M. V. (2021). Convolutional neural networks applied to semantic segmentation of landslide scars. Catena, 201, 105189.

Brammer, H. (1986). Reconnaissance soil and land use survey: Chittagong Hill Tracts (1964–1965). Soil Resources Development Institute, Bangladesh.

Brammer, H. (2012). Physical geography of Bangladesh. The University Press Ltd.

Casagli, N., Intrieri, E., Tofani, V., Gigli, G., & Raspini, F. (2023). Landslide detection, monitoring and prediction with remote-sensing techniques. Nature Reviews Earth & Environment, 4(1), 51-64.

CDMP, I. (2012). Landslide Inventory and Landuse Mapping, DEM Preparation, Precipitation Threshold Value and Establishment of Early Warning Device. Ministry of Food and Disaster Management, Disaster Management and Relief Division, Government of the People's Republic of Bangladesh, 1, 310-324.

Chandra, N., Sawant, S., & Vaidya, H. (2023). An Efficient U-Net Model for Improved Landslide Detection from Satellite Images. PFG–Journal of Photogrammetry, Remote Sensing and Geoinformation Science, 91(1), 13-28.

Chen, G., Weng, Q., Hay, G. J., & He, Y. (2018). Geographic object-based image analysis (GEOBIA): Emerging trends and future opportunities. GIScience & Remote Sensing, 55(2), 159-182.

Chen, Z., Zhang, Y., Ouyang, C., Zhang, F., & Ma, J. (2018a). Automated landslides detection for mountain cities using multi-temporal remote sensing imagery. Sensors, 18(3), 821.

Crosta, G. B., & Frattini, P. (2008). Rainfall-induced landslides and debris flows. Hydrological Processes: An International Journal, 22(4), 473-477.

Das, R., & Wegmann, K. W. (2022). Evaluation of machine learning-based algorithms for landslide detection across satellite sensors for the 2019 Cyclone Idai event, Chimanimani District, Zimbabwe. Landslides, 19(12), 2965-2981.

De, S., Bruzzone, L., Bhattacharya, A., Bovolo, F., & Chaudhuri, S. (2017). A novel technique based on deep learning and a synthetic target database for classification of urban areas in PolSAR data. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 11(1), 154-170.

Deng, J., Dai, K., Liang, R., Chen, L., Wen, N., Zheng, G., & Xu, H. (2023). Interferometric Synthetic Aperture Radar Applicability Analysis for Potential Landslide Identification in Steep Mountainous Areas with C/L Band Data. Remote Sensing, 15(18), 4538.

Fukushima, K., Miyake, S., & Ito, T. (1983). Neocognitron: A neural network model for a mechanism of visual pattern recognition. IEEE transactions on systems, man, and cybernetics, (5), 826-834.

Ghorbanzadeh, O., Shahabi, H., Crivellari, A., Homayouni, S., Blaschke, T., & Ghamisi, P. (2022). Landslide detection using deep learning and object-based image analysis. Landslides, 19(4), 929-939.

Ghorbanzadeh, O., Tiede, D., Wendt, L., Sudmanns, M., & Lang, S. (2021). Transferable instance segmentation of dwellings in a refugee camp-integrating CNN and OBIA. European Journal of Remote Sensing, 54(sup1), 127-140.

Hao, X., Liu, L., Yang, R., Yin, L., Zhang, L., & Li, X. (2023). A Review of Data Augmentation Methods of Remote Sensing Image Target Recognition. Remote Sensing, 15(3), 827.

Hawryło, P., & Wężyk, P. (2018). Predicting growing stock volume of scots pine stands using Sentinel-2 satellite imagery and airborne image-derived point clouds. Forests, 9(5), 274.

Herrera, G., Mateos, R. M., García-Davalillo, J. C., Grandjean, G., Poyiadji, E., Maftei, R., ... & Jensen, O. A. (2018). Landslide databases in the Geological Surveys of Europe. Landslides, 15, 359-379.

Jadon, S. (2020, October). A survey of loss functions for semantic segmentation. In 2020 IEEE conference on computational intelligence in bioinformatics and computational biology (CIBCB) (pp. 1-7). IEEE.

Jelenek, J., & Kopačková-Strnadová, V. (2021). Synergic use of Sentinel-1 and Sentinel-2 data for automatic detection of earthquake-triggered landscape changes: A case study of the 2016 Kaikoura earthquake (Mw 7.8), New Zealand. Remote Sensing of Environment, 265, 112634.

Kamal, A. M., Hossain, F., Rahman, M. Z., Ahmed, B., & Sammonds, P. (2022). Geological and soil engineering properties of shallow landslides occurring in the Kutupalong Rohingya Camp in Cox's Bazar, Bangladesh. Landslides, 19(2), 465-478.

Karimi, B., Yanchuck, M., & Foust, J. (2019). A new landslide inventory and improved susceptibility model for northeastern Pennsylvania. Environmental Geosciences, 26(4), 125-145.

Keyport, R. N., Oommen, T., Martha, T. R., Sajinkumar, K. S., & Gierke, J. S. (2018). A comparative analysis of pixel-and object-based detection of landslides from very high-resolution images. International journal of applied earth observation and geoinformation, 64, 1-11.

Khal, M., Algouti, A., Algouti, A., Akdim, N., Stankevich, S. A., & Menenti, M. (2020). Evaluation of open Digital Elevation Models: estimation of topographic indices relevant to erosion risk in the Wadi M'Goun watershed, Morocco. AIMS Geosci, 6, 231-257.

Kumar, A., Asthana, A. K. L., Priyanka, R. S., Jayangondaperumal, R., Gupta, A. K., & Bhakuni, S. S. (2017). Assessment of landslide hazards induced by extreme rainfall event in Jammu and Kashmir Himalaya, northwest India. Geomorphology, 284, 72-87.

Lanaras, C., Bioucas-Dias, J., Galliani, S., Baltsavias, E., & Schindler, K. (2018). Super-resolution of Sentinel-2 images: Learning a globally applicable deep neural network. ISPRS Journal of Photogrammetry and Remote Sensing, 146, 305-319.

Li, H., He, Y., Xu, Q., Deng, J., Li, W., & Wei, Y. (2022). Detection and segmentation of loess landslides via satellite images: A two-phase framework. Landslides, 19(3), 673-686.

Li, X., Sun, X., Meng, Y., Liang, J., Wu, F., & Li, J. (2019). Dice loss for data-imbalanced NLP tasks. arXiv preprint arXiv:1911.02855.

Lin, Y. N., Chen, Y. C., Kuo, Y. T., & Chao, W. A. (2022). Performance study of landslide detection using multi-temporal SAR images. Remote Sensing, 14(10), 2444.

Lu, H., Ma, L., Fu, X., Liu, C., Wang, Z., Tang, M., & Li, N. (2020). Landslides information extraction using object-oriented image analysis paradigm based on deep learning and transfer learning. Remote Sensing, 12(5), 752.

Lu, W., Hu, Y., Zhang, Z., & Cao, W. (2023). A dual-encoder U-Net for landslide detection using Sentinel-2 and DEM data. Landslides, 1-13.

Ma, J., Chen, J., Ng, M., Huang, R., Li, Y., Li, C., ... & Martel, A. L. (2021a). Loss odyssey in medical image segmentation. Medical Image Analysis, 71, 102035.

Ma, Z., Mei, G., & Piccialli, F. (2021). Machine learning for landslides prevention: a survey. Neural Computing and Applications, 33, 10881-10907.

Maxwell, A. E., Warner, T. A., & Guillén, L. A. (2021). Accuracy assessment in convolutional neural network-based deep learning remote sensing studies—Part 1: Literature review. Remote Sensing, 13(13), 2450.

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, 1937-1950.

Meena, S. R., Nava, L., Bhuyan, K., Puliero, S., Soares, L. P., Dias, H. C., ... & Catani, F. (2023). HR-GLDD: a globally distributed dataset using generalized deep learning (DL) for rapid landslide mapping on high-resolution (HR) satellite imagery. Earth System Science Data, 15(7), 3283-3298.

Meena, S. R., Soares, L. P., Grohmann, C. H., Van Westen, C., Bhuyan, K., Singh, R. P., ... & Catani, F. (2022). Landslide detection in the Himalayas using machine learning algorithms and U-Net. Landslides, 19(5), 1209-1229.

Minaee, S., Boykov, Y., Porikli, F., Plaza, A., Kehtarnavaz, N., & Terzopoulos, D. (2021). Image segmentation using deep learning: A survey. IEEE transactions on pattern analysis and machine intelligence, 44(7), 3523-3542.

Mondini, A. C., Guzzetti, F., & Melillo, M. (2023). Deep learning forecast of rainfall-induced shallow landslides. Nature communications, 14(1), 2466.

Monsieurs, E., Jacobs, L., Michellier, C., Basimike Tchangaboba, J., Ganza, G. B., Kervyn, F., ... & Dewitte, O. (2018). Landslide inventory for hazard assessment in a data-poor context: a regional-scale approach in a tropical African environment. Landslides, 15, 2195-2209.

Mueting, A., & Bookhagen, B. (2023). Tracking slow-moving landslides with PlanetScope data: new perspectives on the satellite's perspective. EGUsphere, 2023, 1-36.

Nafiu, R.F. (2021, July 27). Heavy rains trigger fear of landslides in Bandarban. DhakaTribune. https://www.dhakatribune.com/bangladesh/nation/253479/heavy-rains-trigger-fear-of-landslides-in

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, 14(6), 1449.

Novellino, A., Pennington, C., Leeming, K., Taylor, S., Alvarez, I. G., McAllister, E., Arnhardt, C., & Winson, A. (2024). Mapping landslides from space: A review. Landslides.

Paul, V. K. (2020). Understanding Vulnerability and Resilience of Critical Infrastructure in Extreme Weather Events.

Prakash, N., Manconi, A., & Loew, S. (2020). Mapping landslides on EO data: Performance of deep learning models vs. traditional machine learning models. Remote Sensing, 12(3), 346.

Qi, W., Wei, M., Yang, W., Xu, C., & Ma, C. (2020). Automatic mapping of landslides by the ResU-Net. Remote Sensing, 12(15), 2487.

Qin, S., Guo, X., Sun, J., Qiao, S., Zhang, L., Yao, J., ... & Zhang, Y. (2021). Landslide detection from open satellite imagery using distant domain transfer learning. Remote sensing, 13(17), 3383.

Rabby, Y. W., & Li, Y. (2019). An integrated approach to map landslides in Chittagong Hilly Areas, Bangladesh, using Google Earth and field mapping. Landslides, 16(3), 633-645.

Rabby, Y. W., & Li, Y. (2019a). Landslide inventory (2001–2017) of Chittagong hilly areas, Bangladesh. Data, 5(1), 4.

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18 (pp. 234-241). Springer International Publishing

Sifa, S. F., Mahmud, T., Tarin, M. A., & Haque, D. M. E. (2020). Event-based landslide susceptibility mapping using weights of evidence (WoE) and modified frequency ratio (MFR) model: A case study of Rangamati district in Bangladesh. Geology, Ecology, and Landscapes, 4(3), 222-235.

Su, Z. (2021). Advanced 2D and 3D Computer Vision for a Smarter City-From Image to Point Cloud. Hong Kong University of Science and Technology (Hong Kong).

Tabari, H. (2020). Climate change impact on flood and extreme precipitation increases with water availability. Scientific reports, 10(1), 13768.

Taylor, L., & Nitschke, G. (2018). Improving deep learning with generic data augmentation. In 2018 IEEE symposium series on computational intelligence (SSCI) (pp. 1542-1547). IEEE.

Tehrani, F. S., Calvello, M., Liu, Z., Zhang, L., & Lacasse, S. (2022). Machine learning and landslide studies: recent advances and applications. Natural Hazards, 114(2), 1197-1245.

Tennenholtz, G., Zahavy, T., & Mannor, S. (2018). Train on validation: squeezing the data lemon. arXiv preprint arXiv:1802.05846.

The Hindu (2023). 55 killed in Bangladesh floods in August: officials, website: www.thehindu.com.

Thirugnanam, H., Ramesh, M. V., & Rangan, V. P. (2020). Enhancing the reliability of landslide early warning systems by machine learning. Landslides, 17, 2231-2246.

Varnes, D. J. (1958). Landslide types and processes. Landslides and engineering practice, 24, 20-47.

Wang, H., Zhang, L., Yin, K., Luo, H., & Li, J. (2021). Landslide identification using machine learning. Geoscience Frontiers, 12(1), 351-364.

Wen, D., Huang, X., Bovolo, F., Li, J., Ke, X., Zhang, A., & Benediktsson, J. A. (2021). Change detection from very-high-spatial-resolution optical remote sensing images: Methods, applications, and future directions. IEEE Geoscience and Remote Sensing Magazine, 9(4), 68-101.

Xia, J., Liu, H., & Zhu, L. (2022, April). Landslide hazard identification based on deep learning and sentinel-2 remote sensing imagery. In Journal of Physics: Conference Series (Vol. 2258, No. 1, p. 012031). IOP Publishing.

Xu, H., He, H., Zhang, Y., Ma, L., & Li, J. (2023). A comparative study of loss functions for road segmentation in remotely sensed road datasets. International Journal of Applied Earth Observation and Geoinformation, 116, 103159.

Xu, Y., Ouyang, C., Xu, Q., Wang, D., Zhao, B., & Luo, Y. (2024). CAS Landslide Dataset: A Large-Scale and Multisensor Dataset for Deep Learning-Based Landslide Detection. Scientific Data, 11(1), 12.

Ye, C., Li, Y., Cui, P., Liang, L., Pirasteh, S., Marcato, J., ... & Li, J. (2019). Landslide detection of hyperspectral remote sensing data based on deep learning with constrains. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 12(12), 5047-5060.

Yi, Y., Xu, X., Xu, G., & Gao, H. (2023). Landslide Detection Using Time-Series InSAR Method along the Kangding-Batang Section of Shanghai-Nyalam Road. Remote Sensing, 15(5), 1452.

Zhang, M., Chen, F., Liang, D., Tian, B., & Yang, A. (2020). Use of Sentinel-1 GRD SAR images to delineate flood extent in Pakistan. Sustainability, 12(14), 5784.

Zhang, S., Li, R., Wang, F., & Iio, A. (2019). Characteristics of landslides triggered by the 2018 Hokkaido Eastern Iburi earthquake, Northern Japan. Landslides, 16, 1691-1708.

Appendix

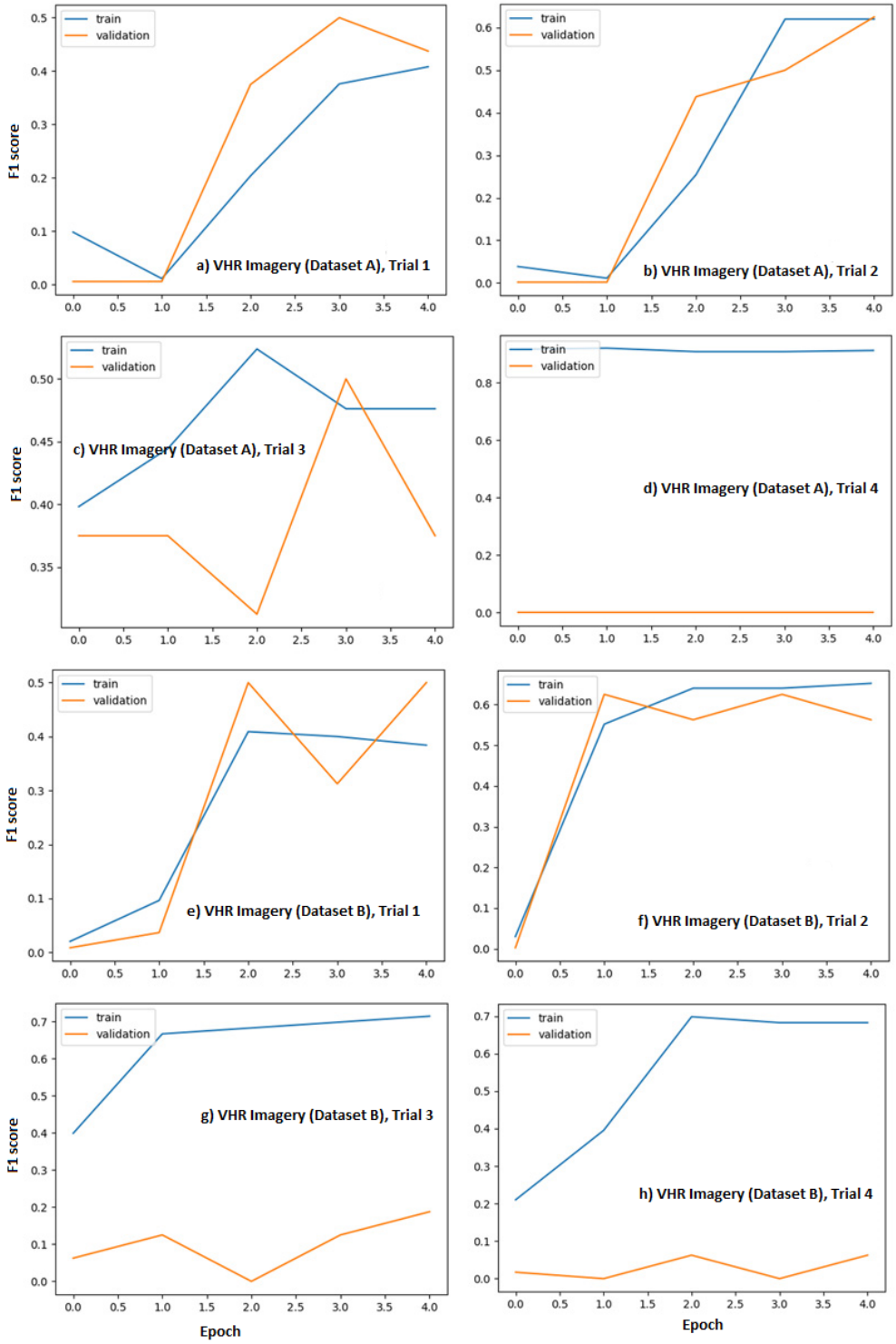

Figure A1: F1 score, over the epoch, of all 4 trials using the Datasets A and B from the best-fitted hyperparameters.

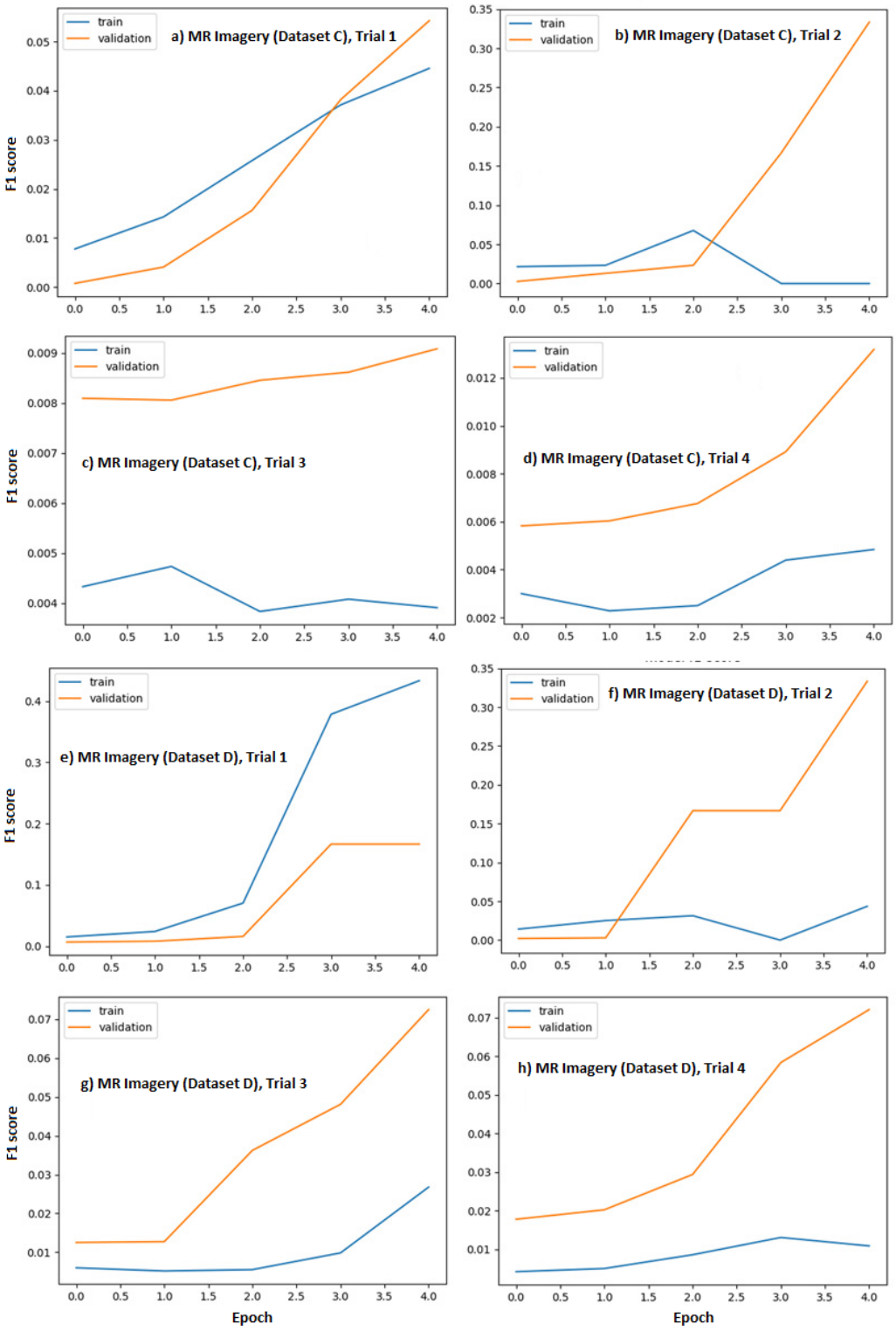

Figure A2: F1 score, over the epoch, of all 4 trials using the Datasets C and D from the best-fitted hyperparameters.

Figure A3: Loss, over the epoch, of all 4 trials using the Datasets A and B from the best-fitted hyperparameters.

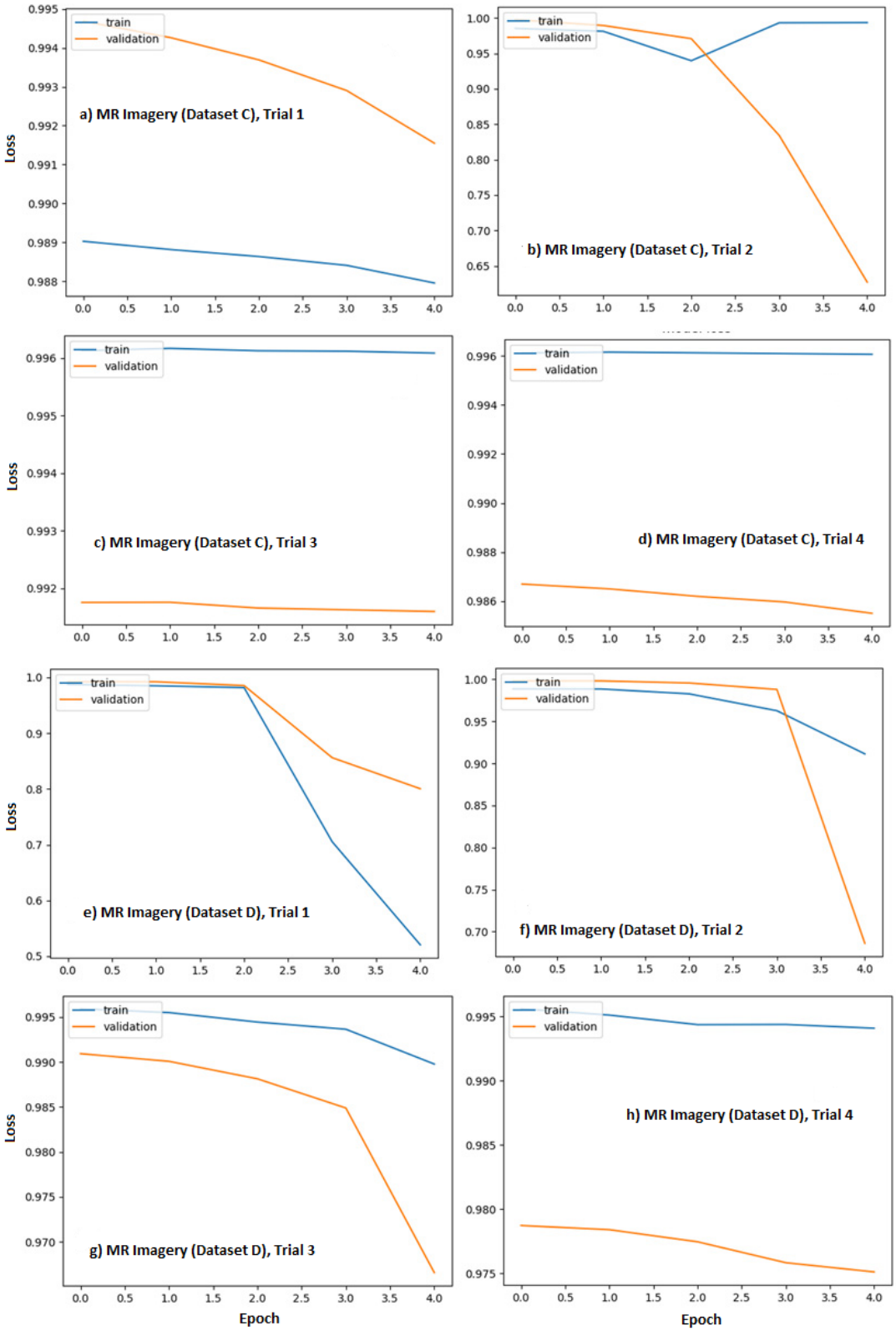

Figure A4: Loss, over the epoch, of all 4 trials using the Datasets C and D from the best-fitted hyperparameters.

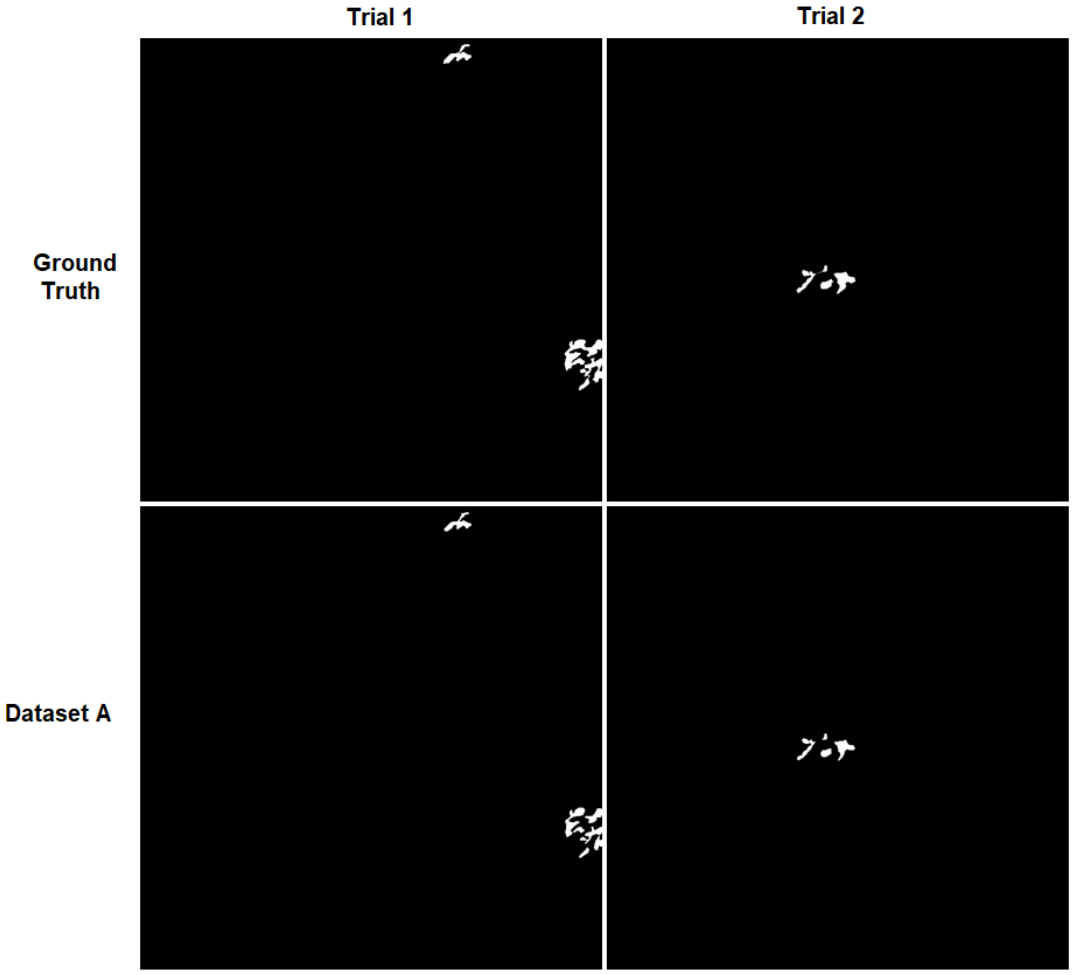

Figure A5: Comparison between the ground truth and the U-net detected landslides polygons (whole image tile) in Trials 1 and 2 from Dataset A.

Masters Program in Geospatial Technologies

Roquia Salam

Supported by:
Education and Culture
ERASMUS MUNDUS