GIScience & Remote Sensing

ISSN: 1548-1603 (Print) 1943-7226 (Online) Journal homepage: www.tandfonline.com/journals/tgrs20

Mapping landslides through a temporal lens: an insight toward multi-temporal landslide mapping using the u-net deep learning model

Kushanav Bhuyan, Sansar Raj Meena, Lorenzo Nava, Cees van Westen, Mario Floris & Filippo Catani

To cite this article: Kushanav Bhuyan, Sansar Raj Meena, Lorenzo Nava, Cees van Westen, Mario Floris & Filippo Catani (2023) Mapping landslides through a temporal lens: an insight toward multi-temporal landslide mapping using the u-net deep learning model, GIScience & Remote Sensing, 60:1, 2182057, DOI: 10.1080/15481603.2023.2182057

To link to this article: https://doi.org/10.1080/15481603.2023.2182057

Published online: 13 Mar 2023.

GISCIENCE & REMOTE SENSING
2023, VOL. 60, NO. 1, 2182057
https://doi.org/10.1080/15481603.2023.2182057

Taylor & Francis
Taylor & Francis Group

Mapping landslides through a temporal lens: an insight toward multi-temporal landslide mapping using the u-net deep learning model

Kushanav Bhuyan [a,b], Sansar Raj Meena [a,b], Lorenzo Nava [a], Cees van Westen [b], Mario Floris [a] and Filippo Catani [a]

[a]Machine Intelligence and Slope Stability Laboratory, Department of Geosciences, University of Padova, Padua, Italy; [b]Centre for Disaster Resilience, Department of Applied Earth Sciences, Faculty of Geo-information Science and Earth Observation (ITC), University of Twente, Enschede, the Netherlands

ABSTRACT

Repeated temporal mapping of landslides is essential for investigating changes in landslide movements, legacy effects of the landslide triggering events, and susceptibility changes in the area. However, in order to perform such investigations, multi-temporal (MT) inventories of landslides are required. The traditional approach of visual interpretation from cloud-free optical remote sensing imageries is time consuming and expensive. Recent endeavors exploring Convolutional Neural Networks and deep learning models have made rapid and accurate mapping of landslides feasible but have not been applied for multi-temporal landslide mapping in the Himalayas, yet. Earlier models used a standard supervised learning approach, with a small landslide inventory over a limited area used for training, which is then utilized to predict landslides in nearby areas. We propose a new strategy, using geographically separate training samples to design a standard approach which can be utilized to create multi-temporal landslide inventories. RapidEye images of 5-meters spatial resolution are used to generate MT landslide inventories in the study area of Rasuwa district, Nepal. We test the effectiveness of the model by training with only 55 landslides and predicting for a different area. Then, using the weights attained from this first training phase, we use transfer learning to map landslides over a time period between 2013 and 2019 in the Rasuwa district attaining an average F1-score of 0.69 for the study area. We also perform a spatial comparison between the manual (observed) and predicted inventories to evaluate the differences between landslide densities and overall landslide statistics of landslide area distribution. The benefit of a transfer learning-based model training is that it circumvents the need for generating annual inventories for training a deep learning. A single event-based inventory is enough to generate landslide inventories over a number of years, at least until landslide conditioning factors do not change significantly. This application can enable automated workflows to generate MT landslide inventories of particular areas as the basis for landslide evolution and movement change analysis.

ARTICLE HISTORY
Received 22 July 2022
Accepted 9 February 2023

KEYWORDS
Multi-temporal; deep learning; U-net; Himalayan; landslide inventories

1. Introduction

Landslides are major causes of loss to life, livelihood, and property due to their destructive nature and dynamic behavior. Slope failures are caused by a combination of inherent slope, soil, and geomorphic process elements as well as triggering factors such rainfall, earthquakes, and human activity (Ozturk et al. 2022; Serey et al. 2019; Wang et al. 2019).

Landslide inventories are the foundation for evaluating the hazard and risk induced by land sliding (Van Westen et al. 2013; Metternicht et al., 2005; Soeters & Van Westen, 1996; Sreedevi & Yarrakula, 2016). With the help of inventories, we can store crucial information related to the time of occurrence, type and the initiation and runout components of landslides. Incomplete and inaccurate landslide inventories can seriously affect the reliability of hazard and risk maps, and the availability of reliable and fast landslide mapping methodologies is fundamental. These inventories carry information about the landslide date of occurrence, triggering factors, volume of the failure, and type of failure, which enables us to understand the spatial-size distribution and characterization of the landslides (Catani et al. 2005; Hervás and Bobrowsky 2009; Korup, Görüm, and Hayakawa 2012; Pokharel, Alvioli, and Lim 2021; Rana, Ozturk, and Malik 2021, 2022). Both point- and polygon-based landslide inventories are essential in applications

CONTACT Sansar Raj Meena s.r.meena@utwente.nl

related to developing early warning systems or near-real time predictive tools but the latter is useful for rapid landslide hazard assessment (Tanyaş et al. 2017).

According to Guzzetti et al. (2012), landslide inventories can be generated using tools and techniques such as 1) field surveying, 2) remote sensing image interpretation, and 3) image processing tools for image classification and/or segmentation. However, here are still not enough effective and reliable ways to quickly gather data on landslide occurrences, despite all the technologies that have been devised and tested, especially to map landslides over a given period of time. Because of this, the majority of landslide inventories that are currently available are either historical (all past and present occurrences are mapped together without information on activation period) or event-based products (only recently failed landslides are mapped as related to a triggering event, without reference to previous occurrences and potential reactivations of the same). An operational capacity for a real multi-temporal (MT) landslide mapping technique is required to circumvent this limitation. Particularly in the most repeatedly susceptible regions, a standard and openly accessible database information with polygon-based recorded landslides with information or documentation of previous landslides is essential. This is where multi-temporal landslide (MTL) inventories come handy and allows for many applications including but not limited to 1) developing dynamic hazard tools, 2) landslide mobility studies, 3) exploring landscape evolution process with respect to the landslides as well as in response to climate changes (Althuwaynee, Pradhan, and Lee 2012; Chen et al. 2019; Dahal et al. 2022; Jones et al. 2021; Tanyaş et al. 2021), that 4) enables studying the spatiotemporal evolution of landslides to understand the "legacy effects" of the triggering event (Guzzetti et al. 2005; Fan et al. 2017; 2019, 2021; Tang et al. 2016; Tanyaş et al. 2021). MTL inventories can also help in designing rainfall thresholds based on inventories of event-based rainfall-induced landslides. Based on the assessment of rainfall thresholds obtained from evaluating rainfall circumstances that have caused landslides to occur, empirical approaches for determining the temporal probability of landslides are used (Jaiswal and van Westen 2009). But the generation of MTL inventories is difficult due to the subjectivity in mapping the landslides manually and the time that goes into producing these inventories (Van Westen et al., 2006; Meena & Piralilou, 2019).

When using automated approaches, the subjectivity is limited as there is a single model that maps the landslides instead of multiple human annotators, thereby reducing the overall bias and time that stems from manual interpretation. In the last few years, landslide mapping has seen a rapid development with techniques using a combination of Earth Observation data, topographic factors (Ghorbanzadeh et al. 2021; Meena et al. 2022), and advanced machine learning (ML) and deep learning (DL) algorithms (Fang et al. 2021; Prakash, Manconi, and Loew 2021). DL algorithms are used both for landslide mapping and landslide susceptibility mapping however, they are quite different in terms of what is actually being modeled. In landslide mapping, the models predict landslide traces or footprints after the event has occurred which effectively becomes an inventory however, the latter deals with predicting the probability of spatial occurrence of landslides in a given region (preferably with the intensity) and thus, talks about spatial probability of landslides in the future (Spinetti et al. 2019; Azarafza et al. 2021). In our work, we focused on mapping or detecting landslides after they have occurred. DL algorithms, especially Convolutional Neural Networks (CNNs), have been successfully employed, demonstrating wider generalization capabilities when compared to other ML models. Ghorbanzadeh et al. (2019) used high-resolution Rapid Eye data to test several ML techniques, including support vector machines (SVMs), random forest (RF), artificial neural networks (ANNs), and deep convolutional neural networks (D-CNN), with CNNs attaining the best results. Meena et al. (2021) used a composite of optical RapidEye images and topographical maps with CNNs to obtain a mean F1-score of 78% to map rainfall-induced landslides. The popular U-Net model (Ronneberger, Fischer, and Brox 2015) and its subsequent variants have been extensively used by the community to map landslides with remote sensing images (Meena et al. 2022; Nava et al. 2022; Tang et al. 2022) due to their strong capability to retain and project discriminative (low-resolution) features to the higher resolution (pixel level) in order to get a dense classification. However, these models are data hungry and require a lot of data to be able to effectively train a deep learning model and detect landslides accordingly.

Deep transfer learning is an approach which assists in the optimization of model performance on a small number of samples as real-world situations make it challenging to obtain adequate training data. The theory of transfer learning to apply information obtained from prior experiences to new contexts. As transfer learning has shown a great potential in solving the data deficiency problem and empowering knowledge transfer between machine learning tasks (Pan and Yang 2010), these objectives are important for addressing the data scarcity issue and enabling knowledge transfer between deep learning tasks. Detecting landslides is considerably more effective and training time is much decreased when the model is retrained on newer regions after learning about edge, shape, and texture detections from prior "experiences" (Ayana, Dese, and Choe 2021). Deep transfer learning was first proposed by Wang, Wang, and Zhang (2022) but for landside susceptibility purposes (which is the spatial probability of landslide occurrence in a given region) and is very different from detection or mapping purposes, and the mapping/detection of landslides both over space and time is still unexplored. Although a wide array of research in the domain of landslide mapping has been published with artificial intelligence, a robust model that is able to map landslides over a period of time has yet to be designed and investigated.

Therefore, we propose with the simplest of model and training approach with the U-Net model to investigate the detection of landslides over a period of years which includes mapping the pre-, co-, and post-seismic landslide, which are the main contribution and motivation of the paper. Illustrated for the Mailung area of Nepal, affected by the 2015 Gorkha earthquake, this study presents an opportunity to map landslides over the years between 2013 and 2019, which could be later helpful for post-seismic landslide evolution studies in the future.

2. Study area and data

2.1. Study area

The study area is located in central Nepal's higher Himalayan district of Rasuwa and is among the most landslide-affected region (see Figure 1). The most common land cover is forest, grassland, shrubland, farmland, and rural regions. With an annual average rainfall of 691 mm, this region's climate is influenced by orographic monsoon precipitation. A Mw 6.7 earthquake occurred the same day as the primary shock, a Mw 6.9 earthquake the next day, and a Mw 7.3 earthquake on 12 May 2015. The primary shock caused significant ground shaking, reaching a maximum Peak Ground Acceleration (PGA) of 0.87 g, according to the United States Geological Survey (USGS) (USGS 2015). Landslides were triggered after the 2015 Gorkha earthquake that dammed the river, which resulted in the formation of multiple lakes behind the dams in various locations. The leading cause of severe flash floods and monsoonal rains is water obstruction behind landslide-induced dams (see Figure 1). Major part of the study area is located in the Langtang national park which consists of several hydropower plants in the Trishuli river that was damaged by the earthquake event. In April 2015, more than 80 lives were perished as a result of landslides and rockfalls near the hydropower project construction camps in Mailung village. The damages resulted in a drop in energy output as well as significant economic damage (Schwanghart, Ryan, and Korup 2018). Local authorities and foreign institutions have made concerted efforts to research the effects of landslides on the human settlements and hydroelectric power stations of the Mailung area. However, field visits were not practical in many inaccessible mountainous places; as a result, remote sensing techniques can supplement field visits coupled with deep learning methods.

Landslides in the areas affected by the Gorkha earthquake were mapped by Kargel et al. (2016), Martha et al. (2016), and Roback et al. (2018). Rosser et al. (2021) also monitored the landslide evolution of new post-seismic landslides by manually generating detailed time-series landslide maps. Therefore, the preface and importance for multi-temporal mapping is witnessed and realized.

2.2. Data

We generated annual landslide inventories of an area of 33 km^2 by manually digitizing landslide polygons (Table 1) on seven scenes of RapidEye images (WGS 1984 UTM Zone 45N) (Planet Team 2017) acquired over the years between 2013 and

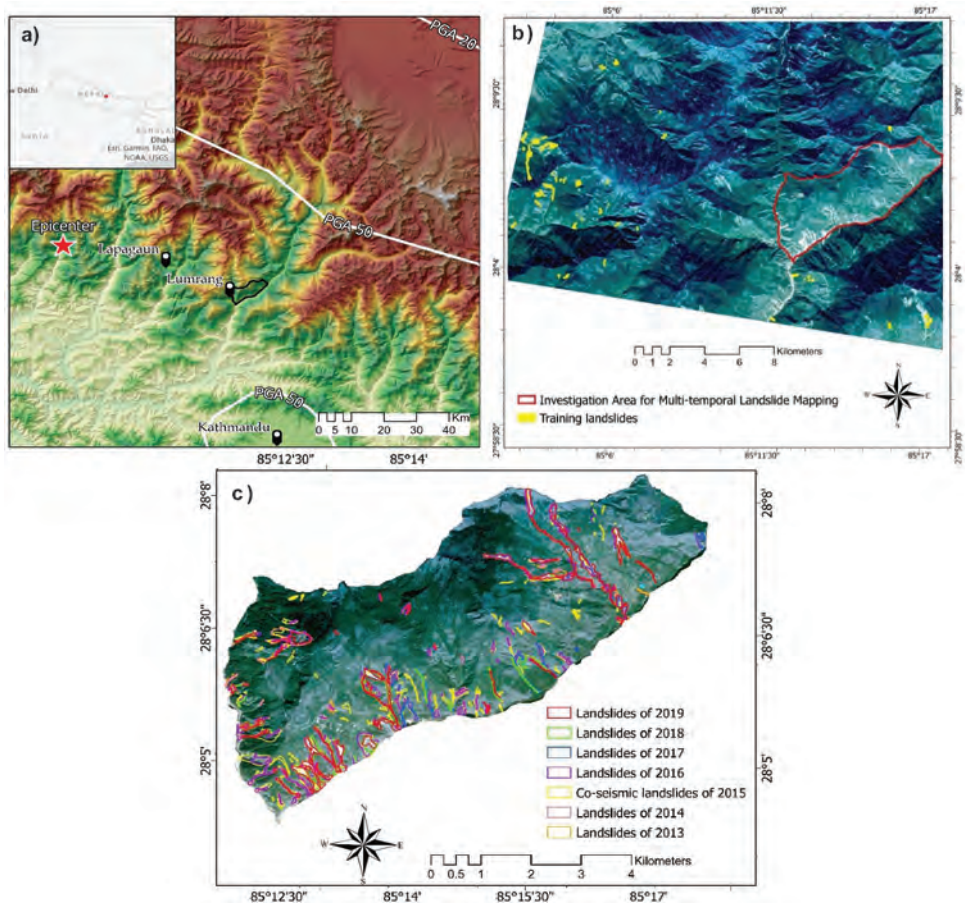

Figure 1. a: Study area location with distance from epicentre. The PGA values are in percentages. b: Investigation area (red outline) with landslides (yellow) used as preliminary training data for the deep learning model. c: MTL manually mapped over the years between 2013 and 2019 in the investigation area.

Table 1. Information about the satellite images from RapidEye and the respective landslides.

Gorkha Earthquake of 2015					
			Landslide Area (m^2)		
	Image acquisition dates	Number of landslides	Total	Minimum	Maximum
Pre-seismic (2013)	07-11-2013	31	513,304.8	216.3	148,948.1
Pre-seismic (2014)	30-11-2014	26	438,778.5	515.7	150,759.3
Co-seismic (2015)	09-11-2015	136	1,855,911.8	276.4	145,380.5
Post-seismic (2016)	04-11-2016	95	1,796,221.2	768.3	137,397.8
Post-seismic (2017)	12-11-2017	63	1,188,037.6	724.5	141,102.3
Post-seismic (2018)	24-10-2018	55	1,315,124.4	724.5	141,102.3
Post-seismic (2019)	10-11-2019	52	1,222,396.5	724.5	112,723.0

2019. For the analysis, we chose the best available cloud- and shadow-free images. The product level for the orthorectified images is "Analytic SR," which means that surface reflectance is represented by pixel values with a bit depth of 16 bits. In order to construct an image that encompassed the study area, we pre-processed the photos from each year's collection by band extraction, sharpening, and mosaicking.

Two geomorphological experts manually digitized the landslides using the satellite imageries. The co-seismic landslides and post-seismic enlargements of the landslide areas, and new failures emerging after the earthquake were all mapped using polygon-based mapping by the geomorphological experts. The seven scenes have 5 meter pixel resolution and five bands of Red, Green, Blue, Red-edge, and Near-infrared. Apart from these landslide inventories,

landslides were also digitized outside the investigation area (yellow polygons in Figure 1) to be used as initial training data. The conceptualization of the research is shown in Figure 2.

NoData is a common occurrence in remote sensing images and is harmful when used in machine learning or deep learning since it lowers model performance as the model trains over time. NoData values can cause gradient explosions since they have no spectral contribution at all. This causes enormous erroneous gradients to build up during model training and causes delayed updates to the weights of the networks, which makes the model unstable and unable to learn from the training data (Challa, Niu, and Johnson 2020). As a result, the NoData sample data were discarded.

3. Methodology

3.1. Curation of the data

We mapped 55 landslides for 2016 for the first training phase which are located outside of the area of investigation (Figure 1b yellow polygons) to avoid bias. We designed a preliminary training data set using landslides that occurred outside the investigation area. Following this, the landslides inside the investigation area are randomly split into *training* (70%) and *testing* (30%) set for the year of 2016. The *training* set is employed for the transfer learning approach, which is then tested on the *testing* set, and we utilize all of the landslides that occurred inside the investigation area as the basis for evaluating the model for the years 2013 till 2019 (except for 2016).

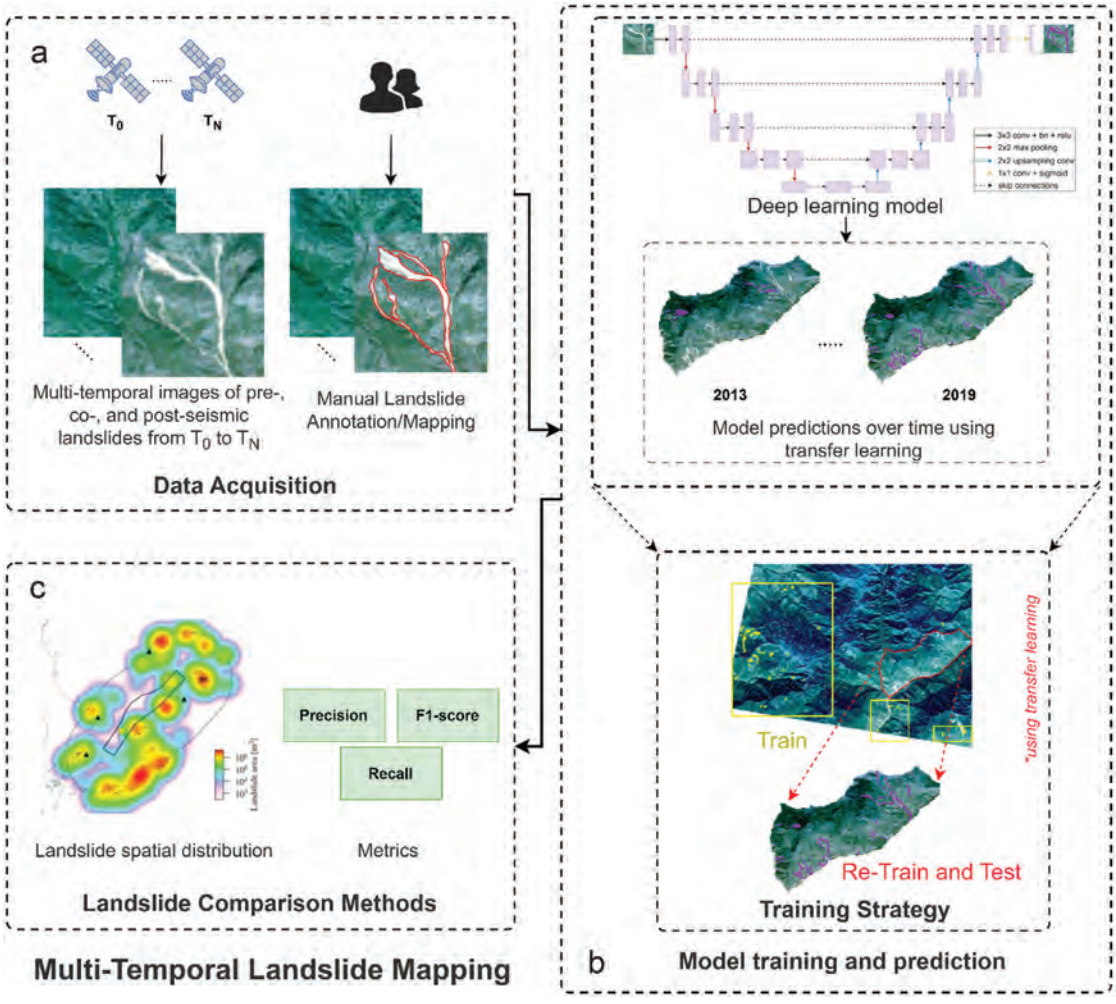

Figure 2. Conceptual diagram of the methodology. a: data acquisition of satellite images between 2013 and 2017, and manually annotated landslides of the same periods. b: model training and prediction using transfer learning. c: comparison methods with classical metrics and landslide spatial distribution.

We extracted patches of 128 × 128 pixels from the input satellite images, as suitable input for the U-Net model based on Ghorbanzadeh et al. (2021) and Prakash, Manconi, and Loew (2021), who reported optimal accuracies of F1-score, Precision, and Recall using this patch size. Furthermore, neither increasing the patch size to 256 × 256 nor decreasing it to 64 × 64 improved the model accuracy in our experiments. By rasterizing the manual inventory of the different years with 5-meter grid cells, the associated binary masks were created. Data augmentation is also adopted to artificially expand the training samples by applying image transformations. The Python library image that implements these strategies was employed (https://imgaug.readthedocs.io/en/latest/). After running multiple tests to determine the most effective combinations, we employed flipping (horizontal and vertical), blurring (gaussian blur: = 0 and 3), sharpening (= 0 and 1), shearing (with factors −20 and 20), and rotating (45° and 90°) to be more exact. These augmentations help in diversifying the training dataset and aid in regularizing the model to better generalize landslide features (Kukačka, Golkov, and Cremers 2017; Shorten and Khoshgoftaar 2019). We applied these augmentations respectively to the satellite images and the corresponding binary masks. In order to reduce the imbalance in the sample distribution between the positive (landslide) and negative (non-landslide) classes, we additionally deleted zero-valued patches (training patches with no landslide information) (Sun, Wong, and Kamel 2009).

So, finally, by using a patch size of 128 × 128 and applying data augmentation, we got a total of 500 patches (located outside of the investigated area) and used them for the first model training. Then a total of 860 image patches were created (after augmentations) for the next re-training process (inside the investigation area). As mentioned in previously, we split the datasets into training (70%) and testing (30%) inside the investigation area in 2016, so the training set consisted of 602 patches and the testing set included 258. But again, since for the other years of 2013–2015 and 2017–2019 we used all the landslides inside the investigation area for testing purposes, we used 860 patches for testing for those years.

3.2. Automating landslide detection

3.2.1. Model architecture and tuning

The U-Net model (Ronneberger, Fischer, and Brox 2015) has been used extensively for landslide detection due to its robust network structure and segmented pixels as outputs (Ghorbanzadeh et al. 2021; Prakash, Manconi, and Loew 2021; Zhang et al., 2018). The model consists of input parameters such as the convolution and max pooling operations that reduce the dimensionality of the input data (called down sampling) to extract complex features like shape, patterns and edges. Since the output needs to be a high-resolution image, we adapt transposed convolution operations, also called deconvolution, coupled with concatenation operations to retain the dimensions and information that were lost during down sampling. The U-Net model (Figure 3) has many advantages. One of these is that it extracts local features through skip connections between the encoder-decoder stages. As spatial details tend to get lost at the deepest end of the encoder stages during model training, the decoder stage with the help of the skip connections retrieve the relevant spatial information from the low-level features and provides per-

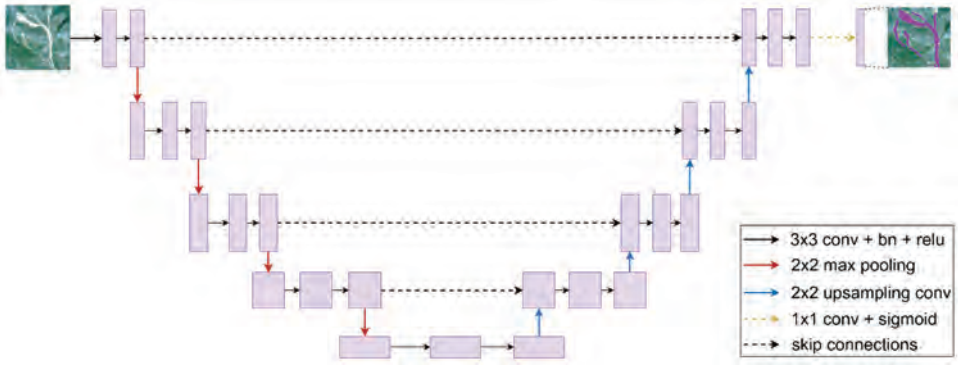

Figure 3. An overview of the U-Net model used in our research based on (Ronneberger, Fischer, and Brox 2015).

pixel segmented results. In this study, we use a deep U-Net model with 5 convolutional blocks containing 10 convolutional layers in total.

In this study, the Adam optimizer was employed as advocated by Bottou (2010) and Pan et al. (2020) due the adaptive learning capability of the optimizer which allows faster convergence to decrease the loss, thereby improving model accuracy. Different learning rates are investigated as well, within the Adam optimizer, to enhance training speed and balance model overfitting. As an outcome of this stage, heat maps of probability pertaining to the classes landslides and non-landslides are generated. After training, the outcome is a binary image that differentiates between *landslides* and *non-landslides* pixels. The U-Net model training was conducted in the Python environment on an NVIDIA RTX 3060 GPU (6 GB VRAM) and 16 GB of RAM.

One of the most critical processes in regulating the model's general behavior is hyper-parameter tuning. The aim is to identify the optimal hyper-parameter combination that minimizes the loss and produces the best result. We investigated different number of filters such as 8, 16, and 32 (which are multiplied by a factor of 2, 4, 8, and 16 in the succeeding convolutional sequences and then divided by a factor of 8, 4, and 2 after/at the deconvolutional/upsampling sequences in the decoder path) as more filters can potentially allow learning more features and patterns from the data, which leads to better performance. The Tversky Loss (equation 1) (Abraham and Khan 2019) function was applied in this study. Using so-called beta weights, the Tversky loss has the benefit of immediately modifying and adjusting the False Positives and False Negatives. The alpha and beta parameters of the Tversky loss function regulate the false positives and false negatives, respectively, thereby impacting the overall prediction capability of the model. This parameter helps decrease model loss when training to obtain improved accuracy by adjusting the imbalance between the data within *landslides* and *non-landslides* classes.

$$Tversky\ Loss = \frac{TP + \varepsilon}{TP + \alpha \times FN + \beta \times FP + \varepsilon} \quad (1)$$

where,

TP = True Positives
FP = False Positives
FN = False Negatives
ε = A constant value of 0.0001 (by default) which prevents the loss from becoming infinite.
α = Alpha parameter that adds weight to the FNs.
β = Beta parameter that adds weight to the FPs.

3.2.2. Training and transfer learning approach

The transfer learned model will adapt on top of these "recognized" features when the model is tasked on the region with different seasons where the radiometric returns will be different for likely the same types of landslides, saving computing time and detecting landslides more effectively by increasing the dataset diversity. Essentially, we intend to increase generalization power while minimizing overfitting by training a model that has already been trained in area A on landslides of region B (plus over different years). We adopt a transfer learning mechanism to train a temporally generalizable model that would be able to detect and map landslides over time. The goal of transfer learning is to transfer the information from previous data and apply what the model has learnt in a new environment, which might be difficult to learn in otherwise. The weights from a prior model can be used in a new region of interest, with the network learning on top of the pre-trained model and retraining an output layer using the target landslide data set. This strategy can reduce the model's training time and increase its effectiveness in a new region (Bai et al. 2012; Xu et al. 2013). The U-Net model was first trained for 250 epochs using the preliminary training data set (yellow boxes in Figure 2b) and then tested within the investigation area (red boundary in Figure 1b) for the year 2016 (one year after the Gorkha earthquake) inclusive of all the landslides. We used checkpoints to save the model with the lowest loss. In the next phase, we first sampled the landslides inside the investigation area into two sets: a training set (70%) (for re-training) and a testing set (30%) just for the year 2016. We then re-trained the pre-trained model on the training set of 2016 and tested on the testing set. After transfer learning on 2016, we performed predictions and tested the model on the entire landslide data within the red boundary for the years between 2013 and 2019 (which varies both temporally and spectrally) except of course for 2016. Weight initialization was not necessary because pre-training weights were already specified. We also choose satellite images from the same late autumn/early winter season with

low cloud cover to maintain similar spectral characteristics. Spectral differences are witnessed in the temporal image acquisitions (Anderson and Perry 1996; Huete 2004), and coupled with image distortions for each acquisition, modeling the landslides was a challenging task as we see later in section 4.

3.3. Accuracy assessment

3.3.1. Classical metrics

The model's prediction outputs are binary maps of landslide areas, which are then compared against manually mapped landslides (ground truth) using the Precision (Equation 2), Recall (Equation 3), and F1-score (Equation 4) accuracy metrics. These metrics are calculated using True Positives (TPs), False Positives (FPs), and False Negatives (FNs), where, TPs are accurately identified landslide areas, FPs are non-landslide areas being detected as landslide areas, and FNs are landslide areas that were missed out by the model. *Precision* here refers to how well the model detects the landslide class. *Recall* is the number of times that the model detects the landslide class, and *F1-score* finally is the harmonic mean of (2) and (3) and acts as a balance between the two.

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{4}$$

3.3.2. Landslide statistics and spatial distribution

The inventories used for training and testing were generated manually using visual image interpretation (See Table 1). Landslide statistics were evaluated for the manually annotated and predicted landslides for each respective year between 2013 and 2019 using information such as the total landslide area, maximum and minimum landslide area. Landslide densities were also analyzed using the number of landslides per square kilometres for each year similar to that of Meena and Piralilou (2019). The landslide density (LD) gives us an idea about the spatial location and changes in the spatial distribution of the pre-, co-, and post-seismic landslides for the study area. Therefore, assessing the differences between the manual and predicted landslide inventories through the lens of LD is important to observe the overall mapping performance of the model on the pre-, co-, and post-seismic years.

4. Results

4.1. Model training and parameterization

We compared the results against manually annotated landslides for accuracy assessment after using transfer learning for the different years. The model was evaluated on the test set using the classical evaluation metrics of Precision, Recall, and F1-score, and using the change in LD. As discussed previously, we trained the model with landslides outside the investigation area (Figure 2) and tested on the investigation area of the year 2016 (one year after the Gorkha event). The results of the metrics can be seen in Table 2. After that, we used transfer learning with pre-trained weights to re-train in our investigation area with newer landslide instances for each year from 2013 till 2019 (see Figure 4). In order to remove and filter out insignificant landslides detections by the model (which show a typical random effect of individual pixels in a so-called salt and pepper effect), we used a threshold area of 200 m^2 to filter out these isolated pixels, and thereby, cleaning the overall results.

The appropriate combination of hyper-parameters allows the model to attain the optimal performance and thus to yield the highest possible mapping accuracy. As we see in Table 2, the best result of F1-score of 78.5% is achieved by the combinations of number of filters 32, batch size 32, and a learning rate of 0.001. This set of combinations were hence chosen for the re-train and testing using transfer learning on the other years inside the investigation area.

Based on numerous experiments for the optimal β value in the Tversky Loss function, $\beta = 0.7$ proved best for the results of mapping the landslides over time. This is because setting β values higher than 0.7 give stronger attention on the FPs thus heavily reducing the Recall and/or reducing the Precision. In our study case, $\beta = 0.7$ gave the best scores in terms of a balanced FP and FN, thereby mitigating the imbalance between Precision and Recall.

We can see in Figure 4 the various landslide footprints that were detected by the model for the years

Table 2. Table of various hyper-parameter combinations based on the preliminary training data and test data of 2016 (**bold** are the best combinations results).

Learning Rate	Number of filters	Batch Size	Loss	Precision	Recall	F1-score
1e-3	8	8	0.227	0.807	0.698	0.748
1e-3	16	8	0.218	0.817	0.702	0.755
1e-3	32	8	0.206	0.845	0.688	0.757
1e-3	8	16	0.229	0.805	0.701	0.749
1e-3	16	16	0.208	0.826	0.719	0.769
1e-3	32	16	0.226	0.865	0.618	0.721
1e-3	8	32	0.233	0.803	0.700	0.746
1e-3	16	32	0.205	0.824	0.725	0.771
1e-3	**32**	**32**	**0.188**	**0.850**	**0.731**	**0.785**
1e-4	8	8	0.559	0.710	0.678	0.694
1e-4	16	8	0.258	0.792	0.651	0.714
1e-4	32	8	0.220	0.821	0.699	0.755
1e-4	8	16	0.602	0.665	0.777	0.716
1e-4	16	16	0.325	0.791	0.607	0.687
1e-4	32	16	0.237	0.813	0.695	0.749
1e-4	8	32	/	/	/	/
1e-4	16	32	0.364	0.755	0.750	0.752
1e-4	32	32	0.242	0.818	0.702	0.755
1e-5	8	8	0.850	0.402	0.958	0.566
1e-5	16	8	0.753	0.547	0.892	0.678
1e-5	32	8	0.432	0.801	0.654	0.720
1e-5	8	16	0.838	0.332	0.955	0.493
1e-5	16	16	0.792	0.527	0.858	0.653
1e-5	32	16	0.509	0.754	0.753	0.752
1e-5	8	32	0.836	0.196	0.989	0.327
1e-5	16	32	0.813	0.477	0.898	0.623
1e-5	32	32	0.710	0.583	0.882	0.702

between 2013 and 2019, tallied against the respective manually annotated landslide footprints.

4.2. Generated multi-temporal inventories and their statistics

The spatial distributions of the manual inventories (MI) and predicted landslide inventories (PI) were compared using landslide statistics and LD for each year. The findings are presented in Table 4.

Table 4 shows that the minimum landslide area is almost always around 200 m^2, as this value was chosen as threshold to avoid mapping of individual pixels. A single landslide in the manual interpretation can sometimes be predicted as multiple instances in the predicted inventory, as the possible functional connection between them (e.g. along a debris flow channel) cannot be analyzed automatically. This is a very common issue as seen in the works of Prakash, Manconi, and Loew (2021) and Zhang, Pun, and Liu (2021) and thus post-processing approaches should be employed to reduce the problem. The opposite is also witnessed in the inventories of 2015 where we see that MI had a higher number of landslides than PI. The reason behind this is the fact that some landslides were not detected by the model, as explained in the low F1-score of 58.6% (Table 3), thereby resulting in lower total number of landslides (T_L). However, the maximum number of landslides ($Max\ A_L$) and the total area of landslides ($T\ A_L$) give a more positive and optimistic overview of the detected landslides in general, since they are comparable across all the investigated years. Moreover, we also obtained more landslides that are actually missing in the inventory (see Figure 5). Although this is not the case all the time but only in certain situations and locations where mapping manually could be quite confusing and are sometimes left out (due to subjectivity of the mapper). For example, Meena and Piralilou (2019) had compared the inventories mapped by four different mappers for the same area and the differences between the manual mapping were quite large. As seen in the example from 2017 in Figure 5, where the blue polygons are the MI and the landslides were not mapped properly. The red polygons are from the PI and we see that the model picked these landslides out quite effectively. So, we see that sometimes the model accurately detects real traces of landslides but since these are not present in the MI, the FPs inadvertently increase, however, checking these FPs can lead to a better understanding of the predictions and often times lead us to detect missed landslides which

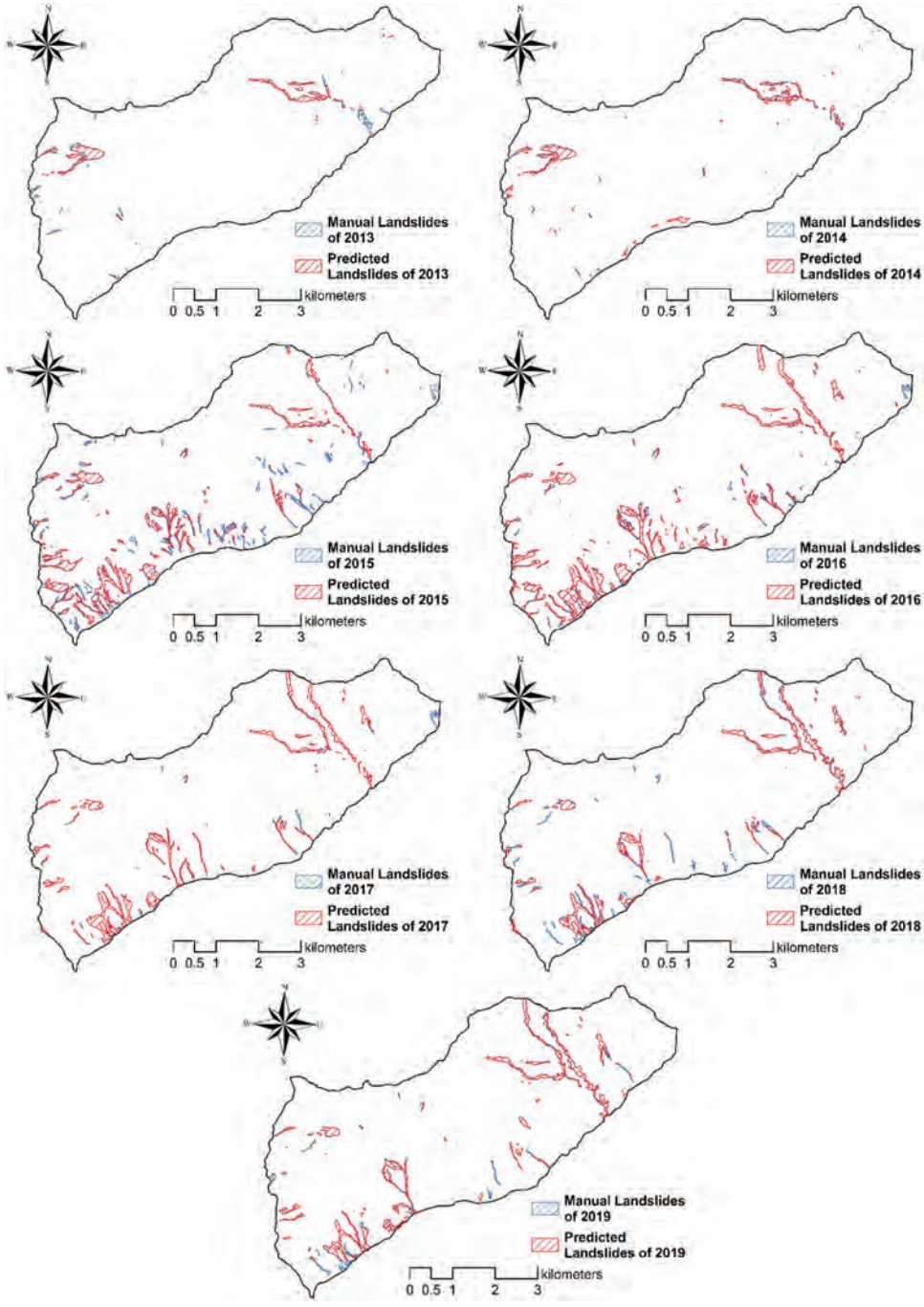

Figure 4. Predicted landslides versus the manually delineated landslides for the years 2013 till 2019.

Table 3. Table of results for each year between 2013 and 2019 (using 32 filters, a batch size of 32 and a learning rate of 0.001).

Year	Loss	Precision	Recall	F1-score	# detected landslides
2013	0.320	0.752	0.651	0.697	38
2014	0.165	0.762	0.696	0.727	53
2015	0.340	0.843	0.450	0.586	108
2016	0.205	0.927	0.603	0.731	117
2017	0.175	0.882	0.724	0.795	73
2018	0.290	0.753	0.633	0.688	75
2019	0.362	0.600	0.759	0.669	87

Table 4. Comparison of landslide statistics for the manually mapped landslide inventories versus the predicted ones.

	Manually Annotated Landslide Inventory (MI)				Predicted Landslide Inventory (PI)			
Year	Total number of landslides (T_L)	Total area of landslides ($T\ A_L$)	Minimum area of landslides ($Min\ A_L$) in m^2	Maximum area of landslides ($Max\ A_L$) in m^2	Total number of landslides (T_L)	Total area of landslides ($T\ A_L$) in m^2	Minimum area of landslides ($Min\ A_L$) in m^2	Maximum area of landslides ($Max\ A_L$) in m^2
2013	31	513,304.8	216.3	148,948.1	38	369,595.9	215.24	186,958
2014	26	438,778.5	515.7	150,759.3	53	437,334.0	206.1	170,278
2015	136	1,855,911.8	276.4	145,380.5	108	1,336,636.5	264.0	224,417
2016	95	1,796,221.2	768.3	137,397.8	117	1,633,912.1	202.6	405,951
2017	63	1,188,037.6	724.5	141,102.3	73	1,048,197.0	217.1	255,057
2018	55	1,315,124.4	724.5	141,102.3	75	1,011,351.6	202.6	254,061
2019	52	1,222,396.5	724.5	112,723.0	87	1,036,097.9	214.6	155,752

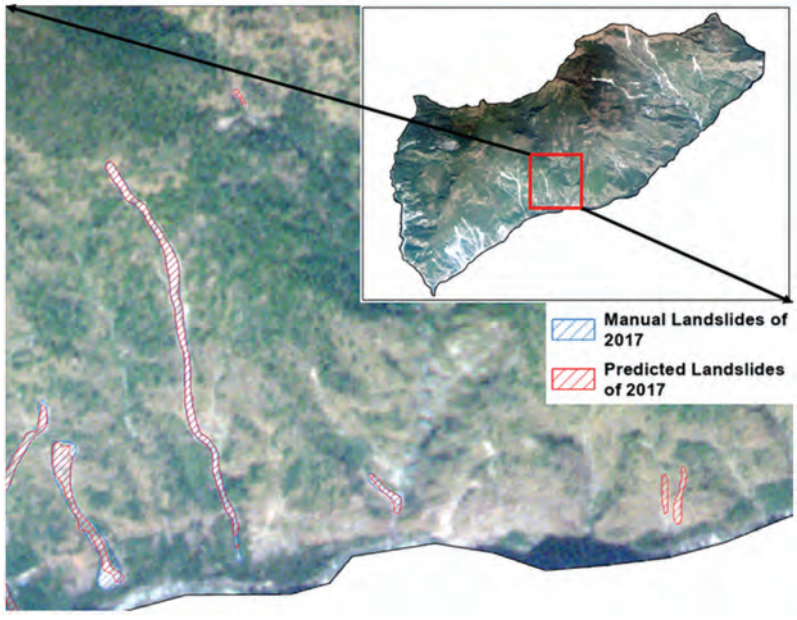

Figure 5. Example of missing landslides in the MI of year 2017. Base image: RapidEye, 2017.

were not present in the MIs. This shows that deep learning in unison with manual mappers can greatly improve inventorying.

Interestingly, we also notice that the trend of the TL and $T\ A_L$ is quite similar between MI and PI. This is also confirmed by looking at the Pearson's correlation between the MI and PI in terms of the T_L and $T\ A_L$ which is 0.85 and 0.98, respectively. The number of landslides increases in 2015 as a result of the Gorkha earthquake, then gradually declines in the subsequent years. This observation is in line with recent post-seismic landslide evolution studies (Fan et al. 2018, 2019). We can also notice that the areas and number of active landslides between MI and PI are different. This can be attributed to the previously discussed fragmentation problem in terms of the model predictions (please see Figure 6). Table 3

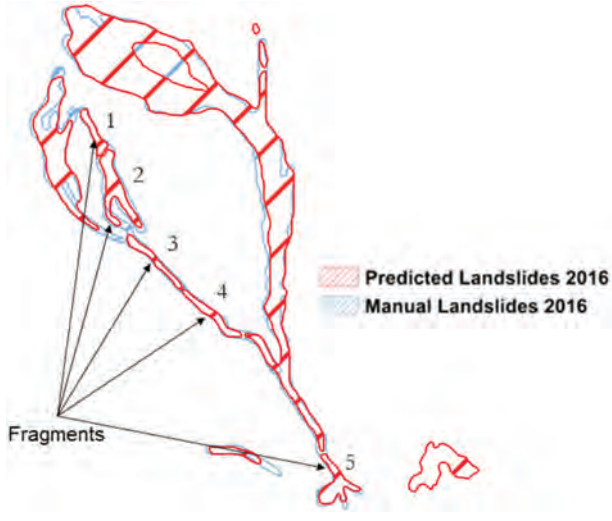

Figure 6. Fragmentation of the landslides of the predicted inventories (red) compared to the manual inventories (blue).

reflects varying results in the F1-scores, although from a spatial point of view, they are related to the same area. This is because the image acquisitions, even in the same seasons (late autumn and early winter), have different spectral reflectance/atmospheric disturbances, which confuses the model even while predicting the same landslides repeatedly. The image of 2015 was the most different spectrally, and that's why we see in Table 3 that the best F1-score is only 58%. However, for the other years, the F1-scores varies between 65% and 80%.

Figure 7 figillustrates the problems discussed above for the year 2015 where the landslide detections are relatively poor compared against the rest. Many landslides seem to be missing in the center and the eastern part of the area, which is indicative of the poor F1-score of 58%. This explains the lower value of T_L in Table 4 for the predicted inventory (108) against the manual inventory (136). As discussed previously, this can occur because of the observed differences in the spectral reflectance which reduces the efficiency of the model to predict the landslides of 2015. Much better predictions are observed in the south-western part of the map where the landslides are mapped almost identically. However, a substantial number of landslides are also missed out in the eastern part of the map for the years 2013, 2015, and 2019. Overall, the prediction over the different years very well captures the general location of the landslide footprints and gives a positive outlook toward employing DL methods for MT inventory generation.The density of the landslides of the MIs versus the PIs for the year 2013 till 2019. The density values are normalized between 0 and 1 for visual comparison purposes between the MIs and PIs.(Continued).

To understand the variations in the spatial distribution of the inventories (both MI and PI) and analyze the LD per square kilometer, we used the centroid of each landslide polygon. The MI of 2013–2019 were generated using satellite images for the same season (right after the monsoon). Based on the results in Table 4, variation in the total number of landslides as well as landslide area can be seen in the manual and predicted inventories. The $T\,A_L$ in manual inventories ranges from 438,778.5 m^2 to 1,855,911.8 m^2 and for the predicted ones, it ranges from 369,595.9 m^2 to 1,633,912.1 m^2. Moreover, the smallest and largest mapped landslide polygon varies as well, both for the manual and predicted inventories (Table 4). The LD distribution varies for manual and predicted inventories ranging between 3.22 and 2.59 landslides/km^2, respectively, in the year 2013. After 2015 earthquake event, the LD increased to 16.46 landslides/km^2 and 11.47 landslides/km^2 for the manual and predicted inventories, respectively. The trend for LD distribution for the same area declined after 2016 for the manual inventory and the same is observed for the predicted inventory (Table 5).

For two cases (2013 and 2015), the LD of the predicted landslide inventories is lower than the manual inventories while, for the rest of the years it is the opposite. While for 2013 the difference is relatively small (−0.63 landslides/km^2), this is much larger in 2015 (−4.99 landslides/km^2) as seen in Table 5. In 2015, the overall prediction is weaker when compared to the other years. In fact, since the Recall is around 45% (Table 3) for 2015, most of the landslide pixels were missed out by the model. However, for the year 2016, we observe that the LD of PI is higher than that of MI, showing a positive difference of 3.17 landslides/km^2. Notice that after the event in 2015, the LD of MI decreased (as expected due to re-vegetation in the terrain) from 16.46 landslides/km^2 to 8.37 landslides/km^2 but this behavior was not reflected in PI as the LD remained almost the same. This phenomenon can be explained by the fragmented predictions made by the model as illustrated in Figure 6. As the model predictions are pixel-based, landslide bodies may be predicted in portions. Thus, a single landslide body can be fragmented into two or more bodies yielding more portions for the same landslide body. This phenomenon leads toward an increase in the overall number of landslides per km^2. Like T_L, the overall trend of the LD for both MI and PI are similar where we first see an increase in the density in 2015 followed by decrease in the following years (except for the 2016 outlier for P).

In Figure 8, we see the scatter plot depicting the area of the landslides in log-scale of the MI in the x-axis and PI in the y-axis. We can observe that the ground truth landslides versus the overall predicted landslide area are pretty close to one another, virtually resembling the data points in a 45-degree line buffer zone. This demonstrates that the majority of each predicted landslide areas was pretty comparable to each landslide's actual MI area.

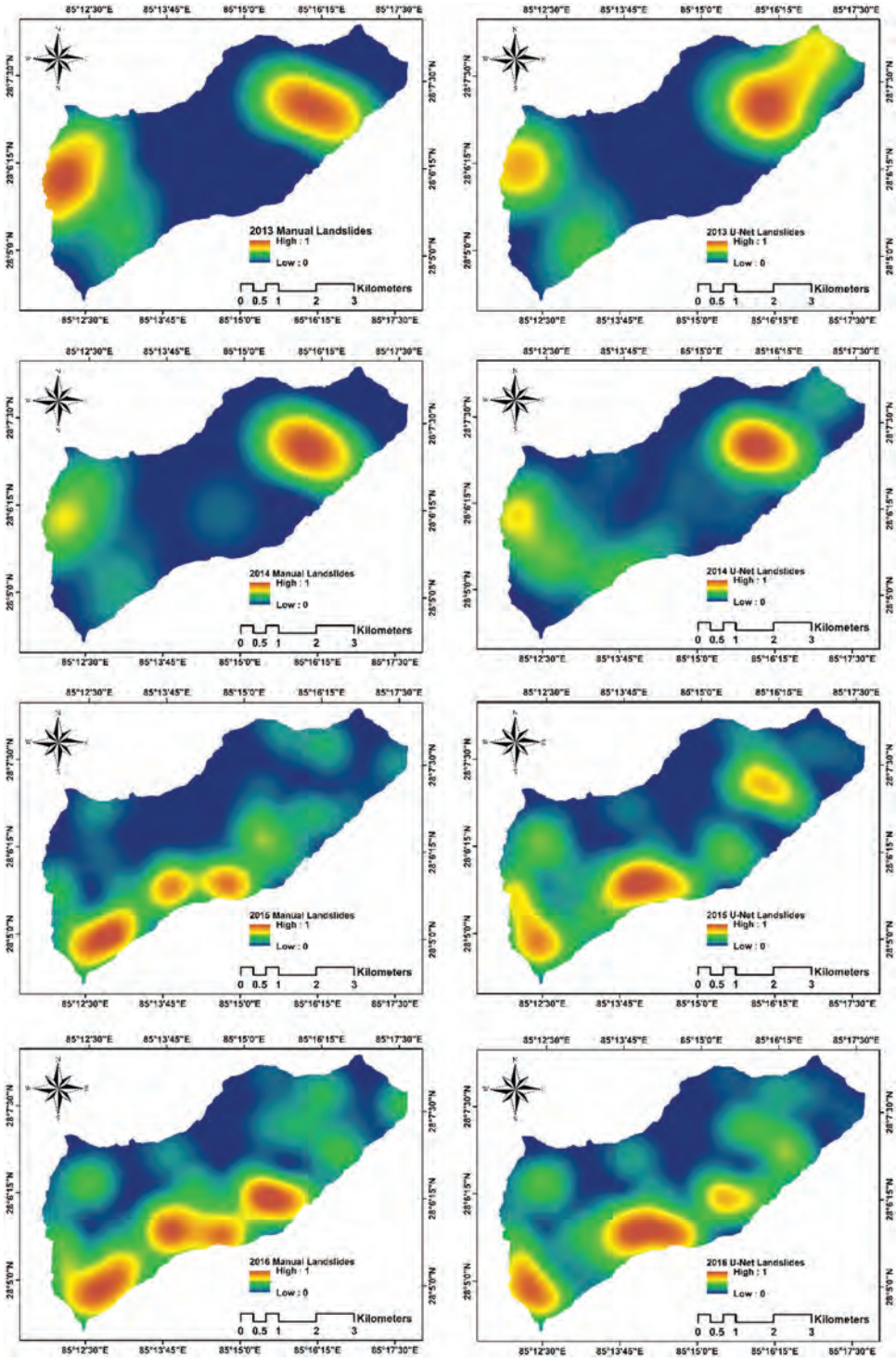

Figure 7. The density of the landslides of the MIs versus the PIs for the year 2013 till 2019. The density values are normalized between 0 and 1 for visual comparison purposes between the MIs and PIs.

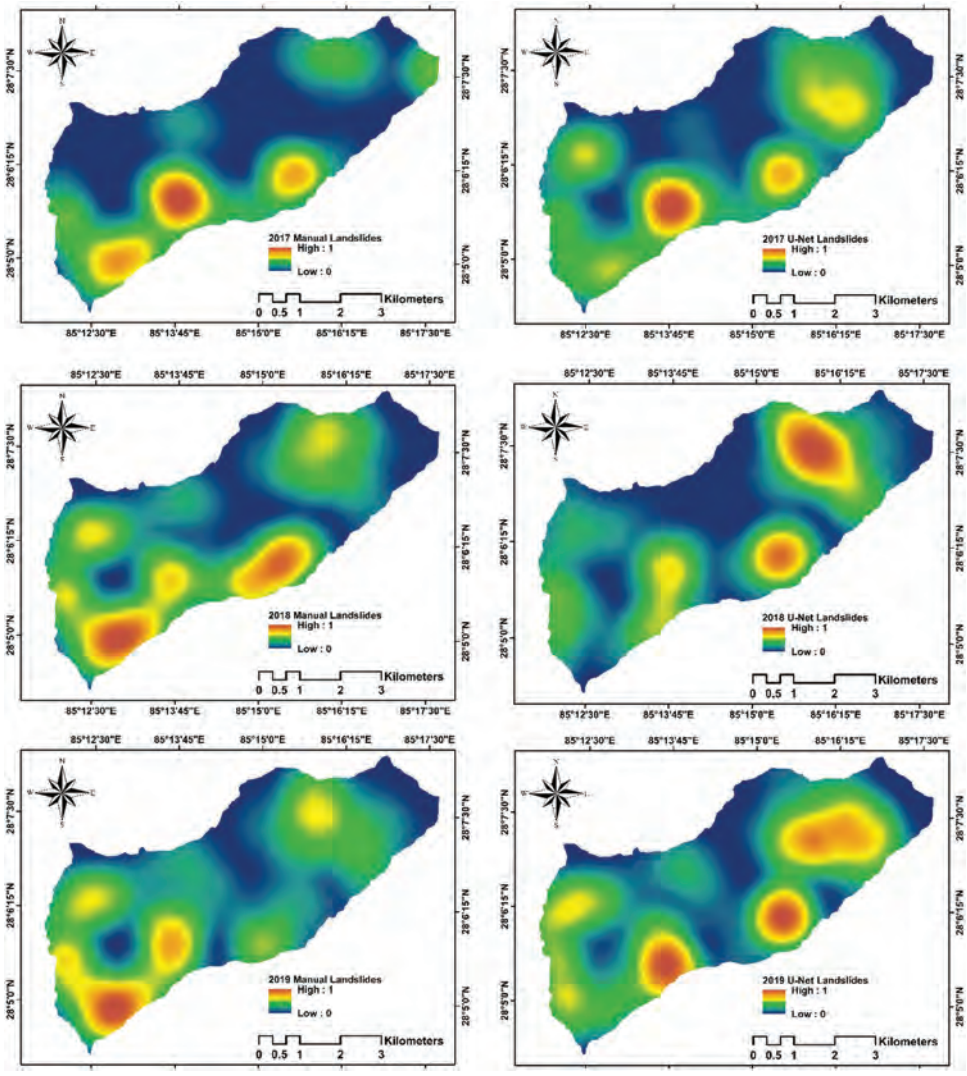

Figure 7. (Continued).

Table 5. Comparison of landslide density for the manually annotated landslide inventories versus the predicted landslide inventories.

Years	Density of the manual landslide inventory MI (Nr/km^2)	Density of the predicted landslide inventory PI (Nr/km^2)	Difference (PI - MI) (Nr/km^2)
2013	3.22	2.59	−0.63
2014	3.62	6.64	3.02
2015	16.46	11.47	−4.99
2016	8.37	11.54	3.17
2017	6.16	7.46	1.30
2018	4.85	7.35	2.50
2019	4.70	6.99	2.25

5. Discussion

5.1. Deep learning as a means of generating multi-temporal inventories

The mapping accuracy in the study area for each year is essential for analyzing the temporal changes of the landslide geometries in the later stages. We noticed that due to fragmentation issues while predicting landslides inhibits the true delineation of a landslide body, therefore detecting "more" landslide bodies in absolute numbers, and under-representing the true area of the modeled landslides. This also led to an increase in the LD in the PIs than the MIs as discussed in section 4.2. The region also suffered from shadows

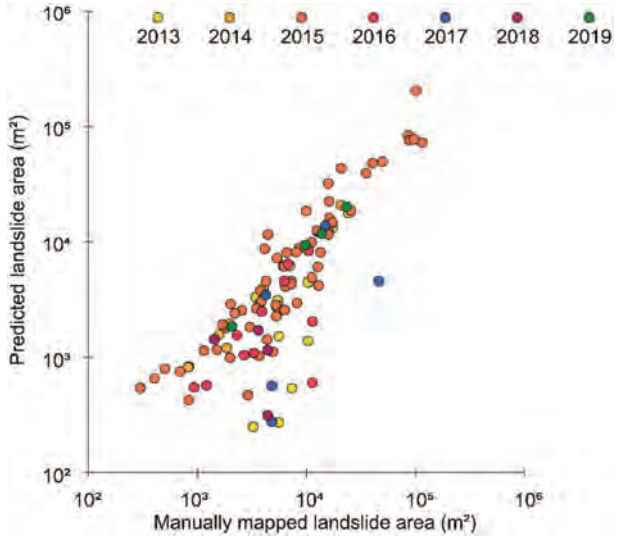

Figure 8. Difference in area between the landslides of MIs versus the PIs. The Pearson correlation between the PI and MI in this graph is 0.985.

and atmospheric noise whereby in some locations, the model simply does not predict the landslides. While the model struggles making the predictions in such conditions, overall, the model performed quite well when it came to detecting the landslide accurately in each year. The use of transfer learning abetted the performance capability of the model to effectively retain and compound on the existing knowledge of landslides from prior training instances and capitalized on detecting the landslides in the pre-seismic and post-seismic years. A typical issue in the successful implementation of a data-driven model for landslide mapping tasks is the shortage of training data. Although the use of 55 training samples can be usually judged as very small for effective training of a deep learning model, it has shown itself effective in our study by applying data augmentation techniques to expand the amount of training samples, thereby allowing the model to generalize better in predicting landslides temporally. A major challenge faced has been the spectral reflectance differences between each image acquisition for each year. Results show that for each year, the overall F1-scores are different because of these variations in the spectral reflectance for each year, however, in general, the landslide footprints are mapped very well for each year. The results also show that using training samples of only 55 landslides from a geographically separate area is enough to detect landslides temporally in interested regions, and also to get more than adequate accuracies to generate MTL inventories. Various gaps and constraints remain despite the fact that this is the first study to attempt in MTL mapping of landslides. Among them: i) the choice of seasonality has to be investigated further, as it is responsible for significant spectral changes in imagery, mostly linked to vegetation.; ii) determining landslide footprints of each respective year while avoiding double counting. Moreover, when compared to studies like Meena and Piralilou (2019), we see that they had compared inventories mapped by four different mappers for the same region and the differences between the manually generated inventories were quite significant. At the same time, taking examples from Figure 8 and instances of false positive predictions that actually turned out to be landslides originally missed by the manual mappers due to drafting errors, shows the strength of transfer learning, at least from the perspective of rapid automated methods.

5.2. *Factors influencing model predictions*

The potential sources of uncertainties in this study can be attributed to 1) the lack of training data prevents the model from accurately predicting the whole landslide body, 2) artifacts in satellite images such as haze, terrain and cloud shadows, and cloud obscuration, 3) the subjectivity and ambiguity in manual mapping of the landslides 4) might result in simply identifying "visible scars" in the terrain as landslides and not taking into consideration, for example, the crown of the landslide that would otherwise be hidden by vegetation cover. Unfortunately, these uncertainties cannot be dealt by the model, however, can easily be addressed by manually adjusting and fixing such errors. The point of using a deep learning or any AI model is to enable quick and accurate mapping of landslides where it is very time consuming to manually delineate the landslides. But posterior human intervention could further improve the automatically generated MTL inventories by simple edits and adjustments.

While the use of the U-Net model in this case works quite well, much room is still left to further improve the capacity of artificial intelligence models to detect landslides in different years. With newer and advanced variants of the U-Net model being used in the landslide detection domain, the use of more advanced models and training schemes can surely

improve results in the future. Furthermore, the next gradual step would be to enhance on the area of investigation and attempt to generate MTL inventories at larger scales and across different geographies. Furthermore, considering the availability of elevation models from missions like SRTM, ALOS-PALSAR, and airborne laser scanners, elevation data can also be used in the future coupled with MT satellite images to further expand and improve the detection results (removing the FPs), particularly in regions which are regularly monitored/surveyed (Kim, Sohn, and Kim 2020).

6. Conclusions

Mapping landslides automatically is a difficult task and much research has been conducted that shows how well DL models can be used to map landslides efficiently and rapidly. But this mapping endeavor through DL models is thus far only explored spatially, and not temporally. We propose the first multi-temporal landslide inventory mapping effort with the U-Net DL model to automatically detect and map landslides over time by using medium resolution RapidEye images of the Nepal Mailung area for the years between 2013 and 2019. The U-Net model is first trained separately outside the investigation area to test model effectiveness in a geographically distinct area and then the weights learnt from this first training phase are utilized to map landslides over time within the investigation area using transfer learning. The model's performance is assessed using classical metrics on the test set, as well as differences in spatial distribution and landslide statistics between the manual and modeled inventories. The significance of the study is the outlook toward automatically generating MTL inventories for locations that has had quite an interesting episodic behavior of past landslide activities which could allow to understand the long-term effects of earthquake events, landscape instability and reactivation of landslides post heavy rainfall, and recovery of the landscape.

Our next focus will be on attempting to detect and map multi-temporal landslides over different topographic regions in order to test how well such models perform in terms of their generalization capability. Moreover, different models will also be experimented to utilize more advanced networks and layers at improving the mapping of landslide footprints.

Highlights

- Releasing source codes, curated data, and predicted multi-temporal landslides inventories.
- Use of remote sensing imagery to generate multi-temporal landslides inventories.
- First artificial intelligence model to map landslides over time.
- Mapping of pre-, co-, and post-seismic landslides of the Gorkha earthquake event of 2015.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research received no external funding.

ORCID

Kushanav Bhuyan http://orcid.org/0000-0002-6173-8696
Sansar Raj Meena http://orcid.org/0000-0001-6175-6491
Lorenzo Nava http://orcid.org/0000-0002-2327-8721
Cees van Westen http://orcid.org/0000-0002-2992-902X
Mario Floris http://orcid.org/0000-0002-8691-2179
Filippo Catani http://orcid.org/0000-0001-5185-4725

Data availability statement

We present the data and codes openly available at https://github.com/kushanavbhuyan/Multi-Temporal-Landslide-Mapping-Nepal to encourage reproducibility of the study. We include a Jupyter Notebook script and the trained model weights, making it straightforward for interested academics to design and test MT landslide inventories in new regions of interest.

Author agreement

As the corresponding Author, I certify that this manuscript is unique, has never been published, and is not under consideration for publication anywhere else. I certify that all listed authors have read and approved the work. I also confirm that we have all given our permission to the authorship order shown in the text.

Author contributions

Conceptualization, K.B., and S.R.M.; data curation, S.R.M, methodology, K.B; analysis and interpretation of results, K.B., S.R.M., and L.N.;; writing – original draft preparation, K.B., and L.N.; writing, review and editing, S.R.M., K.B., L.N., C.V.W., M.F., and F.

C.; supervision, C.V.W., M.F., and F.C.; funding acquisition, S.R.M. The published version of the work has been reviewed and approved by all authors.

References

Abraham, N., and N. M. Khan. 2019. A Novel Focal Tversky Loss Function with Improved Attention U-Net for Lesion Segmentation. *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, April, 683–687. IEEE. 10.1109/ISBI.2019.8759329

Althuwaynee, O. F., B. Pradhan, and S. Lee. 2012. "Application of an Evidential Belief Function Model in Landslide Susceptibility Mapping." *Computers & Geosciences* 44: 120–135. doi:10.1016/j.cageo.2012.03.003.

Anderson, J. E., and J. E. Perry. 1996. "Characterization of Wetland Plant Stress Using Leaf Spectral Reflectance: Implications for Wetland Remote Sensing." *Wetlands* 16 (4): 477–487. doi:10.1007/BF03161337.

Ayana, G., K. Dese, and S. W. Choe. 2021 10 February. "Transfer Learning in Breast Cancer Diagnoses via Ultrasound Imaging." *Cancers* 134: 1–16. Multidisciplinary Digital Publishing Institute. doi:10.3390/cancers13040738.

Azarafza, M., M. Azarafza, H. Akgün, P. M. Atkinson, and R. Derakhshani. 2021. "Deep Learning-Based Landslide Susceptibility Mapping." *Scientific Reports* 11 (1): 1–16. doi:10.1038/s41598-021-03585-1.

Bai, S., J. Wang, Z. Zhang, and C. Cheng. 2012. "Combined Landslide Susceptibility Mapping After Wenchuan Earthquake at the Zhouqu Segment in the Bailongjiang Basin, China." *CATENA* 99: 18–25. doi:10.1016/J.CATENA.2012.06.012.

Bottou, L. 2010. Large-Scale Machine Learning with Stochastic Gradient Descent. In *Proceedings of COMPSTAT 2010 - 19th International Conference on Computational Statistics, Keynote, Invited and Contributed Papers* (pp. 177–186). Physica-Verlag HD. 10.1007/978-3-7908-2604-3_16

Catani, F., N. Casagli, L. Ermini, G. Righini, and G. Menduni. 2005. "Landslide Hazard and Risk Mapping at Catchment Scale in the Arno River Basin." *Landslides* 2 (4): 329–342. doi:10.1007/s10346-005-0021-0.

Challa, H., N. Niu, and R. Johnson .2020. Faulty Requirements Made Valuable: On the Role of Data Quality in Deep Learning. *Proceedings - 7th International Workshop on Artificial Intelligence and Requirements Engineering, AIRE 2020*, 61–69. Institute of Electrical and Electronics Engineers Inc. 10.1109/AIRE51212.2020.00016

Chen, T. H. K., A. V. Prishchepov, R. Fensholt, and C. E. Sabel. 2019. "Detecting and Monitoring Long-Term Landslides in Urbanized Areas with Nighttime Light Data and Multi-Seasonal Landsat Imagery Across Taiwan from 1998 to 2017." *Remote Sensing of Environment* 225: 317–327. doi:10.1016/J.RSE.2019.03.013.

Dahal, A., H. Tanyas, C. van Westen, M. van der Meijde, P. M. Mai, R. Huser, and L. Lombardo. 2022. "Space-Time Landslide Hazard Modeling via Ensemble Neural Networks." Retrieved from. doi:10.31223/x5b075.

Fan, X., G. Domènech, G. Scaringi, R. Huang, Q. Xu, T. C. Hales, L. Dai, Q. Yang, O. Francis. 2018. "Spatio-Temporal Evolution of Mass Wasting After the 2008 Mw 7.9 Wenchuan Earthquake Revealed by a Detailed Multi-Temporal Inventory." *Landslides* 15 (12): 2325–2341. doi:10.1007/s10346-018-1054-5.

Fang, B., G. Chen, L. Pan, R. Kou, and L. Wang. 2021. "GAN-Based Siamese Framework for Landslide Inventory Mapping Using Bi-Temporal Optical Remote Sensing Images." *IEEE Geoscience and Remote Sensing Letters* 18 (3): 391–395. doi:10.1109/LGRS.2020.2979693.

Fan, X., G. Scaringi, G. Domènech, F. Yang, X. Guo, L. Dai, C. He, Q. Xu, R. Huang. 2019 9 January. "Two Multi-Temporal Datasets That Track the Enhanced Landsliding After the 2008 Wenchuan Earthquake." *Earth System Science Data* 111: 35–55. Copernicus GmbH. doi:10.5194/essd-11-35-2019.

Fan, X., Q. Xu, C. J. van Westen, R. Huang, and R. Tang. 2017. "Characteristics and Classification of Landslide Dams Associated with the 2008 Wenchuan Earthquake." *Geoenvironmental Disasters* 4 (1): 1–15. doi:10.1186/s40677-017-0079-8.

Fan, X., A. P. Yunus, G. Scaringi, F. Catani, S. Siva Subramanian, Q. Xu, and R. Huang. 2021 16 January. "Rapidly Evolving Controls of Landslides After a Strong Earthquake and Implications for Hazard Assessments." *Geophysical Research Letters* 481: e2020GL090509. John Wiley & Sons, Ltd. doi:10.1029/2020GL090509.

Ghorbanzadeh, O., T. Blaschke, K. Gholamnia, S. R. Meena, D. Tiede, and J. Aryal. 2019. "Evaluation of Different Machine Learning Methods and Deep-Learning Convolutional Neural Networks for Landslide Detection." *Remote Sensing* 11 (2): 196. doi:10.3390/rs11020196.

Ghorbanzadeh, O., A. Crivellari, P. Ghamisi, H. Shahabi, and T. Blaschke. 2021. "A Comprehensive Transferability Evaluation of U-Net and ResU-Net for Landslide Detection from Sentinel-2 Data (Case Study Areas from Taiwan, China, and Japan)." *Scientific Reports* 11 (1). doi:10.1038/s41598-021-94190-9.

Ghorbanzadeh, O., S. R. Meena, H. Shahabi Sorman Abadi, S. Tavakkoli Piralilou, L. Zhiyong, and T. Blaschke. 2021. "Landslide Mapping Using Two Main Deep-Learning Convolution Neural Network Streams Combined by the Dempster-Shafer Model." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 14: 452–463. doi:10.1109/JSTARS.2020.3043836.

Guzzetti, F., A. C. Mondini, M. Cardinali, F. Fiorucci, M. Santangelo, and K. T. Chang. 2012. "Landslide Inventory Maps: New Tools for an Old Problem." *Earth-Science Reviews* 112 (1–2): 42–66. Elsevier. doi:10.1016/j.earscirev.2012.02.001.

Guzzetti, F., P. Reichenbach, M. Cardinali, M. Galli, and F. Ardizzone. 2005. "Probabilistic Landslide Hazard Assessment at the Basin Scale." *Geomorphology* 72 (1–4): 272–299. doi:10.1016/j.geomorph.2005.06.002.

Hervás, J., and P. Bobrowsky. 2009. "Mapping: Inventories, Susceptibility, Hazard and Risk." *Landslides - Disaster Risk Reduction* 321–349. doi:10.1007/978-3-540-69970-5_19/COVER.

Huete, A. R. 2004. "REMOTE SENSING FOR ENVIRONMENTAL MONITORING." *Environmental Monitoring and Characterization* 183–206. doi:10.1016/B978-012064477-3/50013-8.

Jaiswal, P., and C. J. van Westen. 2009. "Estimating Temporal Probability for Landslide Initiation Along Transportation Routes Based on Rainfall Thresholds." *Geomorphology* 112 (1–2): 96–105. doi:10.1016/j.geomorph.2009.05.008.

Jones, J. N., S. J. Boulton, M. Stokes, G. L. Bennett, and M. R. Z. Whitworth. 2021. "30-Year Record of Himalaya Mass-Wasting Reveals Landscape Perturbations by Extreme Events." *Nature Communications* 12 (1): 1–15. doi:10.1038/s41467-021-26964-8.

Kargel, J. S., G. J. Leonard, D. H. Shugar, U. K. Haritashya, A. Bevington, E. J. Fielding …, and N. Young, M. Geertsema, E. S. Miles, J. Steiner, E. Anderson. 2016. "Geomorphic and Geologic Controls of Geohazards Induced by Nepal's 2015 Gorkha Earthquake." *Science* 351 (6269). doi:10.1126/SCIENCE.AAC8353/SUPPL_FILE/KARGEL-SM.PDF.

Kim, M. K., H. G. Sohn, and S. Kim. 2020. "Incorporating the Effect of ALS-Derived DEM Uncertainty for Quantifying Changes Due to the Landslide in 2011." *Mt Umyeon, Seoul GIScience and Remote Sensing* 57 (3): 287–301. doi:10.1080/15481603.2019.1687133.

Korup, O., T. Görüm, and Y. Hayakawa. 2012. "Without Power? Landslide Inventories in the Face of Climate Change." *Earth Surface Processes and Landforms* 37 (1): 92–99. doi:10.1002/ESP.2248.

Kukačka, J., V. Golkov, and D. Cremers. 2017. *Regularization for Deep Learning: A Taxonomy*. Retrieved from https://arxiv.org/abs/1710.10686v1

Martha, T. R., P. Roy, R. Mazumdar, K. B. Govindharaj, and K. V. Kumar. 2016. "Spatial Characteristics of Landslides Triggered by the 2015 Mw 7.8 (Gorkha) and Mw 7.3 (Dolakha) Earthquakes in Nepal." *Landslides* 14 (2): 697–704. doi:10.1007/S10346-016-0763-X.

Meena, S. R., O. Ghorbanzadeh, C. J. van Westen, T. G. Nachappa, T. Blaschke, R. P. Singh, and R. Sarkar. 2021. "Rapid Mapping of Landslides in the Western Ghats (India) Triggered by 2018 Extreme Monsoon Rainfall Using a Deep Learning Approach." *Landslides* 18 (5): 1937–1950. doi:10.1007/s10346-020-01602-4.

Meena, S. R., and S. T. Piralilou. 2019. "Comparison of Earthquake-Triggered Landslide Inventories: A Case Study of the 2015 Gorkha Earthquake, Nepal." *Geosciences (Switzerland)* 9 (10): 437. doi:10.3390/geosciences9100437.

Meena, S. R., L. P. Soares, C. H. Grohmann, C. van Westen, K. Bhuyan, R. P. Singh, and F. Catani, F. Catani. 2022. "Landslide Detection in the Himalayas Using Machine Learning Algorithms and U-Net." *Landslides* 19 (5): 1209–1229. doi:10.1007/s10346-022-01861-3.

Metternicht, G., L. Hurni, and R. Gogu. 2005. "Remote Sensing of Landslides: An Analysis of the Potential Contribution to Geo-Spatial Systems for Hazard Assessment in Mountainous Environments." *Remote Sensing of Environment* 98 (2–3): 284–303.

Nava, L., K. Bhuyan, S. R. Meena, O. Monserrat, and F. Catani. 2022. "Rapid Mapping of Landslides on SAR Data by Attention U-Net." *Remote Sensing* 14 (6): 1449. doi:10.3390/rs14061449.

Ozturk, U., E. Bozzolan, E. A. Holcombe, R. Shukla, F. Pianosi, and T. Wagener. 2022. "How Climate Change and Unplanned Urban Sprawl Bring More Landslides." *Nature* 608 (7922): 262–265. doi:10.1038/d41586-022-02141-9.

Pan, Z., J. Xu, Y. Guo, Y. Hu, and G. Wang. 2020. "Deep Learning Segmentation and Classification for Urban Village Using a Worldview Satellite Image Based on U-Net." *Remote Sensing* 12 (10): 1574. doi:10.3390/rs12101574.

Pan, S. J., and Q. Yang. 2010. "A Survey on Transfer Learning." *IEEE Transactions on Knowledge and Data Engineering* 22 (10): 1345–1359. doi:10.1109/TKDE.2009.191.

Planet Team. (2017). *Planet Application Program Interface: In Space for Life on Earth*. https://api.planet.com

Pokharel, B., M. Alvioli, and S. Lim. 2021. "Assessment of Earthquake-Induced Landslide Inventories and Susceptibility Maps Using Slope Unit-Based Logistic Regression and Geospatial Statistics." *Scientific Reports* 11 (1): 1–15. doi:10.1038/s41598-021-00780-y.

Prakash, N., A. Manconi, and S. Loew. 2020. "Mapping Landslides on EO Data: Performance of Deep Learning Models Vs. Traditional Machine Learning Models." *Remote Sensing* 12 (3): 346. doi:10.3390/rs12030346.

Prakash, N., A. Manconi, and S. Loew. 2021. "A New Strategy to Map Landslides with a Generalized Convolutional Neural Network." *Scientific Reports* 11 (1): 1–15. doi:10.1038/s41598-021-89015-8.

Rana, K., N. Malik, and U. Ozturk. 2022. "Landsifier V1.0: A Python Library to Estimate Likely Triggers of Mapped Landslides." *Natural Hazards and Earth System Sciences* 22 (11): 3751–3764. doi:10.5194/NHESS-22-3751-2022.

Rana, K., U. Ozturk, and N. Malik. 2021. "Landslide Geometry Reveals Its Trigger." *Geophysical Research Letters* 48 (4). doi:10.1029/2020GL090848.

Roback, K., M. K. Clark, A. J. West, D. Zekkos, G. Li, S. F. Gallen, D. Chamlagain, J. W. Godt. 2018. "The Size, Distribution, and Mobility of Landslides Caused by the 2015 Mw7.8 Gorkha Earthquake, Nepal." *Geomorphology* 301: 121–138. doi:10.1016/J.GEOMORPH.2017.01.030.

Ronneberger, O., P. Fischer, and T. Brox. 2015. "U-Net: Convolutional Networks for Biomedical Image Segmentation." In *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 9351*, 234–241. Springer Verlag. doi:10.1007/978-3-319-24574-4_28.

Rosser, N., M. Kincey, K. Oven, A. Densmore, T. Robinson, D. S. Pujara, and M. R. Dhital, J. Smutny, K. Gurung, S. Lama, M. R. Dhital. 2021. "Changing Significance of Landslide Hazard and Risk After the 2015 Mw 7.8 Gorkha,

Nepal Earthquake." *Progress in Disaster Science* 10: 100159. doi:10.1016/j.pdisas.2021.100159.

Schwanghart, W., M. Ryan, and O. Korup. 2018. "Topographic and Seismic Constraints on the Vulnerability of Himalayan Hydropower." *Geophysical Research Letters* 45 (17): 8985–8992. doi:10.1029/2018GL079173.

Serey, A., L. Piñero-Feliciangeli, S. A. Sepúlveda, F. Poblete, D. N. Petley, and W. Murphy. 2019. "Landslides Induced by the 2010 Chile Megathrust Earthquake: A Comprehensive Inventory and Correlations with Geological and Seismic Factors." *Landslides* 16 (6): 1153–1165. doi:10.1007/S10346-019-01150-6.

Shorten, C., and T. M. Khoshgoftaar. 2019. "A Survey on Image Data Augmentation for Deep Learning." *Journal of Big Data* 6 (1): 1–48. doi:10.1186/s40537-019-0197-0.

Soeters, R., and C. J. Van Westen. 1996. "Slope Instability Recognition, Analysis and Zonation." *Landslides: Investigation and Mitigation* 247: 129–177.

Spinetti, C., M. Bisson, C. Tolomei, L. Colini, A. Galvani, M. Moro, and M. Saroli, V. Sepe. 2019. "Landslide Susceptibility Mapping by Remote Sensing and Geomorphological Data: Case Studies on the Sorrentina Peninsula (Southern Italy)." *GIScience & Remote Sensing* 56 (6): 940–965. doi:10.1080/15481603.2019.1587891.

Sreedevi, N., and K. Yarrakula. 2016. "Different Techniques for Landslide Zonation Mapping and Landslide Assessment: A Review." *Indian Journal of Science and Technology* 9: 47.

Sun, Y., A. K. C. Wong, and M. S. Kamel. 2009. "Classification of Imbalanced Data: A Review." *International Journal of Pattern Recognition and Artificial Intelligence* 23 (4): 687–719. doi:10.1142/S0218001409007326.

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

Tang, C., C. J. Van Westen, H. Tanyas, and V. G. Jetten. 2016. "Analysing Post-Earthquake Landslide Activity Using Multi-Temporal Landslide Inventories Near the Epicentral Area of the 2008 Wenchuan Earthquake." *Natural Hazards and Earth System Sciences* 16 (12): 2641–2655. doi:10.5194/nhess-16-2641-2016.

Tanyaş, H., D. Kirschbaum, T. Görüm, C. J. van Westen, and L. Lombardo. 2021. "New Insight into Post-Seismic Landslide Evolution Processes in the Tropics." *Frontiers in Earth Science* 9: 551. doi:10.3389/feart.2021.700546.

Tanyaş, H., C. J. van Westen, K. E. Allstadt, M. Anna Nowicki Jessee, T. Görüm, R. W. Jibson, J. W. Godt, H. P. Sato, R. G. Schmitt, O. Marc, N. Hovius. 2017. "Presentation and Analysis of a Worldwide Database of Earthquake-Induced Landslide Inventories." *Journal of Geophysical Research: Earth Surface* 122 (10): 1991–2015. doi:10.1002/2017JF004236.

USGS. (2015). M 7.8 - 67 Km NNE of Bharatpur, Nepal. Retrieved August 4, 2022: from United States Geological Survey website https://earthquake.usgs.gov/earthquakes/eventpage/us20002926/executive

Van Westen, C. J., S. Ghosh, P. Jaiswal, T. R. Martha, and S. L. Kuriakose. 2013. "From Landslide Inventories to Landslide Risk Assessment; an Attempt to Support Methodological Development in India." *Landslide Science and Practice: Landslide Inventory and Susceptibility and Hazard Zoning* 1: 3–20. doi:10.1007/978-3-642-31325-7_1.

Wang, F., X. Fan, A. P. Yunus, S. Siva Subramanian, A. Alonso-Rodriguez, L. Dai, Q. Xu, R. Huang. 2019. "Coseismic Landslides Triggered by the 2018 Hokkaido, Japan (Mw 6.6), Earthquake: Spatial Distribution, Controlling Factors, and Possible Failure Mechanism." *Landslides* 16 (8): 1551–1566. doi:10.1007/S10346-019-01187-7.

Wang, H., L. Wang, and L. Zhang. 2022. "Transfer Learning Improves Landslide Susceptibility Assessment." *Gondwana Research*. doi:10.1016/j.gr.2022.07.008.

Xu, C., S. Xu, F. Dai, Z. Wu, H. He, F. Shi, and X. Xu, S. Xu. 2013. "Application of an Incomplete Landslide Inventory, Logistic Regression Model and Its Validation for Landslide Susceptibility Mapping Related to the May 12, 2008 Wenchuan Earthquake of China." *Natural Hazards* 68 (2): 883–900. doi:10.1007/S11069-013-0661-7/FIGURES/7.

Zhang, X., M. O. Pun, and M. Liu. 2021. "Semi-supervised Multi-temporal Deep Representation Fusion Network for Landslide Mapping from Aerial Orthophotos." *Remote Sensing* 13 (4): 1–22. doi:10.3390/rs13040548.