Landslide detection in the Himalayas using machine learning algorithms and U-Net

8 authors, including:

Sansar Raj Meena
University of Padova
32 PUBLICATIONS **573** CITATIONS

Lucas Soares
University of São Paulo
4 PUBLICATIONS **0** CITATIONS

C.J. Van Westen
University of Twente, Faculty of Geo-Information Science and Earth Observation (…
353 PUBLICATIONS **15,580** CITATIONS

Kushanav Bhuyan
University of Padova
7 PUBLICATIONS **14** CITATIONS

Some of the authors of this publication are also working on these related projects:

Project OPERA Civil protection from floods View project

Project CHANGES View project

Sansar Raj Meena · **Lucas Pedrosa Soares** · **Carlos H. Grohmann** ·
Cees van Westen · **Kushanav Bhuyan** · **Ramesh P. Singh** · **Mario Floris** ·
Filippo Catani

Landslide detection in the Himalayas using machine learning algorithms and U-Net

Abstract Event-based landslide inventories are essential sources to broaden our understanding of the causal relationship between triggering events and the occurring landslides. Moreover, detailed inventories are crucial for the succeeding phases of landslide risk studies like susceptibility and hazard assessment. The openly available inventories differ in the quality and completeness levels. Event-based landslide inventories are created based on manual interpretation, and there can be significant differences in the mapping preferences among interpreters. To address this issue, we used two different datasets to analyze the potential of U-Net and machine learning approaches for automated landslide detection in the Himalayas. Dataset-1 is composed of five optical bands from the RapidEye satellite imagery. Dataset-2 is composed of the RapidEye optical data, and ALOS-PALSAR derived topographical data. We used a small dataset consisting of 239 samples acquired from several training zones and one testing zone to evaluate our models' performance using the fully convolutional U-Net model, Support Vector Machines (SVM), K-Nearest Neighbor, and the Random Forest (RF). We created thirty-two different maps to evaluate and understand the implications of different sample patch sizes and their effect on the accuracy of landslide detection in the study area. The results were then compared against the manually interpreted inventory compiled using fieldwork and visual interpretation of the RapidEye satellite image. We used accuracy assessment metrics such as F1-score, Precision, Recall, and Mathews Correlation Coefficient (MCC). In the context of the Nepali Himalayas, employing RapidEye images and machine learning models, a viable patch size was investigated. The U-Net model trained with 128×128 pixel patch size yields the best MCC results (76.59%) with the dataset-1. The added information from the digital elevation model benefited the overall detection of landslides. However, it does not improve the model's overall accuracy but helps differentiate human settlement areas and river sand bars. In this study, the U-Net achieved slightly better results than other machine learning approaches. Although it can depend on architecture of the U-Net model and the complexity of the geographical features in the imagery, the U-Net model is still preliminary in the domain of landslide detection. There is very little literature available related to the use of U-Net for landslide detection. This study is one of the first efforts of using U-Net for landslide detection in the Himalayas. Nevertheless, U-Net has the potential to improve further automated landslide detection in the future for varied topographical and geomorphological scenes.

Keywords Landslides · U-Net · Deep learning · Machine learning · Himalayas

Introduction

Loss of property and human life due to earthquake-triggered landslides is significantly high and is expected to increase due to climate change (Froude and Petley 2016; Gariano and Guzzetti 2016). About 47,000 earthquake-induced landslide casualties were reported from 2004 to 2010 (Petley 2012). Earthquake-induced landslides (EQIL) have direct and indirect long-term socioeconomic and environmental effects (Fan et al. 2018). The direct and indirect effects of landslides, for example, through the formation and breakout of landslide dams, are a significant natural hazard in the mountain regions of the Himalayas (Dhital 2015). Studies show unprecedented loss to both human lives and the economy in the Himalayan regions due to landslides, contributing up to 30% of the world's total landslide-related damage value (Dahal and Hasegawa 2008; Haigh and Rawat 2011). In Northern India, for example, during the recent 2021 Uttarakhand landslides, 24 people were killed by landslides and around 150 were missing (Meena et al. 2021a). A large number of people are affected in the Himalayan regions by small and large-scale landslides, especially during the monsoon seasons (Khanal and Watanabe 2005; Thapa and Dhital 2000; Upreti and Dhital 1996). Although landslides often occur in remote areas, the resulting catastrophic flash floods from landslide dam outbreak cause extensive damage to settlements, hydroelectric projects, and agriculture fields in the downstream areas (Meena and Tavakkoli Piralilou 2019).

To better analyze the frequency and distribution of landslides, there is a growing demand for event-based inventories that can be used to determine the probability of landslide occurrence in space and time as a basis for hazard and risk assessment. There is still insufficient information on landslide occurrences for many areas to make reliable hazard maps (Reichenbach et al. 2018). Landslide susceptibility and hazard modeling require accurate and complete landslide inventory datasets. This inventory dataset is usually used for training hazard models to find potential landslide-prone areas (Guzzetti et al. 2012).

The accuracy and completeness of landslide inventory datasets are essential for making spatial predictions for future events (Hakan and Luigi 2020). The mapping of event-based landslide inventories in remote and mountainous areas makes remote sensing data the primary source of information for mapping these events (Chen et al. 2018).

In terms of detecting landslide boundaries with remote sensing images, classification methods like pixel-based, feature-based, and object-based techniques can be employed (Lu et al. 2020; Su et al.

2020). While pixel-based methods only extract features by classifying each pixel, they do not take the spatial-context into account. However, feature-based methods (like gray level co-occurrence matrix and principal component analysis) (Whitworth et al. 2002) and object-based image analysis (OBIA) explicitly leverage the spatial information from satellite images (Bacha et al. 2020; Hölbling et al. 2012; Martha et al. 2010). During the last decade, deep-learning models and other machine learning models, particularly Convolutional Neural Networks (CNNs), have been applied successfully in a broad range of image segmentation and object detection purposes (Ding et al. 2016; Ghorbanzadeh et al. 2020; Jin et al. 2019; Liu et al. 2019; Shi et al. 2020).

The use of CNN models has yielded promising results for classification of aerial images (Bui et al. 2019; Ghorbanzadeh et al. 2021, 2020; Meena et al. 2021b; Yu et al. 2017). Numerous studies using CNN have been conducted for landslide detection (see Table 1). Many authors used CNN models for automated landslide detection in mountainous regions using multi-temporal high-resolution remote sensing data, mono-temporal medium-resolution image data (Chen et al. 2018), where others optimized their models and compared with existing baseline models such as Fully Convolutional Networks (FCNs) (Lei et al. 2019). Hyperspectral data for landslide detection was first investigated by Ye et al. (2019). In recent studies, different topographical factors like elevation and its derivates like slope, aspect, and curvature combined with remote sensing data for landslide detection were explored to improve landslide detection (Sameen and Pradhan 2019; Liu et al. 2020b; Prakash et al. 2020).

Deep learning models usually require large amount of training data to detect objects efficiently. However, since landslide inventories are generated for small areas using manual interpretation and fieldwork, such inventories commonly have just a few samples and present a limitation for the training of deep learning models (Chen et al. 2020; Liu et al. 2020a; Qi et al. 2020) Therefore, in this study, the main objective was to evaluate and compare the performance of the machine and deep learning models trained with a small dataset composed of only 239 landslide polygons (55 polygons for training and 184 polygons for testing purposes). The fully convolutional U-Net deep learning model and other machine learning models were trained with data from a 5-m RapidEye optical satellite imagery and resampled 12.5-m ALOS PALSAR digital elevation data for landslide detection.

Study area

The study area is located in Rasuwa district Nepal, which is situated in higher Himalayas and is one of the highly affected regions after the 2015 Gorkha earthquake (see Fig. 1). Most of the study area falls in the Langtang national park and there are several hydropower plants projects along the Trishuli River. After the 2015 Gorkha earthquake, a series of landslides triggered by the earthquake caused damage to hydro powerplants, agricultural land, and human settlements. On 15 April 2015, during the Gorkha earthquake, more than 80 people were killed due to EQILs and flood events near the Mailung village hydropower plant camps. Several attempts have been made by local authorities and foreign institutes to study impact of landslide on human settlements and hydro powerplants in the region. However, in many inaccessible hilly areas, field visit

was not feasible hence remote sensing tools can help supplement the field visits. The study area is highly affected by monsoonal rains and every year several deep-seated landslides get reactivated such as the one near Ramche village.

Data used and methodology

Datasets

The landslides were visually interpreted as polygons from RapidEye imagery acquired on 04 November 2016 (Planet Labs Inc.) and field observations. The data has 5 m spatial resolution in five spectral bands: blue (440–510 nm), green (520–590 nm), red (630–685 nm), red-edge (690–730 nm), and near-infrared (760–850 nm) (RapidEye 2011).

A total of 239 landslide polygons were mapped in the entire study area, 55 in the training zones and 184 in the test zone (the training zones are yellow and testing zones are red in color in Fig. 1c). For training the model, 117 sampling points were manually selected along the centerline of the landslide polygons present in the training zone. Other 57 points were selected outside the landslide polygons to represent non-landslide samples (see Fig. 1c). Therefore, a total of 174 sampling points were used to train the models. Those points were used as the centroid to generate the training patches of four sizes: 16×16, 32×32, 64×64, and 128×128 pixels (Fig. 2).

Two datasets were created to train the models. Dataset-1 consists of the five spectral bands (RGB, red-edge, NIR) from the RapidEye satellite. The Dataset-2 consists of the same five bands and two extra topographical bands (elevation and slope). The elevation and slope data were acquired from a digital elevation model (DEM), resampled to 5-m spatial resolution, derived from Phased Array type L-band Synthetic Aperture Radar (PALSAR) of the Advanced Land Observing Satellite (ALOS).

All the models used the same training data to compare the results from the models properly. The deep learning algorithms were trained using the Python libraries TensorFlow 2.0 and the machine learning using Scikit-Learn.

Classifiers

U-Net model

U-Net (Ronneberger et al. 2015) is a state-of-art deep learning model used for semantic segmentation tasks. This model has an encoder-decoder architecture similar to the letter "U" (Fig. 3). The encoder path is composed of blocks of two 3×3 convolutional layers followed by a 2×2 max-pooling layer. The convolutional layers are 3×3 moving windows that translate around the image, calculating a dot product that can be summarized by Eq. 1 (Zhang et al. 2018):

$$O^l = \sigma(O^{l-1} * W^l + b^l) \tag{1}$$

where O^{l-1} refers to the output of the $(l-1)$th layer, W^l represents the weights and b^l represents the bias. σ indicates the non-linear activation function. The rectified linear unit (ReLU) was used as the activation function in this research. ReLU is commonly used as the

Table 1 Overview of some recently published studies on the automated mapping of landslides using deep learning approaches sorted according to the use of topographical features

Main objective	Algorithms used	Topographical feature used	Accuracy evaluation methods	Study
Optimization of (FCN-PP) for landslide inventory mapping	ELSE, RLSE, CDMRF, CDFFCM, CNN, FCN, U-Net, FCN-PP	N/A	Precision, recall, F1-score, OE, accuracy	(Lei et al. 2019)
Automatic mapping of landslides by the ResU-Net	U-Net, ResU-Net	N/A	Precision, recall, F1-score	(Qi et al. 2020)
Landslide recognition by deep convolutional neural network and change detection	CNN, CDCNN	DEM	Precision, recall, F1-score	(Shi et al. 2020)
To detect landslides of Nepal in the year 2015, at a national scale	ResNet-101 CNN	DEM	Precision, recall, F1-score	(Yu et al. 2020)
Landslide detection from an open satellite imagery and digital elevation model dataset using attention boosted convolutional neural networks	VGGi6, ResNet	DEM	Precision, recall, F1-score, Accuracy	(Ji et al. 2020)
An automated approach for landslide detection	DCNN	Slope	CE, DP, QP	(Chen et al. 2018)
Landslide detection using Hyperspectral remote sensing data and comparison of conventional methods with DLWC	DLWC, SVM, SID, SAM	Slope	Overall accuracy, Kappa coefficient, accuracy	(Ye et al. 2019)
Landslide detection using UAV–derived VHR imagery and topographical factors	CNN	Slope	PPV, TPR, F1-score, OPR, UPR, mIOU	(Ghorbanzadeh et al. 2019a)
Post-earthquake landslide extraction using the U-Net model	U-Net, U-Net + ResNet	DSM, slope, slope aspect	Precision, recall, F1-score, mIOU	(Liu et al. 2020b)
Comparison of different sample patch sizes for landslide detection using deep learning and machine learning	CNN, SVM, D-CNN, RF, ANN	Plan curvature, slope, slope aspect	Precision, recall, F1-score, mIOU	(Ghorbanzadeh et al. 2019b)
Landslide detection using residual networks	ResNet, CNN	Altitude, slope, slope aspect, total curvature	Training accuracy, validation accuracy, F1-score, mIOU	(Sameen and Pradhan 2019)
Comparison of pixel-based, object-based, and deep learning methods for landslide detection	RF, ANN, LR, U-Net + ResNet	Hill shade, slope, aspect. Terrain roughness, curvature, valley depth, TWI	Accuracy, F1-score, MCC, POD, POFD	(Prakash et al. 2020)

FCN-PP fully convolutional network within pyramid pooling, *ELSE* employed edge-based level set evolution, *RLSE* region-based level set evolution, *CDMRF* change detection-based on Markov random field, *CDFFCM* change detection-based fast FCM, *CNN* convolutional neural network, *FCN* fully convolutional network, *ResU-Net*, residual and U-Net, *VGGi6* very deep convolutional networks for large-scale image recognition, *UAV* unmanned aerial vehicle, *D-CNN* deep convolution neural network, *DLWC*, deep learning with constrains, *SVM* support vector machine, *SID* spectral information divergence, *SAM* spectral angle match, *RF* random forest, *ANN* artificial neural networks, *LR* logistic regression, *ResNet* residual networks
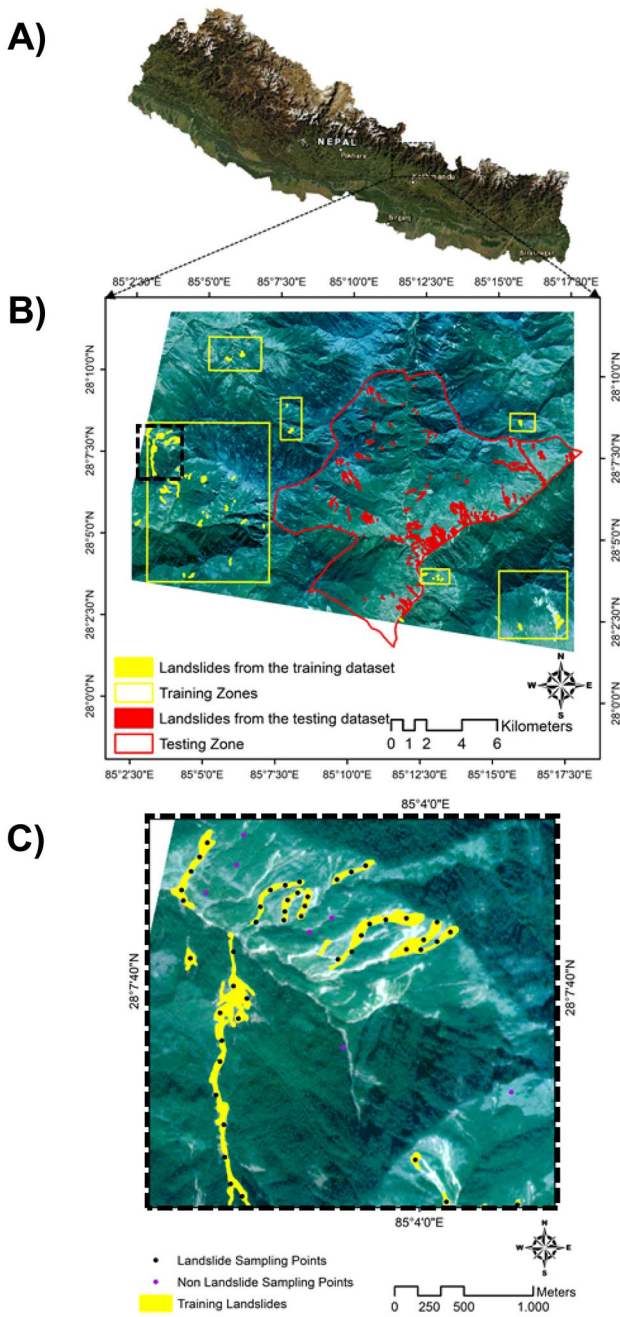
Fig. 1 **A** Location of the study area in Nepal, **B** landslide training and testing zones in the study area, and **C** sampling points along the center line of the landslide polygons (black) and non-landslide class (purple)

operation, the spatial dimension of the feature map is reduced to half of the input size.

The decoder path recovers the spatial location by using up-convolutions and concatenations from the encoder path (Ronneberger et al. 2015). The up-convolution layers increase the dimensions of the feature maps. The layers' output is concatenated with the feature map from the symmetrical position in the encoder path. In the last layer, a sigmoid function was used to output the class predictions in a 0–1 probability range. A threshold of 0.5 was used to determine the positive (> 0.5) and the negative (< 0.5) classes after the prediction.

Several papers describe and explain the U-Net structure and how convolutional neural networks are trained (Ghorbanzadeh et al. 2019a, b; Prakash et al. 2020; Wang et al. 2019). In this study, we use a fully convolutional neural network that is capable of calculating per-pixel probability of comprising a landslide. Unlike previous work conducted by Ghorbanzadeh et al. (2019a, b) where they used a classical convolutional neural network to generate patchwise landslide classification, the neural network used in our study is more efficient for landslide segmentation problems as the result is a binary output with the same size as the input image (Prakash et al. 2020, 2021; Qi et al. 2020). The network hyperparameter tuning process considered different number of filters (8, 16, 32), learning rates (0.01, 0.001, 0.0001), and batch sizes (8, 16, 32). The learning rate value was reduced by a factor of 0.1 when the validation loss function reaches a plateau for more than twenty epochs. The models were saved only when the validation loss function decreased as an attempt to avoid overfitting.
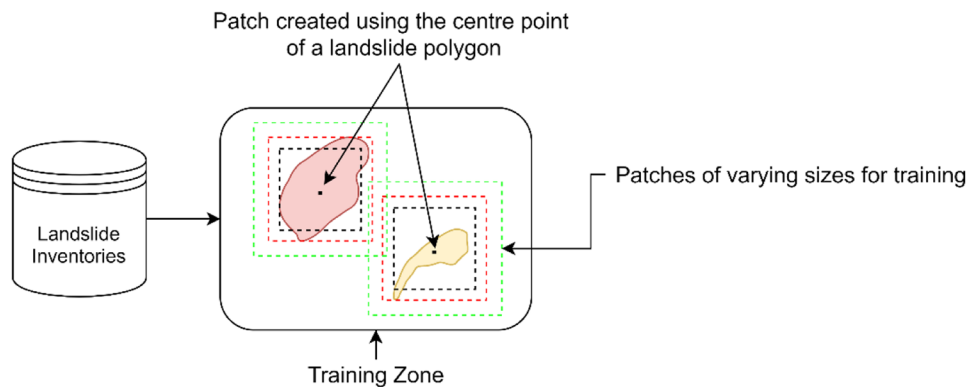
Support vector machine (SVM) model

SVM is a machine learning method that uses kernel functions to map the dataset into a higher dimension to determine a hyperplane that separates the training data feature spaces (Cortes and Vapnik 1995). The margining of the hyperplanes, also known as support vectors, is maximized to be the closest to the training features. This method gained popularity for landslide mapping due to accurate results, even with small datasets and unknown statistical distributions (Moosavi et al. 2014; Mountrakis et al. 2011; Pawłuszek and Borkowski 2016).

The classification result is affected by the kernel function (e.g., linear, sigmoid, polynomial, radial basis). Thus, various kernel functions were evaluated to find the best classifier.

K-nearest neighbors algorithm (KNN) model

K-nearest neighbors is a machine learning algorithm that uses the training data to find the feature space's K-closest neighbors. The algorithm outputs a class probability that reflects the uncertainty with which a given individual item can be assigned to any given class (Marjanovic et al. 2009). In this study, the distance between the feature space points was calculated using the Euclidean distance method. An optimal K value was determined by testing K in a 1–10 range.

activation function because it is more efficient than other functions and reduces the gradient vanishing problem during the training step (Wang et al. 2019). The function returns 0 when the input is negative and the same input value if it is positive. The max-pooling layers keep only the maximum values from the feature maps generated from the convolution operation. Thus, after a max-pooling

Fig. 2 Conceptualization of generating the patches for training the models

Random forests (RF) model

Random forest is an ensemble method widely used for landslide detection (Chen et al. 2014). The method is based on multiple decision trees. Each tree is slightly different since they are trained with the training dataset's random subsets. The technique is less prone to overfitting because each tree's output class is weighted based on a majority voting technique where the class with the most votes becomes the model's prediction.

Multiple input patches

Different patch sizes may affect the model accuracy because landslides have different shapes and sizes, which may not be
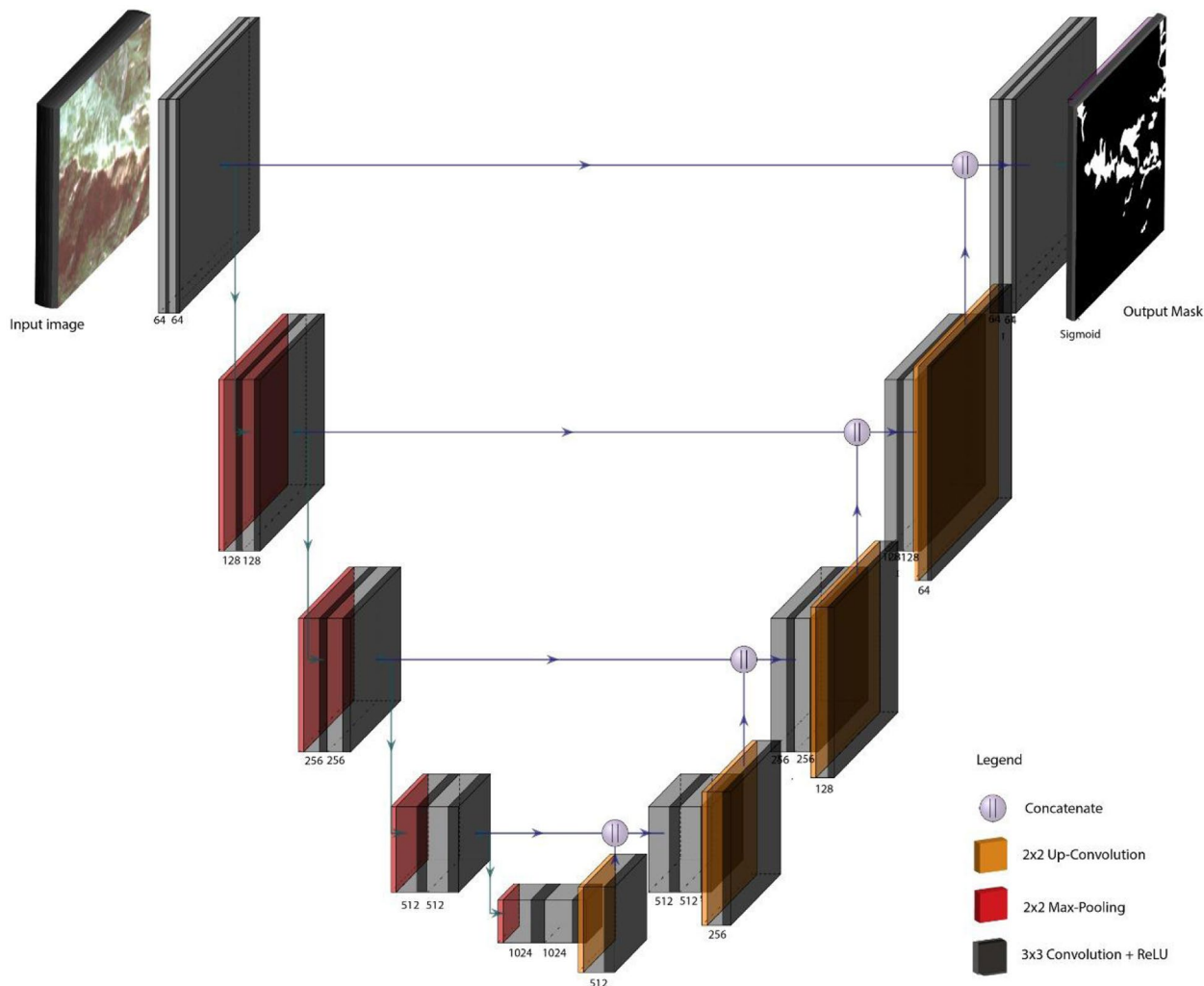
Fig. 3 The architecture of the U-Net model. The numbers below the convolution represent the number of filters used to train the model

well-represented depending on the patch size. Moreover, since the negative class is usually more frequent than the positive class in remote sensing imagery, larger patches may negatively influence the model because they can increase the imbalance between the positive and negative class (Ghorbanzadeh et al. 2019a, b).

In this work, the patches used to train the models were constituted by a multiple of 16 pixels since this is a condition to effectively train the U-Net model. The models were trained with 16 × 16, 32 × 32, 64 × 64, and 128 × 128 pixel patches to compare and evaluate how the different patch sizes affect the accuracy of the model. The models were also trained with 256 × 256 pixels patches. However, since the results were inferior compared to the other patch sizes, only the results achieved with the mentioned patch sizes were considered in the "Results" section.

Results

The machine and deep learning models were trained using only 174 samples to evaluate and compare the performance of the algorithms using small datasets. In total, sixteen result maps were generated for each dataset (dataset-1 and dataset-2). The result maps (Figs. 4a, b; 5a, b; 6a, b; 7a, b) are named based on the algorithm, the patch size, and the dataset used to train the algorithm. Therefore, the map U-Net_16_5 and U-Net_16_7 (Fig. 4a and b) correspond to the U-Net deep learning algorithm trained with the 16 × 16 patch size using the dataset with five optical bands (dataset-1) and seven bands (dataset-2), respectively. The best results were achieved by U-Net models with a learning rate of 0.001; SVM models trained with a polynomial kernel function and a scalable gamma parameter (γ); KNN models trained with nine neighbors; and RF models with 200 trees and depth 8.

Figure 8 portrays the differences in the areas of the landslides detected with the different machine learning models with respect to the influence of the topographical information from dataset-2. As seen in Fig. 8b, the total area in most of the models is relatively higher in dataset-2 than dataset-1 when compared against the manually interpreted ground truth area. This difference is because of the detection of false positives as an influence from the slope and elevation in dataset-2. Although there are improvements in the built-up area and river sand bars, the model gets confused and generates false positives in forests and agricultural areas.

The models were evaluated based on precision, recall, F1-score, and Matthews Correlation Coefficient (MCC) metrics, which are calculated using the value of true positive (TP), false positives (FP), and false negatives (FN) (Fig. 9). Precision (Eq. 2) calculates the proportion of pixels correctly classified as landslides. Recall (Eq. 3) value represents the number of pixels that was correctly classified as landslides from the total pixels representing landslides.

$$\text{Precision} = \frac{TP}{TP + FP} \times 100 \tag{2}$$

$$\text{Recall} = \frac{TP}{TP + FN} \times 100 \tag{3}$$

F1-score (Eq. 4) is a harmonic mean between precision and recall; therefore, the highest values of F1-score correspond to

models with better performance. Landslide datasets usually have an unbalance between the positive (landslides) and negative (background) classes. Thus, the MCC (Eq. 5) metric is better for comparing imbalanced datasets (Baldi et al. 2000).

$$F1score = \frac{2 \times \text{precison} \times \text{recall}}{(\text{precision} + \text{recall})} \times 100 \tag{4}$$

$$MCC = \frac{TP \times TN - FPxPN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \times 100 \tag{5}$$

The results show that among the models trained with dataset-1, the U-Net 128,5 model achieved the highest MCC (71.06) and F1-score (71.12). Nevertheless, compared with the other algorithms, the MCC results are just 0.63, 1.59, and 2.65 higher than the SVM, KNN, and RF algorithms (Table 2). SVM 1285 achieved the highest precision (80.28), while U-Net 16,5 had the highest recall (83.94).

The U-Net also had better performance in dataset-2 (Table 3). However, in dataset-1, the model trained with 128 × 128 patch size achieved the best F1-score and MCC, while in dataset-2, the model trained with 16 × 16 patch size achieved the highest F1-score (69.42) and MCC (69.70). The patch size seems to be more relevant to dataset-2 since all the models trained with 16 × 16 patch size achieved the best results. In dataset-1, the SVM and KNN trained with the 16 × 16 patch size also had the best results; however, the best U-Net and RF model was trained with 128 × 128 and 32 × 32 patch size, respectively.

Comparing the results of both datasets, the models trained with dataset-1 achieved better results compared to the same algorithm over dataset-2. The U-Net 128,5 was the best overall model among both datasets. Similar to what was observed by Ghorbanzadeh et al. (2019a, b) with machine learning models trained in the same area, the topographical layers helped differentiate human settlement areas, which have identical spectral responses to landslides; however, the models generate more false-positive in the steeper areas. Visually evaluating the segmentation of each algorithms (as seen in Fig. 10), the U-Net segmentation is smoother and more continuous, with greater similarity in comparison to the manual annotations than with the other ML methods. SVM, KNN, and RF results show similar segmentation patterns and mistakes.

Discussion

The U-Net deep learning model achieved the best results in this study based on the metrics used to evaluate the models. However, the MCC and F1-score values were similar among all the models. The results highlight that U-Net can achieve robust results even with few training samples. However, since the machine and deep learning achieved similar accuracies, all the algorithms have similar behavior with a small dataset, and it is impossible to define a better algorithm based on the accuracy metrics. However, similarities between the manual annotations and the U-Net model results are noted in terms of landslide prediction smoothness and continuity, demonstrating better segmentation results than the other models. The models evaluated by Ghorbanzadeh et al. (2019a, b) in the same study area were trained with a bigger dataset composed of 3500 samples, which was augmented to 7000 samples. In that study, the CNN model achieved the best results with an F1-score that was

Fig. 4 a Landslide detection results using U-Net model in sampled area in the test zone using dataset-1. **b** Landslide detection results using U-Net model in sampled area in the test zone using dataset-2
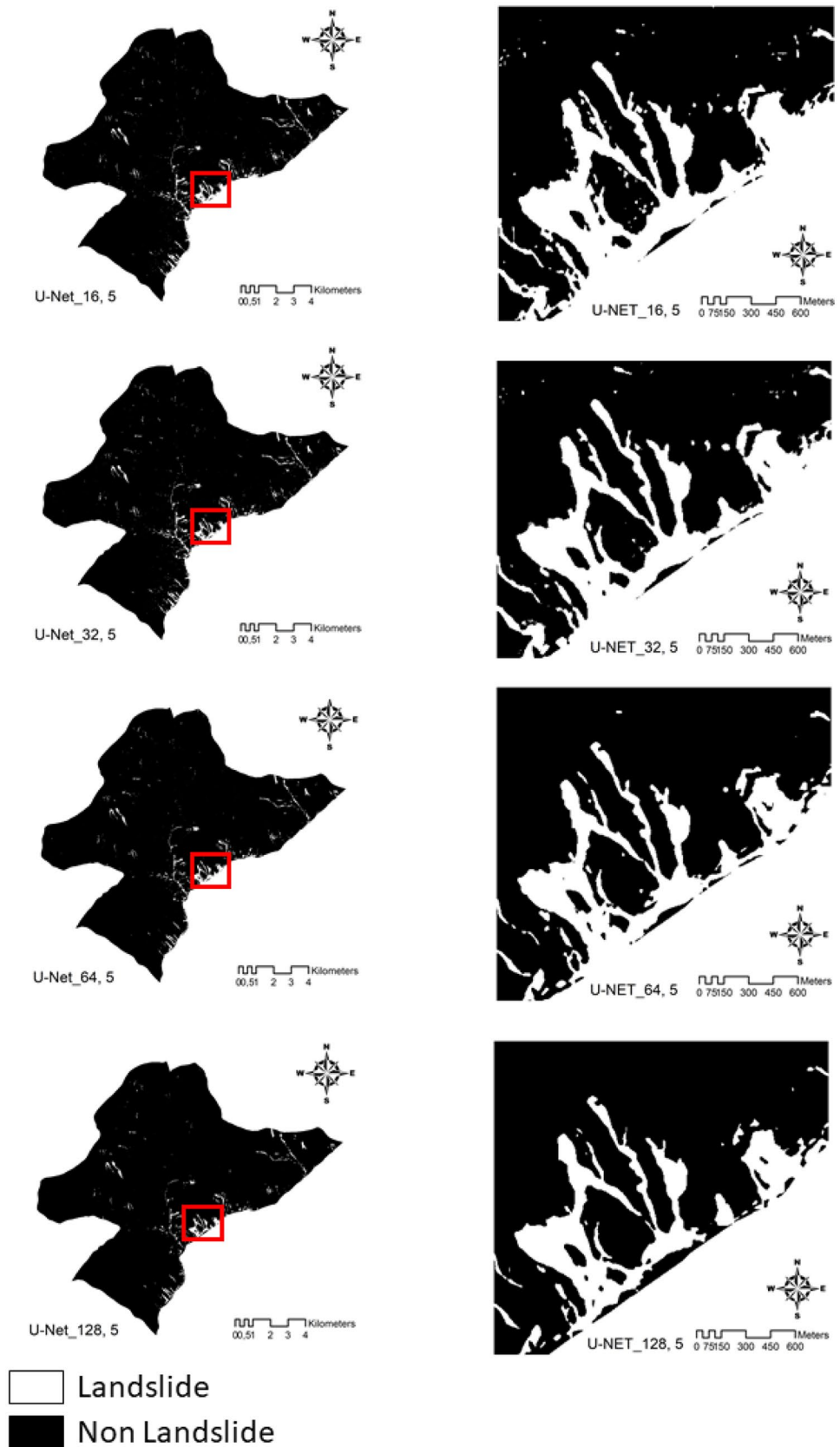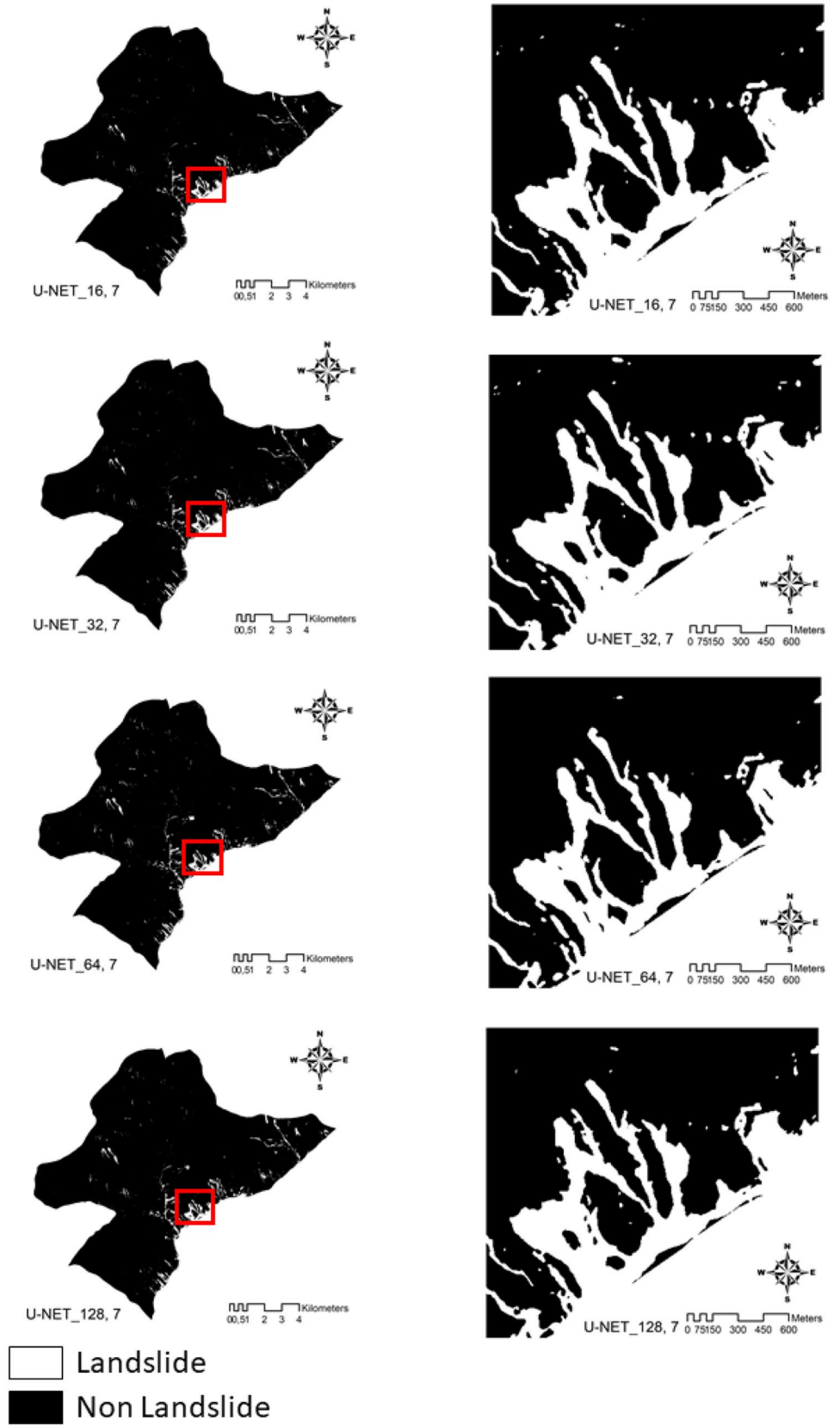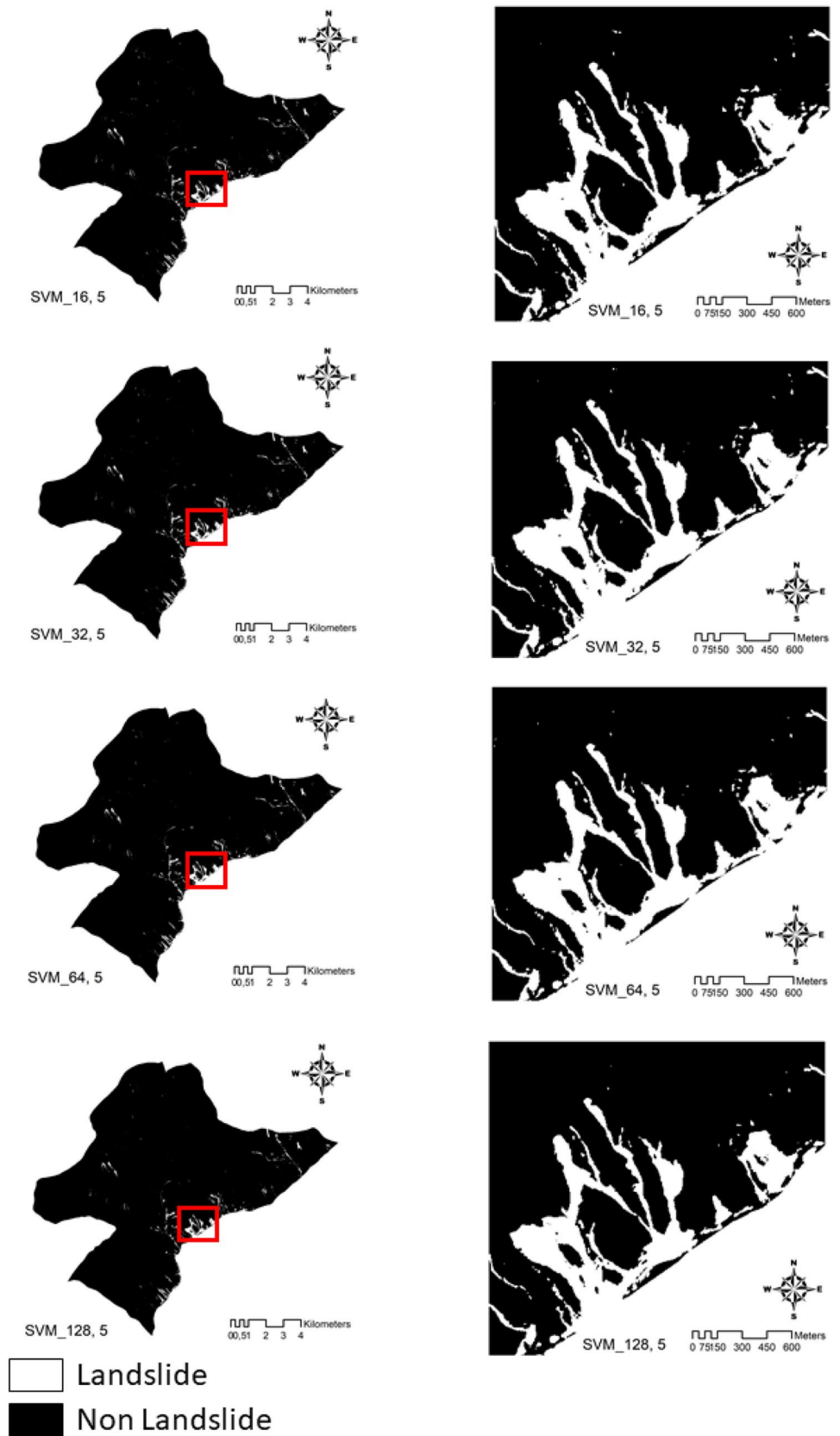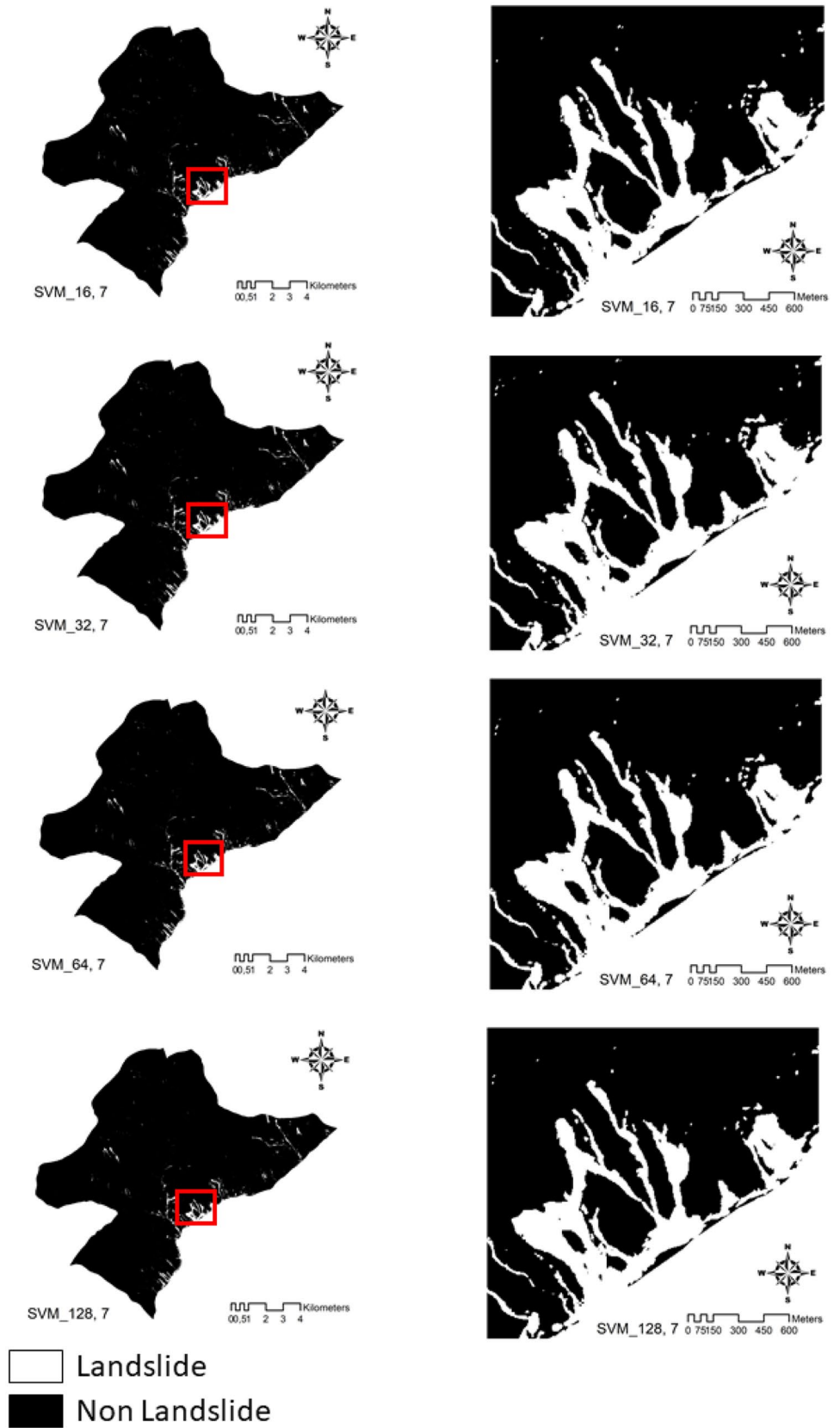

Landslide
Non Landslide

Fig. 4 (continued)

U-NET_16, 7

U-NET_32, 7

U-NET_64, 7

U-NET_128, 7

Landslide
Non Landslide

Fig. 5 a Landslide detection results using SVM model in sampled area in the test zone using dataset-1. **b** Landslide detection results using SVM model in sampled area in the test zone using dataset-2
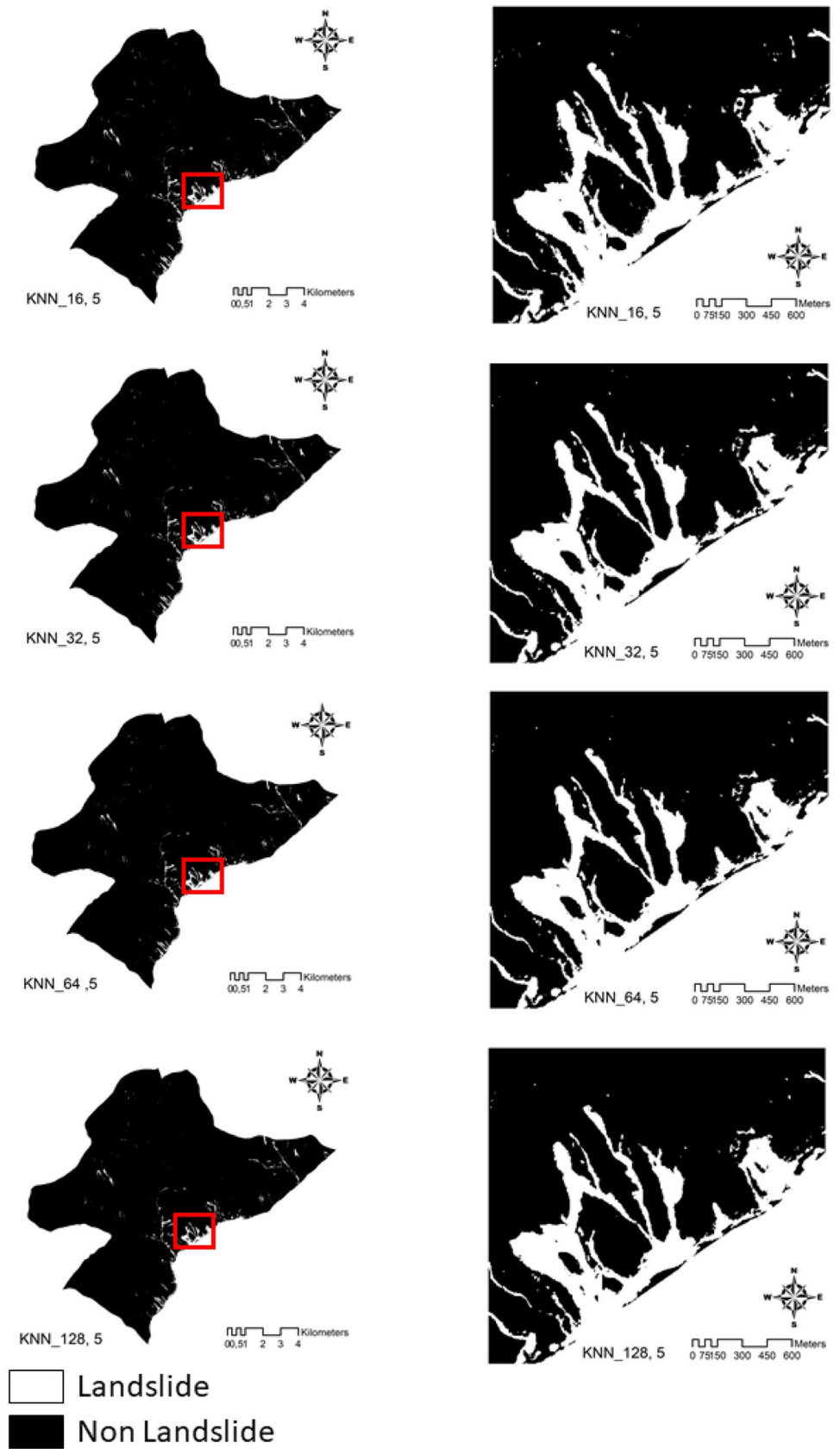
Fig. 5 (continued)

Landslide
Non Landslide

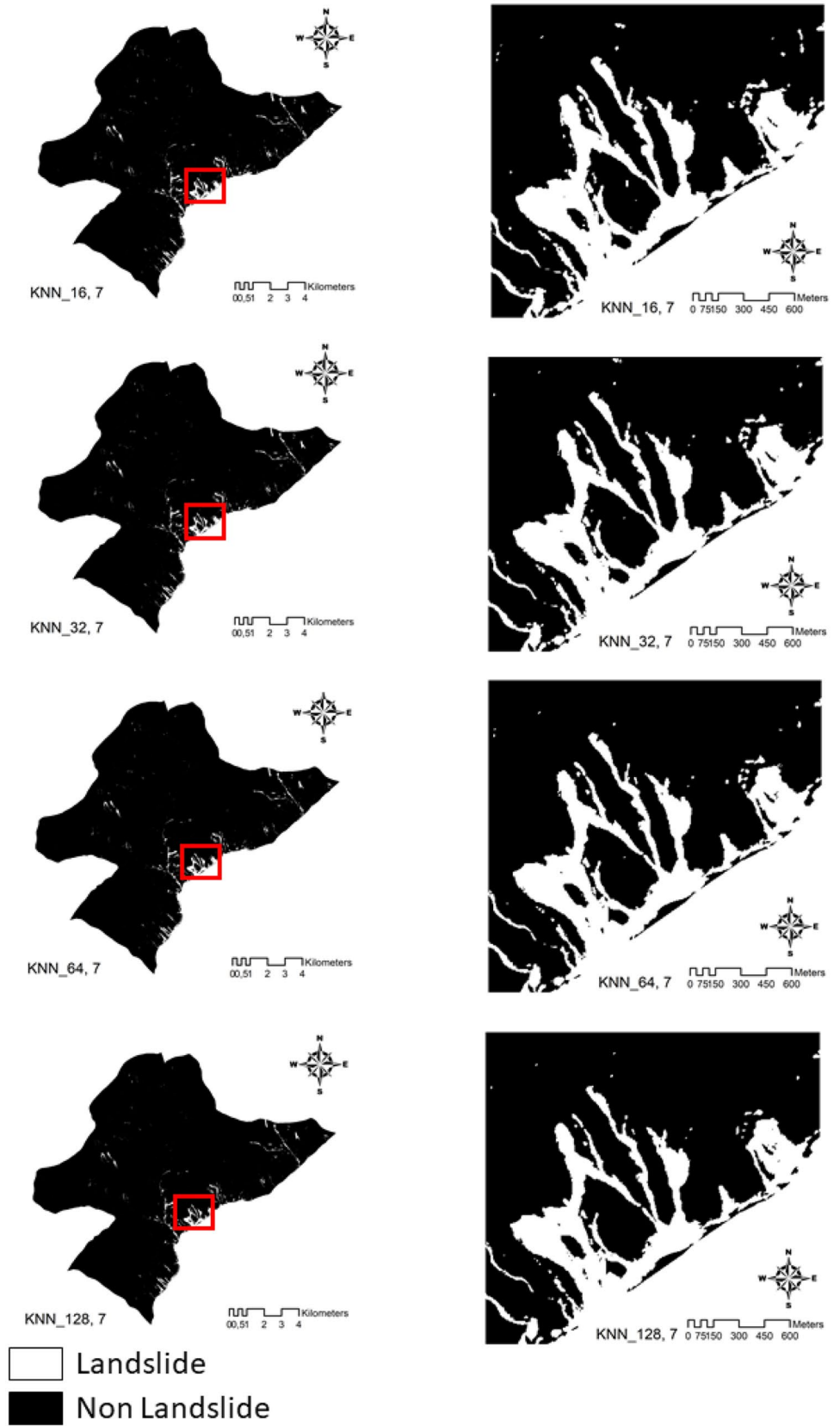**Fig. 6 a** Landslide detection results using KNN model in sampled area in the test zone using dataset-1. **b** Landslide detection results using KNN model in sampled area in the test zone using dataset-2.

KNN_16, 5

KNN_32, 5

KNN_64, 5

KNN_128, 5

KNN_16, 5

KNN_32, 5

KNN_64, 5

KNN_128, 5

Landslide

Non Landslide

Fig. 6 (continued)

KNN_16, 7

KNN_32, 7

KNN_64, 7

KNN_128, 7

Landslide

Non Landslide

Fig. 7 **a** Landslide detection results using RF model in sampled area in the test zone using dataset-1. **b** Landslide detection results using RF model in sampled area in the test zone using dataset-2.
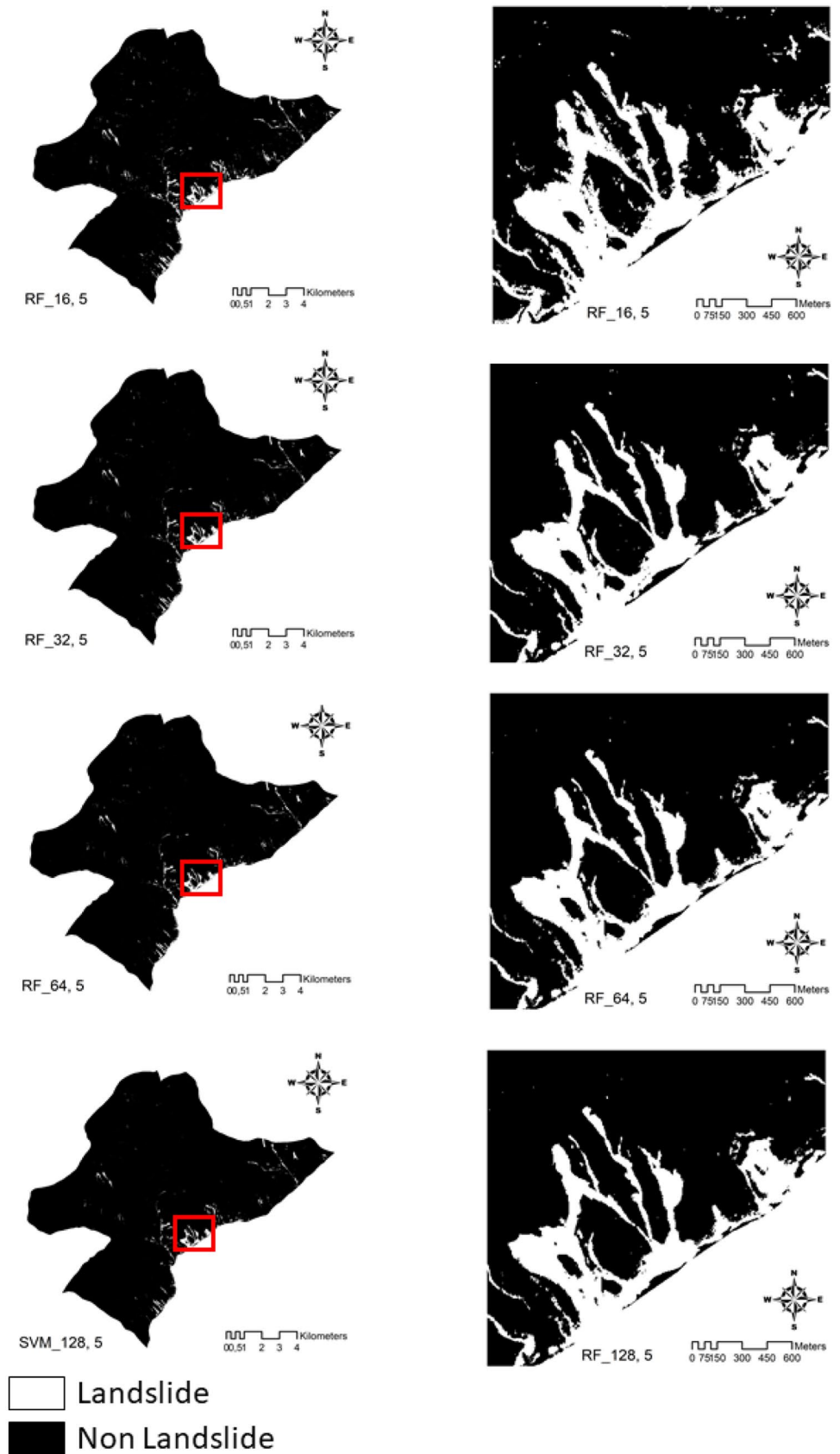

RF_16, 5

RF_32, 5
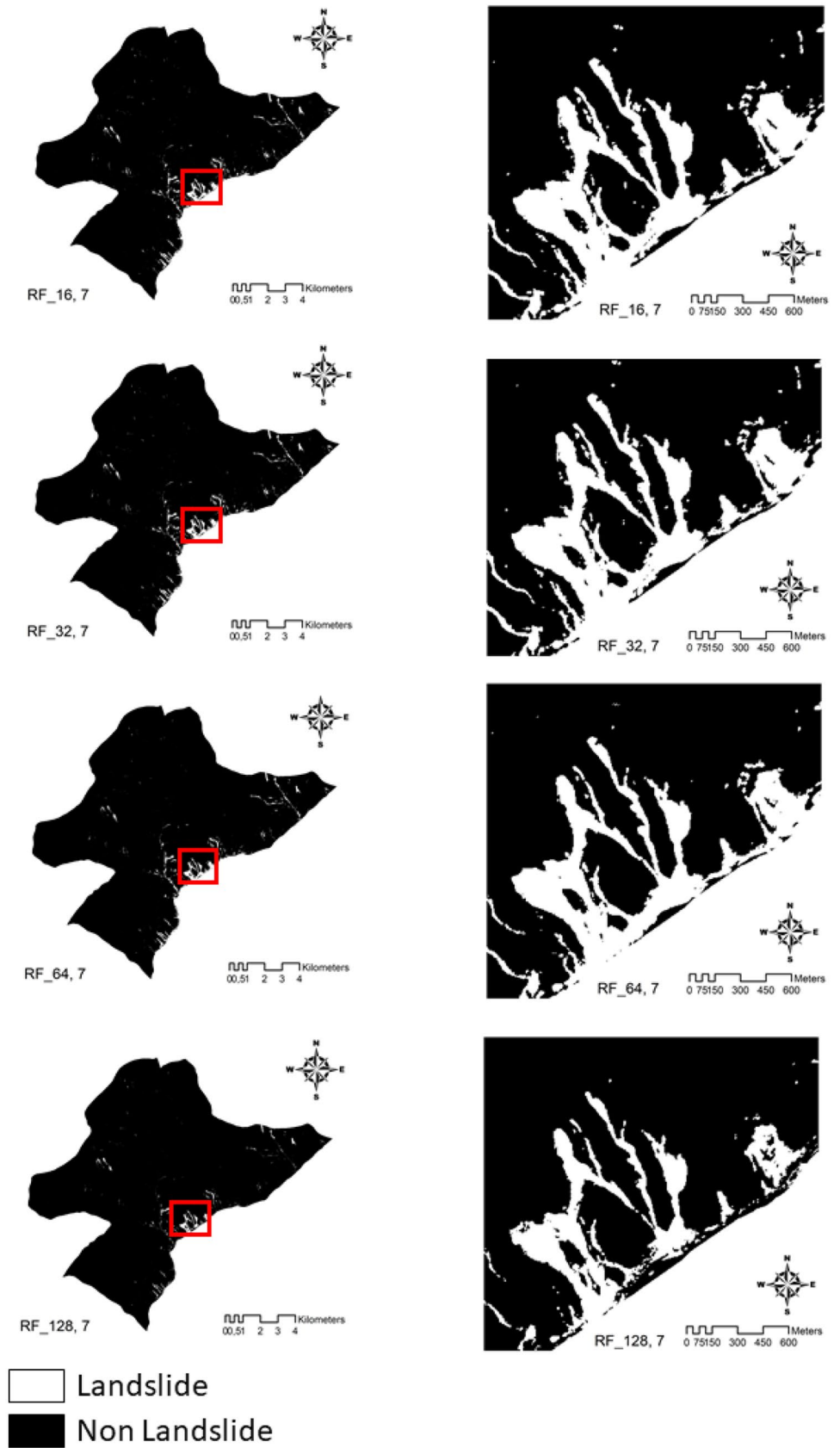
RF_64, 5

SVM_128, 5

RF_16, 5

RF_32, 5

RF_64, 5

RF_128, 5

☐ Landslide
■ Non Landslide

Fig. 7 (continued)

5.73% greater than the best machine learning model. The significant differences in the author's accuracy between the machine and deep learning models highlight the importance of the dataset size. In this study, despite the slightly higher accuracy achieved by the U-Net, the deep learning algorithms were computationally more expensive, needing a GPU (GeForce RTX 2060, 8 GB memory) for the training process, while the machine learning algorithms only used the CPU (Intel I7 10700 K).

The patch size is an important parameter to find the best algorithm since it affects the model's accuracy. The U-Net trained with the optical data showed a similar pattern to what was observed by Soares et al. (2020), where the U-Net models trained with smaller patches (32×32) yield a greater recall while the models trained with the bigger patches (128×128) achieved a greater precision. The models trained with bigger patches became more restrictive (made fewer false-positive errors) than the models trained with smaller patches. Nevertheless, this pattern was not observed in the U-Net models trained with the topographical dataset and on the results achieved by the machine learning models.

The topographical data does not improve the results of the models in this study. This may be related to the resampled DEM used and the samples. Since the dataset is composed of 174 samples, the

Table 2 The results of landslide detection in the study area based on the different ML and U-Net model for dataset-1; accuracies are stated as precision, recall, F1-measure, and MCC. The best values are in bold

Model	Precision	Recall	F1-Score	MCC
U-Net $_{16\_5}$	57.20	83.94	68.03	68.99
U-Net $_{32\_5}$	57.88	81.99	67.85	68.56
U-Net $_{64\_5}$	67.19	71.09	69.09	68.83
U-Net $_{128\_5}$	**76.59**	**66.38**	**71.12**	**71.06**
SVM $_{16\_5}$	75.75	65.95	70.51	70.43
SVM $_{32\_5}$	74.96	62.62	68.23	68.25
SVM $_{64\_5}$	72.43	60.65	66.02	66.00
SVM $_{128\_5}$	80.28	55.84	65.87	66.71
KNN $_{16\_5}$	67.79	71.77	69.73	69.47
KNN $_{32\_5}$	73.66	62.47	67.60	67.57
KNN $_{64\_5}$	73.36	61.39	66.84	66.84
KNN $_{128\_5}$	73.64	60.00	66.16	66.24
RF $_{16\_5}$	58.70	76.54	66.44	66.69
RF $_{32\_5}$	76.64	61.52	68.25	68.41
RF $_{64\_5}$	72.82	60.42	66.04	66.06
RF $_{128\_5}$	73.91	59.73	66.07	66.17

Table 3 The results of landslide detection in the study area based on the different ML and U-Net model for dataset-2; accuracies are stated as precision, recall, F1-measure, and MCC. The best values are in bold

Model	Precision	Recall	F1-Score	MCC
U-Net $_{16\_7}$	**61.46**	**79.74**	**69.42**	**69.70**
U-Net $_{32\_7}$	59.31	80.44	68.28	68.75
U-Net $_{64\_7}$	62.07	77.11	68.78	68.88
U-Net $_{128\_7}$	60.27	78.17	68.35	68.58
SVM $_{16\_7}$	72.20	66.52	69.24	69.04
SVM $_{32\_7}$	66.74	67.34	67.04	66.74
SVM $_{64\_7}$	68.80	63.19	65.88	65.64
SVM $_{128\_7}$	71.32	61.03	65.77	65.69
KNN $_{16\_7}$	66.24	72.58	69.27	69.05
KNN $_{32\_7}$	69.37	66.47	67.89	67.62
KNN $_{64\_7}$	70.25	63.29	66.59	66.39
KNN $_{128\_7}$	71.77	58.72	64.59	64.63
RF $_{16\_7}$	70.22	68.67	69.44	69.17
RF $_{32\_7}$	69.80	66.81	68.27	68.01
RF $_{64\_7}$	69.95	60.04	64.62	64.51

models were not exposed to various topographic features. Therefore, the pattern learned with the training samples may not represent the test area, and consequently, the results were worse. Moreover, the Hughes Phenomenon (Hughes 1968), also known as the Curse of Dimensionality in the field of machine learning, may also be related to the inferior results with the topographical dataset. Since the dataset with two extra topographical bands has a higher dimensionality, a greater number of samples are needed to improve the models' accuracy. The small number of samples used to train the models was not enough for the classifier to reliably classify the landslide areas; therefore, the classification performance degraded with the higher dimensional data. This phenomenon may also justify why the models trained with 16×16 patch size (smaller patch, with lower dimensionality) achieved the best results within this dataset.

The training and test area used in this study have landslides with similar spectral characteristics. Therefore, this may also explain the comparable results achieved with the machine and deep learning models. However, since machine learning algorithms are trained using a one-dimensional vector with pixel values, the spatial pattern of the landslides, such as the shapes, is not learned by those models. Consequently, it is expected that the deep learning method achieves better results in areas with different spectral characteristics than the machine learning algorithm because those models are trained with two-dimensional patches that keep the spatial information of the images. According to the literature, the U-Net like architectures achieve the best results for segmenting landslides in test areas with similar spectral characteristics to the training zones, and test areas with different spectral characteristics highlighting their generalization capacity and good accuracy on landslide segmentation (Qi et al. 2020; Prakasha et al. 2020; Soares et al. 2020; Yi et al. 2020; Prakasha et al. 2021).

Fig. 10 Enlarged maps of sub-area from the test zone. Landslide detection results are overlayed on the inventory data
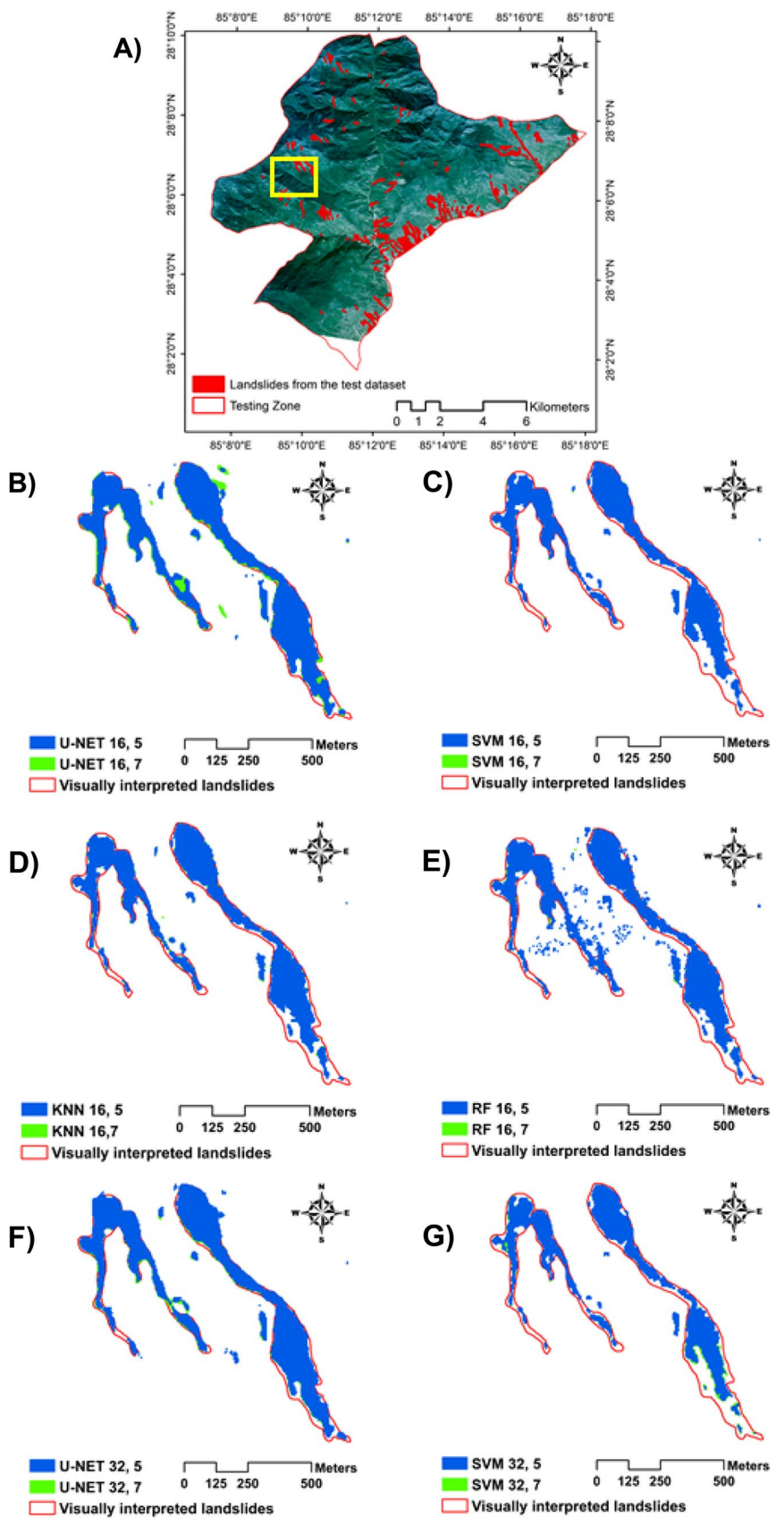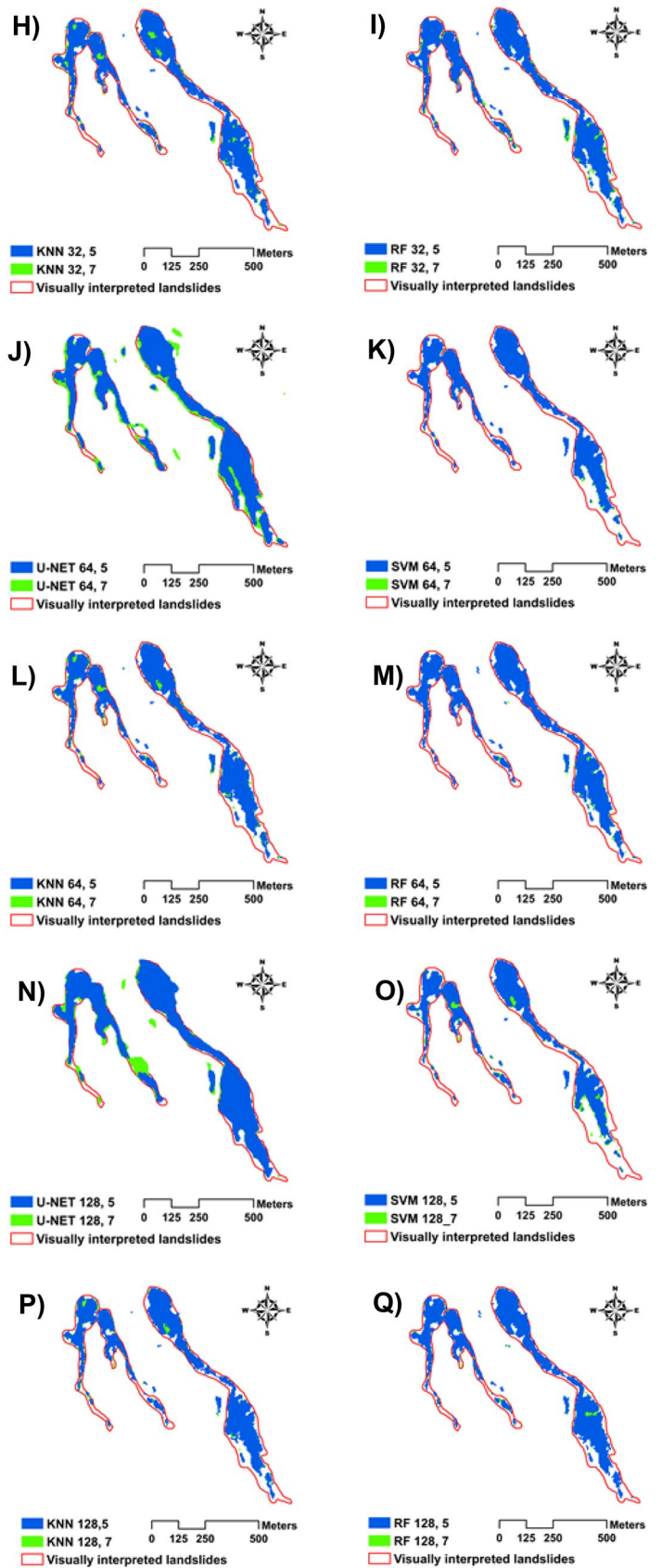
Fig. 10 (continued)

Conclusions

This work evaluates different machine and deep learning model performances trained with small datasets and different patch sizes for landslide segmentation. The U-Net deep learning model achieved the best results on dataset-1 and dataset-2. However, all the models achieved similar MCC and F1-scores, highlighting that deep learning models achieve comparable results to machine learning algorithms with small datasets. The extra topographic features (slope and elevation) did not improve the models' results but yielded improved detection of false-positive such as built-up areas, an error in riverbeds. In this study, U-Net has slightly better results than other machine learning approaches. Although it can depend on the model architecture and the complexity of geographical features in the imagery, the U-Net model is still preliminary when considered for landslide detection. A reason for the U-Net model to perform better is because of the encoder-decoder and skip-connection structure of the model that preserves the structural integrity of the output results even with lower training data (Ronneberger et al. 2015). This exhibits the notion of actually using lesser training data, which is generally the case for new events, and can be then used in training and detecting landslides for newer events.

This study is one of the first efforts of using U-Net for landslide detection in the Himalayas. Nevertheless, U-Net has the potential to further improve automated landslide detection in the future as U-Net excels in producing good results as stated above in regard to the architecture structure but also that since the output is a segmentation result, we are provided with the information of the landslide boundary and the delineation of the landslide body as well. Further adjusting of the encoder part of the model, we can add deeper layers like Virtual Geometry Group (VGG) and Residual Neural Network (ResNet-50) (Simonyan and Zisserman 2014; He et al. 2016) to further improve the results and thereby detecting more landslides with fewer false positives as model complexity overall tends to overcome such artifacts.

The use of only spectral bands can be a limitation for landslide detection since geological and the degree of saturation of the soil directly affect the targets' spectral response. Therefore, areas with higher soil saturation may present darker colors while less saturated areas will have light colors. Moreover, rocks with different weathering conditions will show different spectral responses. Thus, to avoid algorithm misclassifications and improve the results, further studies need to use images covering a more comprehensive range of time and different seasons. This way, the models can learn and predict a broader range of spectral responses of the landslides and achieve better results.

Declarations

Conflict of interests The authors declare no competing interests.

References

Bacha AS, Van der Werff H, Shafique M, Khan H (2020) Transferability of object-based image analysis approaches for landslide detection in the Himalaya Mountains of northern Pakistan. Int J Remote Sens 41:3390–3410. https://doi.org/10.1080/01431161.2019.1701725

Baldi P, Brunak S, Chauvin Y, Andersen CA, Nielsen H (2000) Assessing the accuracy of prediction algorithms for classification: an overview. Bioinformatics 16(5):412–424

Bui TA, Lee PJ, Lum KY, Chen CR, Shiu SH, Ieee (2019) Using BEMD in CNN to identify landslide in satellite image. In: Proceedings of 2019 International Conference on System Science and Engineering. Int Conf Sys Sci Eng 94–97

Chen GY, Li X, Gong WY, Xu H (2020) Recognition of the landslide disasters with extreme learning machine. Int J Comput Sci Eng 21:84–94. https://doi.org/10.1504/ijcse.2020.105215

Chen W, Li X, Wang Y, Chen G, Liu S (2014) Forested landslide detection using LiDAR data and the random forest algorithm: a case study of the Three Gorges. China Remote Sensing of Environment 152:291–301

Chen Z, Zhang Y, Ouyang C, Zhang F, Ma J (2018) Automated landslides detection for mountain cities using multi-temporal remote sensing imagery sensors (Basel) 18.https://doi.org/10.3390/s18030821

Cortes C, Vapnik V (1995) Support-Vector Networks Machine Learning 20:273–297

Dahal RK, Hasegawa S (2008) Representative rainfall thresholds for landslides in the Nepal Himalaya. Geomorphology 100:429–443

Dhital MR (2015) Geology of the Nepal Himalaya: regional perspective of the classic collided orogen. Springer

Ding A, Zhang Q, Zhou X, Dai B (2016) Automatic recognition of landslide based on CNN and texture change detection. In: 2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC). IEEE 444–448

Fan X et al (2018) Coseismic landslides triggered by the 8th August 2017 Ms 7.0 Jiuzhaigou earthquake (Sichuan, China): factors controlling their spatial distribution and implications for the seismogenic blind fault identification Landslides 15:967–983.https://doi.org/10.1007/s10346-018-0960-x

Froude MJ, Petley DN (2016) Global fatal landslide occurrence from 2004 to 2018. Nat Hazard 18:2161–2181. https://doi.org/10.5194/nhess-18-2161-2018

Gariano SL, Guzzetti F (2016) Landslides in a Changing Climate Earth-Science Reviews 162:227–252. https://doi.org/10.1016/j.earscirev.2016.08.011

Ghorbanzadeh O, Blaschke T, Gholamnia K, Meena S, Tiede D, Aryal J (2019a) Evaluation of different machine learning methods and deep-learning convolutional neural networks for landslide detection remote sensing 11. https://doi.org/10.3390/rs1102019a6

Ghorbanzadeh O, Meena SR, Abadi HSS, Piralilou ST, Zhiyong L, Blaschke T (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. https://doi.org/10.1109/JSTARS.2020.3043836

Ghorbanzadeh O, Meena SR, Blaschke T, Aryal J (2019b) UAV-Based Slope Failure Detection Using Deep-Learning Convolutional Neural Networks Remote Sensing 11:2046

Ghorbanzadeh O, Tiede D, Wendt L, Sudmanns M, Lang S (2020) Transferable instance segmentation of dwellings in a refugee camp - integrating CNN and OBIA European. J Remote Sensing:1–14. https://doi.org/10.1080/22797254.2020.1759456

Guzzetti F, Mondini AC, Cardinali M, Fiorucci F, Santangelo M, Chang K-T (2012) Landslide Inventory Maps: New Tools for an Old Problem Earth-Science Reviews 112:42–66. https://doi.org/10.1016/j.earscirev.2012.02.001

Haigh M, Rawat JS (2011) Landslide causes: human impacts on a Himalayan landslide swarm Belgeo Revue belge de géographie 201–220

Hakan T, Luigi L (2020) Completeness index for earthquake-induced landslide inventories Engineering geology 264. https://doi.org/10.1016/j.enggeo.2019.105331

He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition 770–778

Hölbling D, Füreder P, Antolini F, Cigna F, Casagli N, Lang S (2012) A Semi-Automated Object-Based Approach for Landslide Detection Validated by Persistent Scatterer Interferometry Measures and Landslide Inventories Remote Sensing 4:1310–1336. https://doi.org/10.3390/rs4051310

Hughes G (1968) On the mean accuracy of statistical pattern recognizers. IEEE Trans Inf Theory 14(1):55–63

Ji SP, Yu DW, Shen CY, Li WL, Xu Q (2020) Landslide detection from an open satellite imagery and digital elevation model dataset using attention boosted convolutional neural networks. Landslides 17:1337–1352. https://doi.org/10.1007/s10346-020-01353-2

Jin B, Ye P, Zhang X, Song W, Li S (2019) Object-oriented method combined with deep convolutional neural networks for land-use-type classification of remote sensing images. Journal of the Indian Society of Remote Sensing 47:951–965. https://doi.org/10.1007/s12524-019-00945-3

Khanal N, Watanabe T (2005) Landslide and debris flow hazards induced by heavy precipitation in Nepal. In: Proceedings of Himalaya International Symposium on Landslide Hazard in Orogenic Zone from the Himalaya to Island Arc in Asia 111–118

Lei T, Zhang Y, Lv Z, Li S, Liu S, Nandi AK (2019) Landslide inventory mapping from bitemporal images using deep convolutional neural networks. IEEE Geosci Remote Sens Lett 16:982–986. https://doi.org/10.1109/lgrs.2018.2889307

Liu P, Wei Y, Wang Q, Chen Y, Xie J (2020a) Research on post-earthquake landslide extraction algorithm based on improved U-Net model Remote sensing 12.https://doi.org/10.3390/rs12050894

Liu P, Wei YM, Wang QJ, Chen Y, Xie JJ (2020b) Research on Post-earthquake landslide extraction algorithm based on improved U-Net model Remote sensing 12.https://doi.org/10.3390/rs12050894

Liu S, Qi Z, Li X, Yeh AG-O (2019) Integration of convolutional neural networks and object-based post-classification refinement for land use and land cover mapping with optical and SAR data Remote sensing 11:690

Lu H, Ma L, Fu X, Liu C, Wang Z, Tang M, Li NW (2020) landslides information extraction using object-oriented image analysis paradigm based on deep learning and transfer learning Remote sensing 12. https://doi.org/10.3390/rs12050752

Marjanovic M, Bajat B, Kovacevic M (2009) Landslide susceptibility assessment with machine learning algorithms. In: 2009 International Conference on Intelligent Networking and Collaborative Systems. IEEE 273–278

Martha TR, Kerle N, Jetten V, van Westen CJ, Kumar KV (2010) Characterising spectral, spatial and morphometric properties of landslides for semi-automatic detection using object-oriented methods. Geomorphology 116:24–36.https://doi.org/10.1016/j.geomorph.2009.10.004

Meena S, Chauhan A, Bhuyan K, Singh RP (2021a) Impact of the Chamoli disaster on flood plain and water quality along the Himalayan rivers. Paper presented at the EGU General Assembly

Meena SR, Ghorbanzadeh O, van Westen CJ, Nachappa TG, Blaschke T, Singh RP, Sarkar R (2021b) Rapid mapping of landslides in the Western Ghats (India) triggered by 2018 extreme monsoon rainfall using a deep learning approach Landslides.https://doi.org/10.1007/s10346-020-01602-4

Meena SR, Tavakkoli Piralilou S (2019) Comparison of earthquake-triggered landslide inventories: a case study of the 2015 Gorkha earthquake. Nepal Geosciences 9:437

Moosavi L, Mahyuddin N, Ab Ghafar N, Ismail MA (2014) Thermal Performance of Atria: an Overview of Natural Ventilation Effective Designs Renewable and Sustainable Energy Reviews 34:654–670

Mountrakis G, Im J, Ogole C (2011) Support vector machines in remote sensing: a review. ISPRS J Photogramm Remote Sens 66:247–259

Pawłuszek K, Borkowski A (2016) Landslides identification using airborne laser scanning data derived topographic terrain attributes and support vector machine classification. In: The international archives of the photogrammetry, remote sensing and spatial information sciences, XXIII ISPRS Congress

Petley D (2012) Global patterns of loss of life from landslides. Geology 40:927–930

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.https://doi.org/10.3390/rs12030346

Prakash N, Manconi A, Loew S (2021) A new strategy to map landslides with a generalized convolutional neural network. Sci Rep 11(1):1–15

Qi W, Wei M, Yang W, Xu C, Ma C (2020) Automatic mapping of landslides by the resU-Net Remote sensing 12. https://doi.org/10.3390/rs12152487

Reichenbach P, Rossi M, Malamud BD, Mihir M, Guzzetti F (2018) A Review of Statistically-Based Landslide Susceptibility Models Earth-Science Reviews 180:60–91. https://doi.org/10.1016/j.earscirev.2018.03.001

Ronneberger O, Fischer P, Brox T (2015) U-Net: convolutional networks for biomedical image segmentation. International Conference on Medical image computing and computer-assisted intervention. Springer, pp 234–241

RapidEye AG (2011) RapidEye satellite imagery product specifications. Version 3.2. RapidEye AG, product manual

Sameen MI, Pradhan B (2019) Landslide detection using residual networks and the fusion of spectral and topographic information IEEE. Access 7:114363–114373. https://doi.org/10.1109/access.2019.2935761

Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556

Shi W, Zhang M, Ke H, Fang X, Zhan Z, Chen S (2020) Landslide recognition by deep convolutional neural network and change detection IEEE transactions on geoscience and remote sensing 1–19. https://doi.org/10.1109/TGRS.2020.3015826

Soares LP, Dias HC, Grohmann CH (2020) Landslide segmentation with U-Net: Evaluating different sampling methods and patch sizes. arXiv preprint arXiv:2007.06672

Su Z, Chow JK, Tan PS, Wu J, Ho YK, Wang Y-H (2020) Deep convolutional neural network–based pixel-wise landslide inventory mapping Landslides.https://doi.org/10.1007/s10346-020-01557-6

Thapa P, Dhital M (2000) Landslide and debris flows of 19–21 July 1993 in the Agra Khola watershed of central. Nepal Jour Nepal Geol Soc 21:5–20

Upreti BN, Dhital M (1996) Landslide studies and management in Nepal

Wang Y, Wu X, Chen Z, Ren F, Feng L, Du Q (2019) Optimizing the predictive ability of machine learning methods for landslide susceptibility mapping using SMOTE for Lishui City in Zhejiang Province, China. Int J Environ Res Public Health 16. https://doi.org/10.3390/ijerph16030368

Whitworth M, Giles D, Murphy W Identification of landslides in clay terrains using Airborne Thematic Mapper (ATM) multispectral imagery.

In: Remote Sensing for Environmental Monitoring, GIS Applications, and Geology, 2002. Int Soc Optics and Photonics 216–224

Ye CM et al (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 12:5047–5060. https://doi.org/10.1109/jstars.2019.2951725

Yi Y, Zhang Z, Zhang W, Jia H, Zhang J (2020) Landslide susceptibility mapping using multiscale sampling strategy and convolutional neural network: A case study in Jiuzhaigou region. Catena 195:104851

Yu B, Chen F, Xu C (2020) Landslide detection based on contour-based deep learning framework in case of national scale of Nepal in 2015 Computers & geosciences 135.https://doi.org/10.1016/j.cageo.2019.104388

Yu H, Ma Y, Wang LF, Zhai YS, Wang XQ, Ieee (2017) A landslide intelligent detection method based on CNN and RSG_R. 2017 Ieee Int Conf Mechatronics and Automation

Zhang P, Ke Y, Zhang Z, Wang M, Li P, Zhang S (2018) Urban Land Use and Land Cover Classification Using Novel Deep Learning Models Based on High Spatial Resolution Satellite Imagery Sensors 18:3717

Sansar Raj Meena (✉) · **Cees** van **Westen**
Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Enschede, The Netherlands
Email: s.r.meena@utwente.nl

Sansar Raj Meena · Kushanav Bhuyan · Mario Floris · Filippo Catani
Machine Intelligence and Slope Stability Laboratory, Department of Geosciences, University of Padova, Padova, Italy
Email: s.r.meena@utwente.nl

Lucas Pedrosa Soares
Institute of Geosciences, University of São Paulo (IGc-USP), São Paulo 05508-080, Brazil

Carlos H. Grohmann
Institute of Energy and Environment, University of São Paulo (IEE-USP), São Paulo 05508-010, Brazil

Lucas Pedrosa Soares · Carlos H. Grohmann
Spatial Analysis and Modelling Lab (SPAMLab) -, São Paulo, Brazil

Ramesh P. Singh
School of Life and Environmental Sciences, Schmid College of Science and Technology, Chapman University One University Drive, Orange, CA, USA