Article

Matrix SegNet: A Practical Deep Learning Framework for Landslide Mapping from Images of Different Areas with Different Spatial Resolutions

Bo Yu [1]**, Fang Chen** [1,2,3,4,*]**, Chong Xu** [5] **, Lei Wang** [4] **and Ning Wang** [1,2]

[1] Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China; boyu@radi.ac.cn (B.Y.); wangning@radi.ac.cn (N.W.)

[2] University of Chinese Academy of Sciences, Beijing 100049, China

[3] Hainan Key Laboratory of Earth Observation, Aerospace Information Research Institute, Chinese Academy of Sciences, Sanya 572029, China

[4] State Key Laboratory of Remote Sensing Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China; wanglei@radi.ac.cn

[5] National Institute of Natural Hazards, Ministry of Emergency Management of China, Beijing 100085, China; chongxu@ninhm.ac.cn

[*] Correspondence: chenfang_group@radi.ac.cn

Abstract: Practical landslide inventory maps covering large-scale areas are essential in emergency response and geohazard analysis. Recently proposed techniques in landslide detection generally focused on landslides in pure vegetation backgrounds and image radiometric correction. There are still challenges in regard to robust methods that automatically detect landslides from images with multiple platforms and without radiometric correction. It is a significant issue in practical application. In order to detect landslides from images over different large-scale areas with different spatial resolutions, this paper proposes a two-branch Matrix SegNet to semantically segment input images by change detection. The Matrix SegNet learns landslide features in multiple scales and aspect ratios. The pre- and post- event images are captured directly from Google Earth, without radiometric correction. To evaluate the proposed framework, we conducted landslide detection in four study areas with two different spatial resolutions. Moreover, two other widely used frameworks: U-Net and SegNet, were adapted to detect landslides via the same data by change detection. The experiments show that our model improves the performance largely in terms of recall, precision, F1-score, and IOU. It is a good starting point to develop a practical, deep learning landslide detection framework for large scale application, using images from different areas, with different spatial resolutions.

Keywords: landslide detection; Matrix nets; different spatial resolutions

1. Introduction

Landslides have become devastating hazards around the globe, especially in mountainous regions. They are associated with climate change, and can cause substantial loss of life and property [1,2]. Diving into the mechanisms of landslides for accurate landslide prediction calls for practical landslide mapping from large scale areas, with complicated background objects. Many researchers have developed advanced technologies in landslide detection, but they still face challenges in transferring the technology for practical application [3–6].

Generally, landslide detection technology from remotely sensed images can be grouped into two categories: image enhancement based [7] and machine learning based [8]. The image enhancement based category enhances landslide information over other background objects through typical image processing techniques, such as Markov random field [5,9], morphological operations [10], object-oriented image segmentation [11,12], and the image enhancement method [8]. One main drawback of the methods abovementioned is the

parameter tuning process, which has to be conducted by trial and error. In terms of building machine learning models in detecting landslides, the manual parameter tuning process can be avoided. Random forest is a popular model that is flexible and has less probability of overfitting [13]. It has been widely used in landslide detection by classification [14,15] or regression [16]. Apart from random forest, support vector machine (SVM), decision tree, and k-means clustering methods have been applied to detect landslides [17–19]. The main obstacle in applying machine learning models in practical applications is the time-consuming feature engineering. It is quite difficult to design appropriate feature to distinguish landslide from bare soil, because soil takes similar spectral and textural characteristics with landslides. Change detection is a commonly used framework in landslide detection, by exploring the spectral or textural changes between images from different time domains [20]. However, due to different imaging conditions, the captured remotely sensed images mostly require radiometric correction to synchronize [5].

The advent of deep convolutional neural networks (CNNs) makes it possible to overcome the shortages abovementioned, by fitting target functions to achieve reasonable detection performance using convolution layers. It has been widely used in computer vision, achieving state-of-the-art performance approval [21,22]. Several studies have explored the application capabilities of deep learning frameworks in landside detection [23,24]. Ding et al. [25] applied a six-layer CNN architecture to detect one small landslide event in Shenzhen and discovered a low commission error. A seven-layer CNN architecture, synthesized with an improved region growing method, was adopted to detect two landslide events [26], achieving high sensitivity. Omid, et al. [23] explored the performance of different technologies in landslide detection by conducting a comparison of artificial neural networks, support vector machines, random forest, an eight-layer CNN, and a five-layer CNN with different input image sizes. The best mean intersection-over-union (mIOU) of 78.26% was achieved from the five-layer deep learning method with an input image size of 16×16 pixel. Apart from application in high spatial resolution images, there is also research exploring landslide detection from hyperspectral remote sensing images using a deep belief network to extract features of landslides, which are further used for detection by a logistic regression classifier [24]. In accordance with the development of deep learning architectures, Bo Yu et al., [27] proposed an object-based pyramid scene parsing network (PSPNet) [28] to detect landslides in Nepal in 2015, from the annually synthesized Landsat 8 image, obtained from Google Earth Engine. They saw a 44% improvement in recall and 15% improvement in precision compared with the traditional method. Nikhil et al. proposed a modified U-Net [29] to detect landslides from high spatial resolution images and achieved a detection rate of 0.72.

The studies above are generally conducted on images with the same spatial resolution after radiometric correction with the training data. It largely simplifies the cases in evaluating the practicability of the proposed model in different areas for actual application, such as emergency response. In this paper, we proposed an end-to-end deep semantic segmentation framework of a two-branch Matrix SegNet to detect landslides. It was assembled from state-of-the-art key point detector Matrix Nets (xNets) [21]. The training and evaluating images of the proposed model were both achieved from different study areas with different spatial resolutions. Semantic segmentation is a typical application branch in computer vision [30], widely used in unmanned driving [31] and robotic domains [32]. It assigns a semantic label to each pixel in the image. Given two input images in different time domains, our pipeline extracts feature maps with a backbone of ResNet-50, respectively, and enhances the more significant features using a squeeze-and-excitation (SE) module [33]. The modified xNets module, inspired by Matrix Nets [21], is further adopted to learn features with various scales and aspect ratios. A binary segmentation layer is added up to generate a binary image with pixels labeled as landslides or background objects. The proposed deep learning framework is evaluated by four different areas, with multiple scales of landslides and complicated background objects. Furthermore, we compare the

detection performance with the widely used network architecture U-Net and SegNet. The main contributions of our manuscript are as follows:

(1) We proposed a practical deep learning framework for landslide mapping from different areas with different spatial resolution images.
(2) We explored the capability of the proposed model in detecting landslides with complicated background objects.
(3) We performed landslide detection by learning features with multiple scales and aspect ratios, without radiometric correction on the input images.

The rest of our manuscript is organized as follows: the related works with the proposed network architecture and recent developments in semantic segmentation are presented in Section 2. Section 3 introduces our proposed framework and experimental settings. The study area and data collection are demonstrated in Section 4. Sections 5 and 6 illustrate the evaluations and discussions, respectively. Our conclusions are presented in the final section.

2. Related Works

2.1. Semantic Segmentation Development

Generally, semantic segmentation has two executing paradigms [34], one is two-stage segmentation, the other is one-stage. The two-stage segmentation method segments images after detecting the bounding box of each object. Mask R-CNN [35] is a typical method, and lays the foundation for the development of other methods, such as PANet [36], and Mask Scoring R-CNN [35]. Two-stage segmentation can achieve high accuracy but low efficiency. The one-stage end-to-end network architecture is more practical in actual application cases (we adopt it in this study). It mostly takes the structure of the encoder–decoder; encoding the input images into feature maps is conducted by spatial pooling or atrous convolution and decoding the feature maps to restore segmentation detail is achieved by deconvolution or up-sampling operations. SegNet [37], U-Net [29], PSPNet [28], and the series of DeepLab models [38,39] all take such typical network structures. In terms of landslide detection, U-Net is a commonly used network architecture [40] because it is easy to train and is highly efficient [41].

2.2. Matrix Nets (xNets)

The xNets module was originally proposed in an object detection framework [21]. It was designed to detect key-points by learning heat maps, corner regression, or center regression, simultaneously. xNets, inspired by the Feature Pyramid Networks proposed for object detection [42], was adopted to down-sample the input feature maps by convolution in horizontal, vertical, and diagonal directions, separately. Therefore, it can learn features with multiple scales and aspect ratios, which are necessary in landslide detection from images with different spatial resolutions. Synthesizing the works in [21,42], a two-branch Matrix SegNet is proposed in this manuscript, modifying xNets module to detect landslides. The images before and after the landslide event are drawn as inputs of the two-branch Matrix SegNet and encoded to the feature maps separately, without radiometric correction, as required in [5].

3. Proposed Architecture

Due to the large proportion of background objects in remotely-sensed images, we first adopted potential landslide detection on the pre- and post-event images for each study area, to remove the background objects that were easily separable. A raw image example with a spatial resolution of 19 m is shown in Figure 1. Since a landslide takes similar spectral characteristics with bare soil in the image, we assigned all of the pixels with intensity values in the green channel smaller than the red channel as a potential landslide, and others as background objects. Such operation was conducted because the red channel was more sensitive to bare soil than the green and blue channels [43]. By

eliminating enormous background object pixels, potential landslide detection can save much computation consumption.
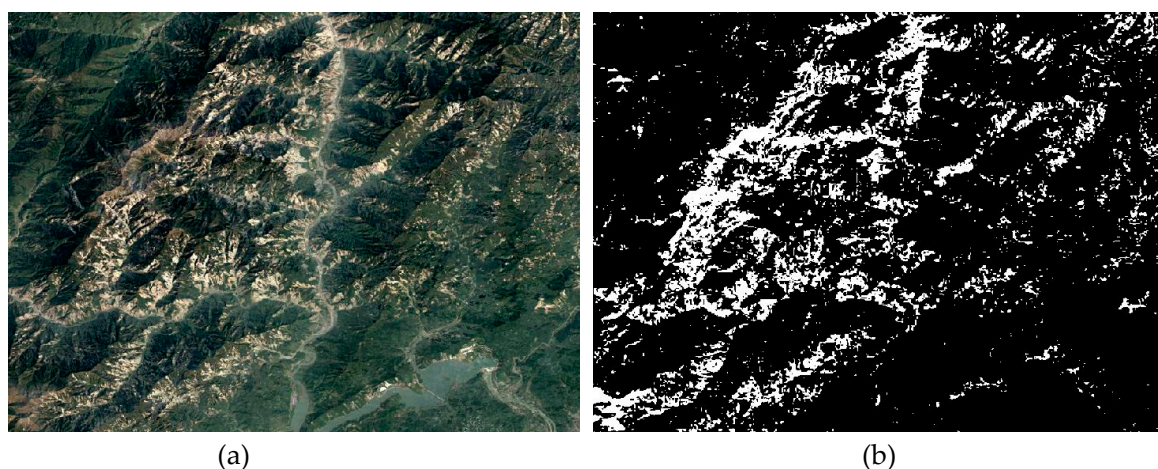

(a)
(b)

Figure 1. Example of a raw image and potential landslide extracted based on a raw image: (**a**) raw image; (**b**) potential landslide extracted image.

Based on the detected potential landslide image, we calculated connective contour for the potential landslide pixel groups, according to their distribution. Each contour is recognized as a potential landslide. Among the potential landslide contours, bare soil was a major background object that was difficult to recognize. Therefore, we proposed a two-branch Matrix SegNet to enhance the landslide feature learning ability in distinguishing landslides from confusable background objects. The proposed model is trained on image patches that are cropped from the original images based on each connective contour. This strategy could enhance the distinguishing ability of the proposed model more directly and avoid the unbalanced sample distribution issue commonly confronted in natural hazard detection, where the number of landslide pixels is smaller than that of background objects in magnitudes.

As shown in Figure 2, our proposed semantic segmentation network takes a twin-tower structure to explore the land cover change of the images before and after the landslide event.
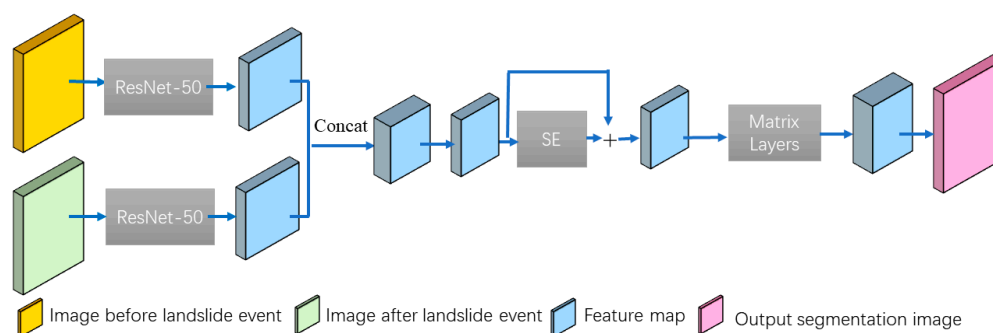

Figure 2. Proposed semantic segmentation network architecture, wherein the concat operation indicates the concatenation of two feature maps, + means the addition operation of the input feature map and output feature map of the SE module.

The twin-tower architecture is employed according to the conclusion in [44]; it is better than concatenating the input images directly. The two input images are encoded into two feature maps with a backbone of ResNet-50 [45], respectively. ResNet-50 is a 50-layer residual network, a widely used network backbone in learning features [29,37,42]. It was proposed to solve the problem of model degradation by adding an identity mapping

in each network building block. Model degradation is a common issue raised by the continuously increasing layers of the network. The encoded feature maps by backbone are further concatenated and enhanced by the squeeze-and-excitation (SE) module, as demonstrated in Figure 3. SE enhances features by learning the weight for each channel of feature maps, multiplying the learnt weight with each corresponding feature map channel. The enhanced feature maps are further used for learning features with multiple scales and aspect ratios using the matrix convolution module. Figure 4 shows the detailed network structure of the matrix convolution module. It consists of 5×5 matrix of convolution operations, wherein vertical convolution (shown in yellow) down-samples the input feature maps vertically, horizontal convolution (shown in red) down-samples the input feature maps horizontally, and diagonal convolution (shown in green) down-samples the input feature maps in both directions. By sampling the feature maps in three different ways simultaneously, the matrix convolution can extract features in multiple scales and aspect ratios and generate an output feature map *Fmo*. It takes the same size with that of the input feature map *Fmi.* To enlarge the feature scale extraction, we concatenate the feature maps *Fmi* and *Fmo* for final convolution to produce a semantic segmentation result image.
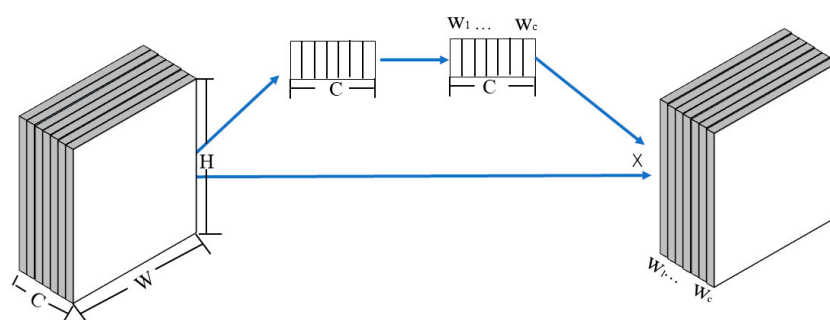

Figure 3. Detailed network structure of the squeeze-and-excitation (SE) module, wherein C is the channel of the feature map, W and H are width and height of the feature maps, and $W_1, \ldots W_c$ is the weight of each channel of the feature map.
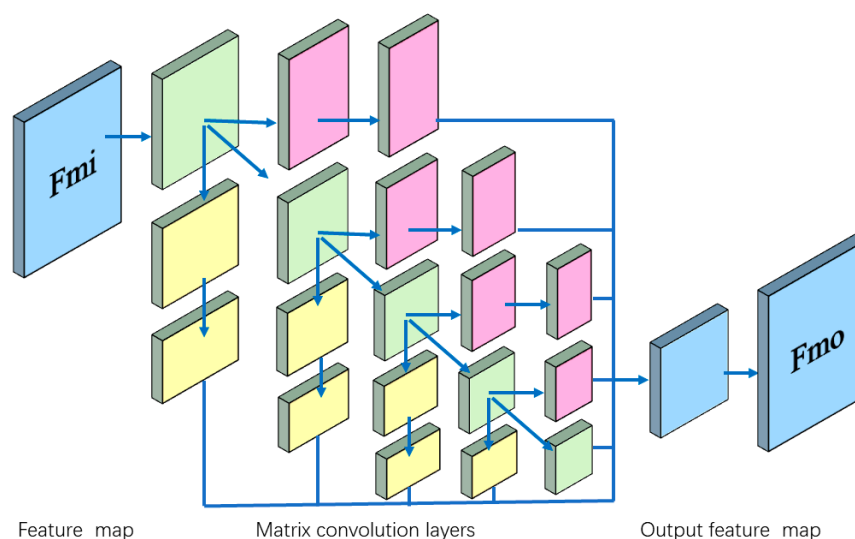
Figure 4. Network structure of the matrix convolution layers in our framework, wherein Fmi is the input feature map, Fmo is the output feature map, yellow feature maps are obtained by vertical convolution from the layer above, pink feature maps are calculated by horizontal convolution from the layer above, and green feature maps are achieved by diagonal convolutions.

Inspired by RetinaNet [46], focal loss function is adopted in our framework training pipeline. It was proposed to deal with the unbalanced sample distribution between positive

and negative samples, which is commonly confronted in landslide detection. Focal loss is modified from binary cross entropy loss, as stated in Equation (1), wherein y_{gt} indicates the ground truth label, and p_{pred} stands for the probability calculated by the model that input sample belongs to label 1 (binary 0/1 classification task). The weight of negative samples is reduced in the convergence of the training model by adding weight factor α, as shown in Equation (2). Moreover, focal loss focuses on difficult samples by adding focusing index β. It can adjust the ratio of easy examples, being assigned a small weight using β. Following the works in [46], β is set as 2 and α is set as 0.25.

$$L_{entropy} = -y_{gt}logp_{pred} - (1 - y_{gt}) \log(1 - p_{pred}) \tag{1}$$

$$L_{focal} = -\alpha(1 - p_{pred})^{\beta}y_{gt}logp_{pred} - (1 - \alpha)p_{pred}^{\beta}(1 - y_{gt}) \log(1 - p_{pred}) \tag{2}$$

4. Study Area and Dataset Preparation

In order to evaluate the transferability and robustness of our proposed two-branch Matrix SegNet, it was applied to detect landslides in four research areas, including the Lushan earthquake impacted area (Lushan in short), the Jiuzhaigou earthquake impacted area (Jiuzhaigou in short), the Central Nepal area (Nepal in short), and Southern Taiwan (Taiwan in short). All images before and after the landslide events were collected from Google earth (because they are public and free). The images collected were directly used for the training model, without radiometric correction. It added more difficulty in evaluating the proposed model. Due to different imaging times and limited data provided from Google Earth, there were still some high spatial resolution images missing after the landslide event. Therefore, we adopted different resolutions for different study areas, as long as they generally covered the impacted area, with an acquisition time of within 2 years after the event. The unified resolution for each study area was selected as high as possible. Table 1 presents general information of the images used for each landslide event, and the detailed process of each corresponding dataset construction can be referred to in the following part.

Table 1. General information of images used for each landslide event.

Dataset	Pre-Event Image Time	Post-Event Image Time	Image Size	Spatial Resolution	Landslide/Non-Landslide Pixel Ratio (%)	No. of Landslides
Lushan	31 December 2010	31 December 2013	5626 × 5088	19 m	0.084	11,754
Jiuzhaigou	7 December 2015	13 August 2017 14 August 2017	8999 × 9890	2.39 m	0.55	3817
Nepal	12 December 2014	27 September 2019 9 November 2015	6627 × 5985	2.39 m	0.26	1126
Taiwan	31 December 2008	31 December 2009	2685 × 4105	19 m	0.60	359

4.1. Lushan Earthquake-Induced Landslide

The magnitude of the 7.0 Ms Lushan earthquake occurred in Lushan County, Sichuan Province, on 20 April 2013 (shown in Figure 5). It triggered tens of thousands of landslides [47] and led to severe casualties and loss of wealth. The landslides were manually interpreted from high spatial resolution images on the GIS platform [47]. With the landslide inventories interpreted in [47], we collected Google Earth images covering the impacted area before and after the earthquake, with a spatial resolution of 19 m. The images were acquired on 31 December 2010, and 31 December 2013, respectively. From Figure 5, we can recognize that forestry and road networks occupy the majority part of the study area. Landslides are distributed intensively along the side of road networks. Since the landslide inventories in [47] are visually interpreted from images with resolutions of 1 to 15 m, they cover more details than what we can achieve from 19 m resolution images. Therefore, we adjusted the landslide inventories and removed the landslides that could not be visually interpreted from the 19 m resolution images. After modification, the number of landslide

inventories reached 11,754, which could still provide abundant training landslide pixel samples to build up a landslide detection model.
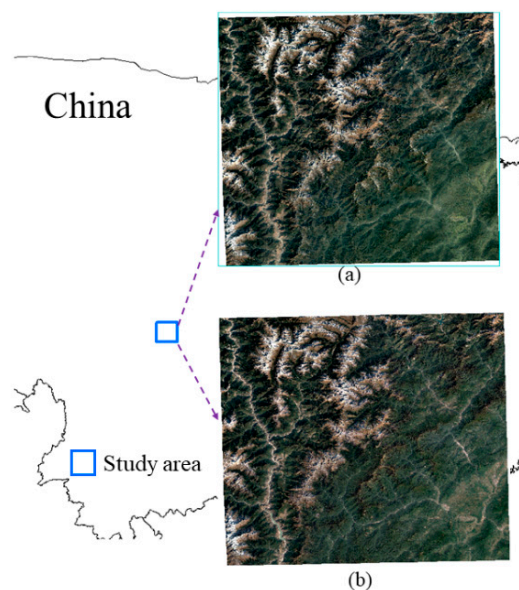

Figure 5. Study area demonstration: (**a**) image of earthquake impacted area in Lushan County, taken on 31 December 2010; (**b**) image taken on 31 December 2013.

4.2. Jiuzhaigou Earthquake-Induced Landslide

On 8 August 2017, an earthquake with a magnitude of 6.5 Mw occurred in Jiuzhaigou County. Synthesizing field investigations and visual interpretations from high spatial resolution images, 4834 earthquake-triggered landslide inventories are mapped in [48]. The images covering the Jiuzhaigou earthquake-impacted area before and after the event were downloaded with a resolution of 2.39 m. However, the resolution is still different from that used to interpret landslide inventories in [48]. Therefore, the landslide inventories were adjusted to match the images we captured from Google Earth, as well by visual interpretation. The number of landslide inventories maintained in our study images is 3817. Moreover, there is one point where the images (after the event collected from Google Earth) are largely covered by clouds, as shown in Figure 6a,b. We collected one more image taken on 27 September 2019, in Figure 6c. Some parts are updated in Figure 6c, while the majority parts are maintained (the same as in Figure 6b). Concerning the disturbing clouds, we synthesized the three images by selecting the smallest pixel intensity in the green channel to maintain soil information and remove clouds. The synthesized image is shown in Figure 6d, and is directly used for landslide detection, although there is still some remaining clouds. From Figure 6d, we can recognize that landslides are distributed intensively in the central part of the study area and sparsely along the side of road networks. The main background objects mainly comprise of road networks, forestry, vegetation, bare soil, rocks, and clouds. This complicated object distribution pattern can be used to evaluate the transferability of the proposed model.

4.3. Central Nepal Landslide

Nepal is a country prone to multiple geohazards. In 2015, Nepal went through a series of deadly earthquakes and experienced thousands of landslides [49]. As shown in Figure 7, we randomly selected one mountainous spot with numerous landslides in Central Nepal and collected the corresponding images from Google Earth with a spatial resolution of 2.39 m. The landslides are mainly distributed along the side of road networks. Ground truth landslide polygons used for training and evaluating models are visualized by two experienced experts.
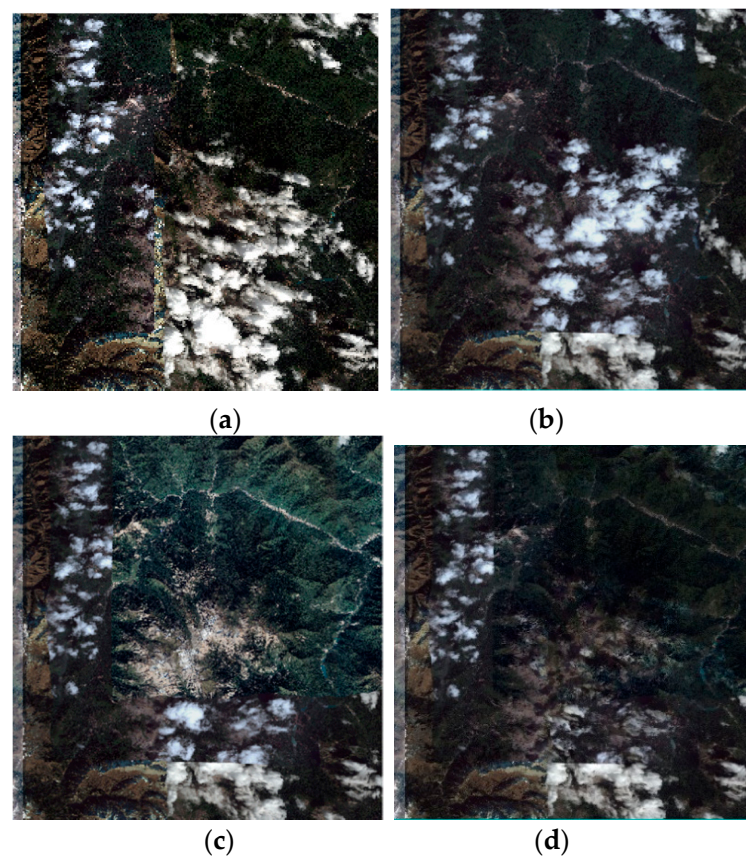
Figure 6. Post-event image demonstration collected from Google Earth: (**a**) image taken on 13 August 2017; (**b**) image taken on 14 August 2017; (**c**) image taken on 27 September 2019; (**d**) synthesized image.
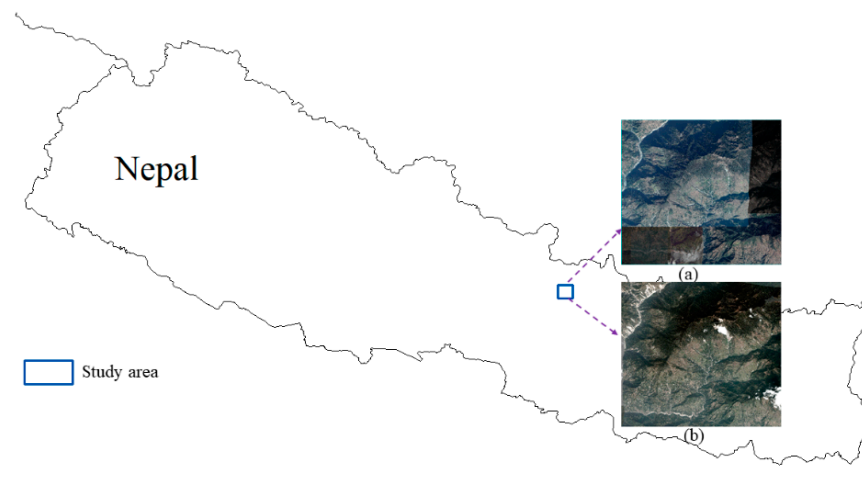
Figure 7. Experimental image in Central Nepal: (**a**) image captured on 12 December 2014; (**b**) image captured on 9 November 2015.

4.4. Southern Taiwan Landslide

Southern Taiwan is vulnerable of rainfall-induced landslides [50]. Typhoon Morakot hit Taiwan on 7 August 2009, and resulted in about 18,000 landslides. There has been considerable research (in regard to detecting landslides) for that event [51], but that study area is purely cut out from rural areas, and the background objects are generally pure vegetation. On the contrary, our study area (shown in Figure 8) is a rectangle, including oceans, an urban area with construction, and a rural area with bare soil and road networks.

The landslides are distributed intensively in the central part of the study area. That adds more difficulty in detecting landslides, but one can evaluate the model more practically and objectively. The images we collected from Google Earth have a resolution of 19 m, and the corresponding ground truth landslide polygons are visually interpreted by two experienced experts.
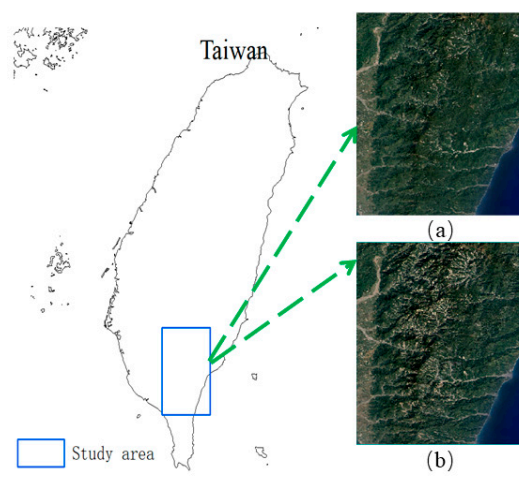

Figure 8. Study area demonstration: (**a**) image captured on 31 December 2008; (**b**) image captured on 31 December 2009.

5. Experiments and Evaluations

5.1. Experimental Settings

Our experiment was conducted on the PyTorch deep learning framework, and the proposed model was trained on three TITAN $\times$ GPUs from NVidia. Each GPU had a memory storage of 12 GB. The strategies of random scaling, random colorization, and random cropping were adopted to enlarge the data variability and enhance the generalization ability of the model. As introduced in Section 3, a potential landslide was first detected from the collected image to remove background objects (as many as possible). Based on the detected potential landslide image, four patches with a size of 512×512 pixels were generated in four directions, respectively, for each connective contour to cover more neighboring background objects, as demonstrated in Figure 9. For the cases where patch regions exceed image boundaries, they would be filled with zero intensity to maintain a size of 512×512. For cases where the potential landslide contour exceeds 512×512, the patch would take the same size of the bounding box of the corresponding connective contour. The total number of cropped patches of landslide events in Lushan is 7140, in Jiuzhaigou—19,088, in Nepal—6037, and in Taiwan—1160.
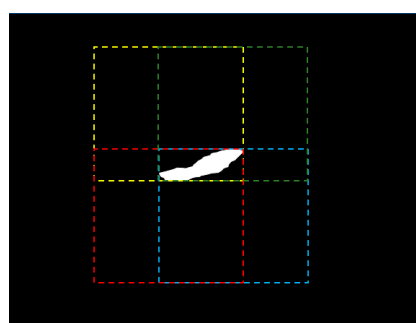

Figure 9. Patch generation of each potential landslide in four directions (labeled in each color), wherein white contour regions represent potential landslides.

The Adam (adaptive moment estimation) optimizer [52] is used to optimize our proposed framework. Following the work in [21], the initial learning rate is set to 5×10^{-5}, and has a decay rate of 10. We trained our model on 70% of the cropped patches generated in Lushan earthquake-induced landslide images and evaluated the model structure on the rest patches. In order to evaluate the transferability of the proposed network structure, we randomly selected 30% of the cropped patches from the other three events (Jiuzhaigou earthquake-induced landslide event, Nepal landslide event, and Taiwan rainfall-induced landslide event) to fine-tune the model, and we evaluated the model with the rest patches, respectively. The number of patches used for evaluation in each of the corresponding four datasets are 4998 (Lushan), 13,362 (Jiuzhaigou), 4226 (Nepal), and 812 (Taiwan), and the corresponding number of ground truth landslide inventories for each dataset is 654 (Lushan), 1286 (Jiuzhaigou), 483 (Nepal), and 62 (Taiwan).

5.2. Evaluations

To verify our proposed model, U-Net and SegNet were adopted to detect landslides on the four evaluation areas. Since U-Net and SegNet are both originally proposed with a single branch, they are designed to process input images with a single time domain. In order to conduct change detection from images with two different time domains, both U-Net and SegNet were adapted to take the same twin-tower structures before the backbone of ResNet-50, to carry on a fair performance comparison with our proposed model. Moreover, the modified U-Net and SegNet were trained on the same data, with the same strategies and loss functions. We randomly selected two test sites for each evaluation study area and demonstrated the corresponding detection results by U-Net, SegNet, and the proposed two-branch Matrix SegNet in Figure 10, respectively.

A visual comparison of Figure 10 reveals that the matrix convolution module and the squeeze-and-excitation module adopted in the two-branch Matrix SegNet work well in capturing multi-scale features and enhancing useful features from images with different spatial resolutions. Most landslides can be well detected by the two-branch Matrix SegNet in all four study areas, except for some minor landslides omitted in each image, especially in the cases of Lushan. As shown in Figure 10(c1,c2), SegNet fails to detect most landslides in the images of Lushan. U-Net omits more small landslides than our two-branch Matrix SegNet, as shown in the yellow circle in Figure 10(d1). In Jiuzhaigou, SegNet is able to detect some landslides, but omits more landslides (referring to Figure 10(c3,c4). U-Net performs much better than SegNet in detecting most landslides, nevertheless, it misclassifies more tiny landslides as background objects compared with the two-branch Matrix SegNet in Figure 10(a3–d3,a4–d4). The landslide distributions of images in Nepal and Taiwan are comparatively simpler, and the detection performances by the three methods are largely enhanced. However, omission is still an important issue in small landslide detection, especially by SegNet and U-Net. Moreover, in Taiwan (Figure 10(a7–d7,a8–d8)), commission is largely raised by SegNet and U-Net as well. Faced with complicated background objects in different spatial resolutions, the proposed two-branch Matrix SegNet behaves generally well in landslide detection by learning multi-scale features through the matrix convolution module, enhancing useful features through the squeeze-and-excitation module. That further verifies the effectiveness of our model in detecting various landslides from complicated background objects without radiometric correction.

To present a general statistical and objective evaluation on each of the four study areas, we calculated recall, precision, F1-measure, and intersection over union (IOU), according to Equations (3)–(6) for each study site, and listed the statistics in Tables 2–5, respectively. TP indicates the number of ground truth landslide pixels correctly classified to landslides. TN indicates the number of ground truth background object pixels correctly classified as background objects. FP indicates the number of ground truth background object pixels misclassified as landslides. FN indicates the number of ground truth landslide pixels misclassified as background object pixels. IOU and F1—measure are commonly used as comprehensive evaluation indexes, balancing between recall and precision.
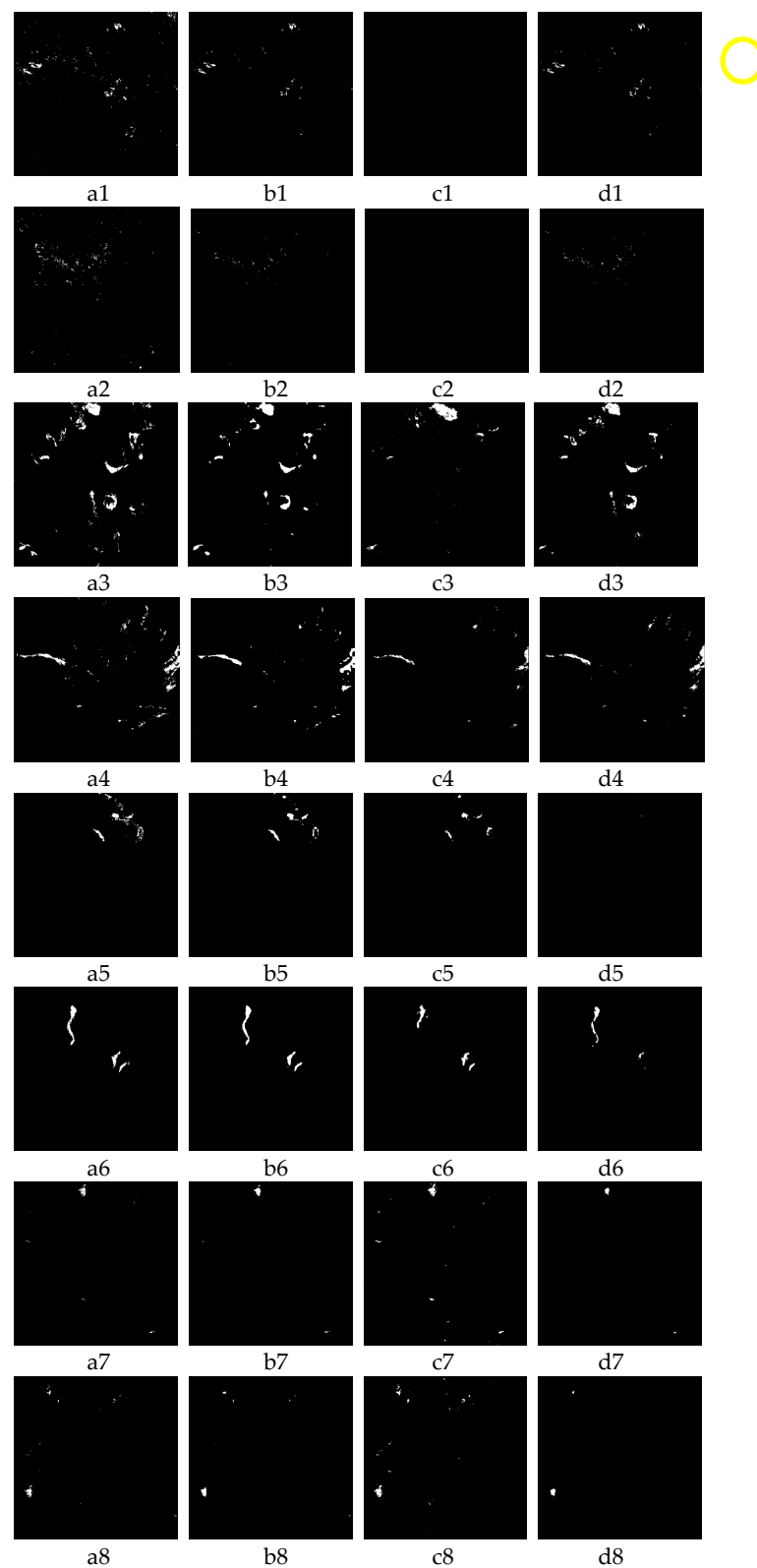
Figure 10. Landslide detection results of two test samples in each dataset: the images in column (a) indicate ground truth images of landslides, labeled as white, and background objects are labeled as black; the images in column b–d demonstrate the detection results by our two-branch Matrix SegNet, SegNet, and U-Net, respectively; indexes (*).1–(*).2 represent the test images in Lushan; indexes (*).3–(*).4 represent the test images in Jiuzhaigou; indexes (*).5–(*).6 represent the test images in Nepal and indexes (*).7–(*).8 represent the test images in Taiwan, * indicates a–d.

$$IOU = \frac{TP}{FP + FN + TP} \tag{3}$$

$$recall = \frac{TP}{TP + FN} \tag{4}$$

$$precision = \frac{TP}{TP + FP} \tag{5}$$

$$F1\text{-measure} = \frac{2 \times precision \times recall}{precision + recall} \tag{6}$$

Table 2. Evaluation statistics (%) in landslide detection from Lushan study area.

	Recall	Precision	F1-measure	IOU
U-Net	20.60	**96.85**	33.97	20.46
SegNet	0.0487	75.12	0.0973	0.0487
Proposed network	**22.33**	73.49	**34.26**	**20.67**

Table 3. Evaluation statistics (%) in landslide detection from Jiuzhaigou study area.

	Recall	Precision	F1-measure	IOU
U-Net	61.03	**89.86**	72.69	57.10
SegNet	47.60	74.29	58.03	40.87
Proposed network	**70.46**	82.75	**76.11**	**61.44**

Table 4. Evaluation statistics (%) in landslide detection from Nepal study area.

	Recall	Precision	F1-measure	IOU
U-Net	31.37	73.40	43.96	28.17
SegNet	62.29	72.76	67.12	50.51
Proposed network	**68.43**	**78.30**	**73.04**	**57.53**

Table 5. Evaluation statistics (%) in landslide detection from Taiwan study area.

	Recall	Precision	F1-measure	IOU
U-Net	28.30	**81.72**	42.04	26.61
SegNet	**90.97**	43.15	58.54	41.38
Proposed network	49.39	72.59	**58.79**	**41.63**

It is apparent that the proposed two-branch Matrix SegNet has the best general performance with the highest F1-measure and IOU among the three methods in detecting landslides for the four study areas. U-Net and SegNet both gain low recall and high precision in Lushan datasets. Synthesizing images in Figure 10, we can see that U-Net and SegNet are more sensitive to small landslides. They are more likely to misclassify small landslide areas as background objects. We should note that the spatial resolution of Lushan and Taiwan datasets is lower than that of Nepal and Jiuzhaigou datasets. As shown in Figure 10, the landslides in Lushan dataset, especially, are distributed more intensively with a smaller size. It indicates that small landslides can easily be excluded with continuous convolution operations in the encoding parts of both U-Net and SegNet structures, since they both take the typical encoding–decoding network structures. Moreover, comparing landslide distribution of different datasets from ground truth images in Figure 10, we can

recognize that landslide distribution is most complicated in Lushan dataset, with the most intensive distribution of the smallest landslides. Therefore, F1-measure and IOU of U-Net, and our two-branch Matrix SegNet, are no higher than 35%. However, our proposed Matrix SegNet still has a slightly higher accuracy, by 0.2%. In terms of the Taiwan dataset, which has the same spatial resolution with the Lushan dataset, U-Net still omits small landslides (as shown in Figure 10(c7–d7), while SegNet misclassifies many small background areas as landslides (as shown in Figure 10(c8–d8). It further validates the high sensitivity of a typical encoder–decoder network architecture in detecting small landslides from high spatial resolution images.

The performances of the three methods are apparently better in Nepal and Jiuzhaigou datasets than the other two datasets with at least 15% higher F1-measure and IOU. It indicates that more details in higher spatial resolution images are better at providing more detailed ground object information for landslide detection, especially in cases of small landslides. Our proposed Matrix SegNet can achieve up to 20% progress in detection accuracy in the datasets of Jiuzhaigou and Nepal.

6. Discussions

Synthesizing the evaluation performances of U-Net, SegNet, and our proposed two-branch Matrix SegNet on the test images of four different study areas, our model has better transferability than U-Net and SegNet. Faced with evaluation images with different spatial resolutions, the parameter IOU of our proposed two-branch Matrix SegNet is at least 0.21% higher than U-Net and SegNet in Lushan and Taiwan datasets with a spatial resolution of 19 m, and over 7% progress than U-Net and SegNet in the Nepal and Jiuzhaigou datasets with a spatial resolution of 2.35 m. Similar with parameter IOU, the comprehensive evaluation parameter F1-measure also witnessed similar improvement by our proposed two-branch Matrix SegNet with almost 0.2% progress in Lushan and Taiwan datasets and 7% progress in Nepal and Jiuzhaigou datasets. The performance of our model improves more on images having a spatial resolution of 2.35 m than those with a spatial resolution of 19 m. This mostly attributes to the spectral and textural details demonstrated of landslides in images with higher spatial resolution. The matrix learning module used in our model is more applicable for multiple landslide detection from high spatial resolution images by capturing landslide features horizontally, vertically, and diagonally. However, the traditional layer-wise convolution is likely to omit multi-scale features of landslides by U-Net and SegNet.

From the evaluation statistics, we recognize that the precision and recall of our proposed two-branch Matrix SegNet are more balanced than those of U-Net and SegNet in all cases of the evaluation datasets. One possible driving factor of this phenomenon is the high omission of small landslides by U-Net and SegNet. They can mostly be well-detected by our Matrix SegNet. However, regardless of the improvement, the recall by our proposed model is smaller than 71% in the evaluation cases. The disturbing bare soil, as the background objects of landslides, is an issue confronted in our model as well. This may be overcome by modifying the network structure to focus more on enhancing the shape discrimination ability between landslides and background objects in future work.

We should note that the accuracy gained by our experiments is lower than that obtained in [5], mainly because of the large differences of experimental images. Images in [5] were mostly captured from areas with pure vegetation and there was high contrast between landslides and background objects in spectral characteristics. However, images in our experiment were mostly taken with various imaging radiances from different study areas, and the ground objects were distributed complicatedly to mimic practical applications.

7. Conclusions

This study proposed a practical trial of landslide detection by adopting raw pre- and post-event images captured from Google Earth directly. In our proposed two-branch Matrix

SegNet, the matrix convolution module was adopted to learn landslide features with multiple aspect ratios and scales, horizontally and vertically. Landslide features were further enhanced by incorporating the squeeze-and-excitation modules. The proposed model structure was applied to detect landslides from images with different spatial resolutions from four different study areas with various ground object distribution patterns. Two widely used semantic segmentation frameworks, U-Net and SegNet, were adopted as comparisons with the proposed model. The statistical evaluations verify the efficiency of our proposed two-branch Matrix SegNet in detecting multiple landslides. It shows a larger improvement in landslide detection in F1-measure and the IOU of datasets, with a spatial resolution of 2.39 m over the performances by U-Net and SegNet, than that in datasets with a spatial resolution of 19 m over U-Net and SegNet. The matrix convolution module has a stronger ability to capture multi-scale landslide features. However, a high landslide omission rate is an issue in our proposed model, especially in cases with small landslides in complicated background objects. In future work, we will work on modifying the model structures to enhance the distinguishing ability of shape characteristics of landslides and bare soils, to reduce the omission error. Our proposed framework switches on a path to develop reliable and applicable methods to detect landslides from large-scale research areas with complicated background objects.

Author Contributions: Conceptualization, B.Y. and F.C.; methodology, B.Y.; software, C.X.; validation, N.W. and C.X.; formal analysis, B.Y.; investigation, L.W.; resources, B.Y.; data curation, L.W.; writing—original draft preparation, F.C.; writing—review and editing, B.Y.; visualization, L.W.; supervision, F.C.; project administration, C.X.; funding acquisition, F.C. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: The images are available at https://pan.baidu.com/s/1uurwlYtcika9 InSemQqU9w (accessed on 29 July 2021), with a code of 3r5q.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Haque, U.; da Silva, P.F.; Devoli, G.; Pilz, J.; Zhao, B.; Khaloua, A.; Wilopo, W.; Andersen, P.; Lu, P.; Lee, J. The human cost of global warming: Deadly landslides and their triggers (1995–2014). *Sci. Total. Environ.* **2019**, *682*, 673–684. [CrossRef] [PubMed]
2. Gariano, S.L.; Guzzetti, F. Landslides in a changing climate. *Earth Sci. Rev.* **2016**, *162*, 227–252. [CrossRef]
3. Deijns, A.A.J.; Bevington, A.R.; van Zadelhoff, F.; de Jong, S.M.; Geertsema, M.; McDougall, S. Semi-automated detection of landslide timing using harmonic modelling of satellite imagery, buckinghorse river, canada. *Int. J. Appl. Earth Obs. Geoinf.* **2020**, *84*, 101943. [CrossRef]
4. Lu, P.; Bai, S.; Tofani, V.; Casagli, N. Landslides detection through optimized hot spot analysis on persistent scatterers and distributed scatterers. *ISPRS J. Photogramm. Remote. Sens.* **2019**, *156*, 147–159. [CrossRef]
5. Lu, P.; Qin, Y.; Li, Z.; Mondini, A.C.; Casagli, N. Landslide mapping from multi-sensor data through improved change detection-based markov random field. *Remote. Sens. Environ.* **2019**, *231*, 111235. [CrossRef]
6. Wood, J.L.; Harrison, S.; Reinhardt, L.; Taylor, F.E. Landslide databases for climate change detection and attribution. *Geomorphology* **2020**, *355*, 107061. [CrossRef]
7. Yu, B.; Chen, F.; Muhammad, S.; Li, B.; Wang, L.; Wu, M. A simple but effective landslide detection method based on image saliency. *Photogramm. Eng. Remote. Sens.* **2017**, *83*, 351–363. [CrossRef]
8. Yu, B.; Chen, F. A new technique for landslide mapping from a large-scale remote sensed image: A case study of central nepal. *Comput. Geosci.* **2017**, *100*, 115–124. [CrossRef]
9. Li, Z.; Shi, W.; Lu, P.; Yan, L.; Wang, Q.; Miao, Z. Landslide mapping from aerial photographs using change detection-based markov random field. *Remote. Sens. Environ.* **2016**, *187*, 76–90. [CrossRef]
10. Aksoy, B.; Ercanoglu, M. Landslide identification and classification by object-based image analysis and fuzzy logic: An example from the azdavay region (kastamonu, turkey). *Comput. Geosci.* **2012**, *38*, 87–98. [CrossRef]
11. Moosavi, V.; Talebi, A.; Shirmohammadi, B. Producing a landslide inventory map using pixel-based and object-oriented approaches optimized by taguchi method. *Geomorphology* **2014**, *204*, 646–656. [CrossRef]
12. Pradhan, B.; Jebur, M.N.; Shafri, H.Z.M.; Tehrany, M.S. Data fusion technique using wavelet transform and taguchi methods for automatic landslide detection from airborne laser scanning data and quickbird satellite imagery. *IEEE Trans. Geosci. Remote. Sens.* **2016**, *54*, 1610–1622. [CrossRef]

13.   Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

14.   Chen, F.; Yu, B.; Li, B. A practical trial of landslide detection from single-temporal landsat8 images using contour-based proposals and random forest: A case study of national nepal. *Landslides* **2017**, *2018*, 453–464. [CrossRef]

15.   Cheng, Y.S.; Yu, T.T. Potential landslide detection with fractal and roughness by lidar data in taiwan. In Proceedings of the EGU General Assembly Conference Abstracts, Vienna, Austria, 12–17 April 2015.

16.   Youssef, A.M.; Pourghasemi, H.R.; Pourtaghi, Z.S.; Al-Katheeri, M.M. Landslide susceptibility mapping using random forest, boosted regression tree, classification and regression tree, and general linear models and comparison of their performance at wadi tayyah basin, asir region, saudi arabia. *Landslides* **2016**, *13*, 839–856. [CrossRef]

17.   Aimaiti, Y.; Liu, W.; Yamazaki, F.; Maruyama, Y. Earthquake-induced landslide mapping for the 2018 hokkaido eastern iburi earthquake using palsar-2 data. *Remote. Sens.* **2019**, *11*, 2351. [CrossRef]

18.   Mustafa, M.; Biswajeet, P.; Hossein, R. Improving landslide detection from airborne laser scanning data using optimized dempster–shafer. *Remote. Sens.* **2018**, *10*, 1029.

19.   Tran, C.J.; Mora, O.E.; Fayne, J.V.; Lenzano, M.G. Unsupervised classification for landslide detection from airborne laser scanning. *Geosciences* **2019**, *9*, 221. [CrossRef]

20.   Li, Z.; Shi, W.; Myint, S.W.; Lu, P.; Wang, Q. Semi-automated landslide inventory mapping from bitemporal aerial photographs using change detection and level set method. *Remote. Sens. Environ.* **2016**, *175*, 215–230. [CrossRef]

21.   Rashwan, A.; Kalra, A.; Poupart, P. Matrix nets: A new deep architecture for object detection. In Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), Seoul, Korea, 27–28 October 2019; pp. 2025–2028.

22.   Xie, E.; Sun, P.; Song, X.; Wang, W.; Liu, X.; Liang, D.; Shen, C.; Luo, P. Polarmask: Single shot instance segmentation with polar representation. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*; IEEE: Seattle, WA, USA, 2019.

23.   Ghorbanzadeh, O.; Blaschke, T.; Gholamnia, K.; Meena, S.; Tiede, D.; Aryal, J. Evaluation of different machine learning methods and deep-learning convolutional neural networks for landslide detection. *Remote Sens.* **2019**, *11*, 196. [CrossRef]

24.   Ye, C.; Li, Y.; Cui, P.; Liang, L.; Pirasteh, S.; Marcato, J.; Gonçalves, W.N.; Li, J. Landslide detection of hyperspectral remote sensing data based on deep learning with constrains. *IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens.* **2019**, *12*, 5047–5060. [CrossRef]

25.   Ding, A.; Zhang, Q.; Zhou, X.; Dai, B. Automatic recognition of landslide based on cnn and texture change detection. In Proceedings of the 2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC), Wuhan, China, 11–13 November 2016; pp. 444–448.

26.   Yu, H.; Ma, Y.; Wang, L.; Zhai, Y.; Wang, X. A landslide intelligent detection method based on cnn and rsg_r. In Proceedings of the 2017 IEEE International Conference on Mechatronics and Automation (ICMA), Takamatsu, Japan, 6–9 August 2017; pp. 40–44.

27.   Yu, B.; Chen, F.; Xu, C. Landslide detection based on contour-based deep learning framework in case of national scale of nepal in 2015. *Comput. Geosci.* **2020**, *135*, 104388. [CrossRef]

28.   Zhao, H.; Shi, J.; Qi, X.; Wang, X.; Jia, J. Pyramid scene parsing network. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*; IEEE: Honolulu, HI, USA, 2016.

29.   Ronneberger, O.; Fischer, P.; Brox, T. *U-Net: Convolutional Networks for Biomedical Image Segmentation*; Springer International Publishing: Cham, Switzerland, 2015; pp. 234–241.

30.   Wang, N.; Chen, F.; Yu, B.; Qin, Y. Segmentation of large-scale remotely sensed images on a spark platform: A strategy for handling massive image tiles with the mapreduce model. *ISPRS J. Photogramm. Remote. Sens.* **2020**, *162*, 137–147. [CrossRef]

31.   Melekhov, I.; Tiulpin, A.; Sattler, T.; Pollefeys, M.; Rahtu, E.; Kannala, J. Dgc-net: Dense geometric correspondence network. In *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*; IEEE: Waikoloa, HI, USA, 2019; pp. 1034–1042.

32.   Wolf, D.; Prankl, J.; Vincze, M. Enhancing semantic segmentation for robotics: The power of 3d entangled forests. *IEEE Robot. Autom. Lett.* **2016**, *1*, 49–56. [CrossRef]

33.   Hu, J.; Shen, L.; Albanie, S.; Sun, G.; Wu, E. Squeeze-and-excitation networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2019**, *42*, 2011–2023. [CrossRef] [PubMed]

34.   Yu, B.; Yang, L.; Chen, F. Semantic segmentation for high spatial resolution remote sensing images based on convolution neural network and pyramid pooling module. *IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens.* **2018**, *11*, 3252–3261. [CrossRef]

35.   He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask r-cnn. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *42*, 386–397. [CrossRef] [PubMed]

36.   Liu, S.; Qi, L.; Qin, H.; Shi, J.; Jia, J. Path aggregation network for instance segmentation. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*; IEEE: Salt Lake City, UT, USA, 2018.

37.   Badrinarayanan, V.; Kendall, A.; Cipolla, R. Segnet: A deep convolutional encoder-decoder architecture for scene segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 2481–2495. [CrossRef] [PubMed]

38.   Chen, L.C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE Trans. Pattern Anal. Mach. Intell.* **2016**, *40*, 834–848. [CrossRef]

39.   Chen, L.C.; Papandreou, G.; Schroff, F.; Adam, H. Rethinking atrous convolution for semantic image segmentation. *arXiv* **2017**, arXiv:1706.05587.

40.   Prakash, N.; Manconi, A.; Loew, S. Mapping landslides on eo data: Performance of deep learning models vs. Traditional machine learning models. *Remote Sens.* **2020**, *12*, 346. [CrossRef]

41.   Zhao, X.; Yuan, Y.; Song, M.; Ding, Y.; Lin, F.; Liang, D.; Zhang, D. Use of unmanned aerial vehicle imagery and deep learning unet to extract rice lodging. *Sensors* **2019**, *19*, 3859. [CrossRef] [PubMed]

42. Lin, T.Y.; Dollár, P.; Girshick, R.; He, K.; Hariharan, B.; Belongie, S. Feature pyramid networks for object detection. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*; IEEE: Honolulu, HI, USA, 2016.

43. Deng, Y.; Wu, C.; Li, M.; Chen, R. Rndsi: A ratio normalized difference soil index for remote sensing of urban/suburban environments. *Int. J. Appl. Earth Obs. Geoinf.* **2015**, *39*, 40–48. [CrossRef]

44. Xu, J.Z.; Lu, W.; Li, Z.; Khaitan, P.; Zaytseva, V. Building damage detection in satellite imagery using convolutional neural networks. *arXiv* **2019**, arXiv:1910.06444.

45. He, K.; Zhang, X.; Ren, S.; Jian, S. *Deep Residual Learning for Image Recognition*; IEEE: Las Vegas, NV, USA, 2016.

46. Lin, T.; Goyal, P.; Girshick, R.; He, K.; Dollár, P. Focal loss for dense object detection. In Proceedings of the IEEE international conference on computer vision, Venice, Italy, 22–29 October 2017; pp. 2980–2988.

47. Lina, L.; Chong, X.; Jian, C. Landslide factor sensitivity analysis for landslides triggered by 2013 lushan earthquake using gis platform and certainty factor method. *J. Eng. Geol.* **2014**, *22*, 11.

48. Tian, Y.; Xu, C.; Ma, S.; Xu, X.; Wang, S.; Zhang, H. Inventory and spatial distribution of landslides triggered by the 8th august 2017 mw 6.5 jiuzhaigou earthquake, china. *J. Earth Sci.* **2019**, *30*, 206–217. [CrossRef]

49. Kargel, J.; Leonard, G.; Shugar, D.; Haritashya, U.; Bevington, A.; Fielding, E.; Fujita, K.; Geertsema, M.; Miles, E.; Steiner, J.; et al. Geomorphic and geologic controls of geohazards induced by nepal's 2015 gorkha earthquake. *Science* **2016**, *351*. [CrossRef] [PubMed]

50. Mondini, A.C.; Marchesini, I.; Rossi, M.; Chang, K.-T.; Pasquariello, G.; Guzzetti, F. Bayesian framework for mapping and classifying shallow landslides exploiting remote sensing and topographic data. *Geomorphology* **2013**, *201*, 135–147. [CrossRef]

51. Lin, C.W.; Chang, W.S.; Liu, S.H.; Tsai, T.T.; Lee, S.P.; Tsang, Y.C.; Shieh, C.L.; Tseng, C.M. Landslides triggered by the 7 august 2009 typhoon morakot in southern taiwan. *Eng. Geol.* **2011**, *123*, 3–12. [CrossRef]

52. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.