



Article

Landslide Extraction Using Mask R-CNN with Background-Enhancement Method

Ruilin Yang ¹, Feng Zhang ^{1,2,*} , Junshi Xia ³ and Chuyi Wu ¹

¹ School of Earth Sciences, Zhejiang University, Hangzhou 310027, China; yangruilin@zju.edu.cn (R.Y.); wuchuyi@zju.edu.cn (C.W.)

² Zhejiang Provincial Key Laboratory of Geographic Information Science, Hangzhou 310028, China

³ Geoinformatics Unit, RIKEN Center for Advanced Intelligence Project, Tokyo 103-0027, Japan; junshi.xia@riken.jp

* Correspondence: zfcarnation@zju.edu.cn; Tel.: +86-571-8827-3287

Abstract: The application of deep learning methods has brought improvements to the accuracy and automation of landslide extractions based on remote sensing images because deep learning techniques have independent feature learning and powerful computing ability. However, in application, the quality of training samples often fails the requirement for training deep networks, causing insufficient feature learning. Furthermore, some background objects (e.g., river, bare land, building) share similar shapes, colors, and textures with landslides. They can be confusing to automatic tasks, contributing false and missed extractions. To solve the above problems, a background-enhancement method was proposed to enrich the complexity of samples. Models can learn the differences between landslides and background objects more efficiently through background-enhanced samples, then reduce false extractions on background objects. Considering that the environments of disaster areas play dominant roles in the formation of landslides, landslide-inducing attributes (DEM, slope, distance from river) were used as supplements, providing additional information for landslide extraction models to further improve the accuracy of extraction results. The proposed methods were applied to extract landslides that occurred in Ludian county, Yunnan Province, in August 2014. Comparative experiments were conducted using a mask R-CNN model. The experiment using both background-enhanced samples and landslide-inducing information showed a satisfying result with an F1 score of 89.08%. Compared with the F1 score from the experiment using only satellite images as input data, it was significantly improved by 22.38%, underscoring the applicability and effectiveness of our background-enhancement method.

Keywords: landslide extraction; background enhancement; deep learning; landslide-inducing factors; Mask R-CNN



Citation: Yang, R.; Zhang, F.; Xia, J.; Wu, C. Landslide Extraction Using Mask R-CNN with Background-Enhancement Method. *Remote Sens.* **2022**, *14*, 2206. <https://doi.org/10.3390/rs14092206>

Academic Editor: Sandro Moretti

Received: 23 March 2022

Accepted: 2 May 2022

Published: 5 May 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Landslides occur frequently worldwide, and they often result in severe casualties and property damage, causing huge loss to transportation systems, industry, and agricultural production. In addition, landslides may trigger emergencies if they occur close to human settlements [1]. An accurate and quick detection of a landslide can offer a landslide's location and the scope of its reach, then assist in releasing evacuation information, formulating an emergency rescue plan, and assessing disaster loss.

The conventional access to landslide information is based on field surveys. Since landslides often happen in mountainous areas with complex terrain, field surveys can be very time-consuming and dangerous [2]. Nowadays, satellite images with high spatial resolution can be obtained from many platforms, and this has boosted the broad application of landslide detections based on satellite images. A traditional visual interpretation method can provide high accuracy, but it highly relies on expert knowledge and heavy artificial

work [3]. There are also methods such as pixel-based and object-oriented extraction methods [4]. Both of them have been integrated with different machine learning models (e.g., random forests, support vector machine) to improve the efficiency of landslide extraction tasks and have certain achievements in application [5–8]. However, the pixel-based method often has tattered and discontinuous extraction results. The object-based method requires artificial adjustment on features and parameters and has a resulting model with low portability.

In recent studies, deep learning has great performance in tasks such as classification, object detection, instance segmentation, and so on [9]. Compared with the above methods, deep learning methods can independently learn features from input data, thus reducing subjective impact brought about by artificial feature extractions, and can also improve efficiency [10]. Under this trend, landslide extractions have also been shifted from conventional methods to more automatic methods based on deep learning. Romeo et al. [11] used CNNs for landslide detections and proved that networks with deeper structures have stronger feature learning ability. Nava et al. [12] was the first to combine CNNs and SAR data for landslide detection and has achieved almost as much accuracy as with landslide detections based on optical satellite images. Ju et al. [13] proved the applicability of deep learning methods for old loess landslide detection tasks by experiments in Northwest China.

Based on these studies, we can see deep learning as a reliable tool to address landslide extraction tasks using remote sensing images of various types, and it has exceeded other traditional approaches for high efficiency and automation [1,14]. However, landslides show great diversity in shapes, textures, colors, and sizes, meanwhile some background objects (e.g., river, bare land, buildings, roads) share similar features with landslides [13,15]. These can be disturbances for landslide extraction models. In application, landslide extractions often have training samples with unsatisfying quality, where deep neural networks cannot be trained properly, causing insufficient feature learning [15,16]. The false extractions on background objects are common contributors to low extraction accuracy [17]. An effective way to solve the above problems is to extract precise, expressive features from training data. In the field of landslide extractions, some efforts have already been made to improve the feature learning step of deep learning methods. These studies can be roughly divided into two types: optimizations on deep learning model structures and data enhancement on training samples.

(1) Optimizations on deep learning model structures

Modifications on network structures are common ways to improve feature extraction and increase accuracy. Qi et al. [18] built ResU-Net by combining U-Net and residual modules, and conducted experiments in Tianshui, China to validate the model's effectiveness. Yi et al. [16] further strengthened feature learning, and built a cascaded end-to-end LandsNet, whose F1 score was improved by 7% compared with ResU-Net. Liu et al. [10] replaced the feature extraction layer of the Mask R-CNN model with ResNext network and added edge loss function to improve extraction results of landslide boundaries and tiny landslides. Aiming to make models more concentrated on important information, some researchers introduced the attention mechanism into landslide extraction studies to achieve better results. For example, a 3D attention mechanism [19] was proposed by Ji et al. for landslide detection and was combined with various backbones to prove its feasibility and robustness. To avoid false extractions on background objects, Zhu et al. [17] fused local and non-local features to preserve contextual information, meanwhile adding a scale attention module into the U-Net model. Based on these strategies, the F1 score was improved by 14.62%. Cheng et al. [20] reconstructed a YOLO-SA model based on YOLOv4 with an attentional mechanism and fewer parameters to improve landslide detection accuracy while maintaining the detection speed.

(2) Data enhancement and supplement for training samples

Training samples are the basis of deep learning and have a dominant impact on landslide extraction accuracy. Some simple data enhancement methods, such as rotating and flipping, are already applied in many studies [10,20–22], but excessive rotations or flips on single samples may cause overfitting [21]. In comparison, data enhancement using multiple samples can provide more useful information [23]. Jiang et al. [15] made simulated hard samples by utilizing background objects which share similar features with landslides, and the improvements on false extractions were proved by the extraction results of Bijie City and Tianshui City in China. Apart from data enhancement, many researchers also used geoscience data as supplement for remote sensing images. For example, Ghorbanzadeh et al. [24] proved that using slope gradient data as an additional input layer can help models distinguish landslides and background objects with similar spectral characteristics. Generally speaking, the proportion of landslides in a large-scale image is relatively small compared with complex background. Given this phenomenon, Yu et al. [4] based on NDVI (vegetation index) and DEM (digital elevation model), removed background objects in a large extent and obtained potential landslide areas for further accurate semantic segmentation. Liu et al. [21] combined DSM, slope, aspect, and optical satellite images as a six-channel input to better extract landslides using the U-Net model, and the F1 score was improved by 4.13% based on additional spatial information.

Both optimizations on network structures or training data can improve the accuracy of landslide extractions to a certain extent. By contrast, enhancing training samples and adding additional information are optimizations on data level, which are applicable for deep learning models with different structures theoretically. How to provide more useful information through modifications on training samples to improve the feature learning step of deep learning methods is a direction worth working on. Aiming for that, we developed a background-enhancement method which can provide comparisons between landslides and confusing background objects for model training. Furthermore, landslide-inducing factors have already been integrated with various deep learning methods to analyze landslide susceptibility and have certain achievements [25–27]; hence it can be considered that landslide-inducing factors can provide valid landslide-related information for deep learning models. Therefore, this study also uses landslide-inducing factors as additional layers to provide auxiliary information. Comparative experiments with and without background-enhancement operations are made based on Mask R-CNN [28] model and post-landslide satellite images of Ludian in 2014 to validate the effectiveness of our proposed methods.

This paper is organized mainly in five sections. The reviews of recent research and the main objective of our study are in Section 1. Section 2 gives basic information of the study area and prepares the dataset for experiments. Section 3 introduces the Mask R-CNN model, meanwhile specifies our background-enhancement method and the use of landslide-inducing information. Section 4 shows the extraction results and discussions of the comparative experiments. Finally, Section 5 concludes our work and the direction of future research.

The main contributions of this study are as follows:

- (1) Developing a background-enhancement method based on image splicing and a modified CutMix [29]. While increasing the amount of training data, the backgrounds of the samples are more complex, assisting models to distinguish landslide and background objects.
- (2) Adding landslide-inducing topographic factors (DEM, slope, distance from river) into input training data as auxiliary information. Using landslide formation mechanism as a reference for landslide extractions.
- (3) Evaluating the applicability and effectiveness of our proposed methods by comparative experiments using Mask R-CNN and Ludian landslide data.

2. Study Area and Data

2.1. Study Area

Ludian county of Zhaotong city, Yunnan Province, lies in the north the of Yunnan–Guizhou Plateau. The location of Ludian is shown in Figure 1a. Its special geological structures and complex terrains have resulted in frequent geological disasters, including earthquakes, landslides, and debris flows. On 3 August 2014, an earthquake occurred in Ludian county [30]. Although the magnitude was only Ms 6.5, the earthquake triggered massive secondary disasters such as landslides and collapses, causing further casualties and property damage [31]. According to a newly updated inventory, the earthquake has triggered more than 10,000 landslides [32], even causing large-scale barrier lakes [33].

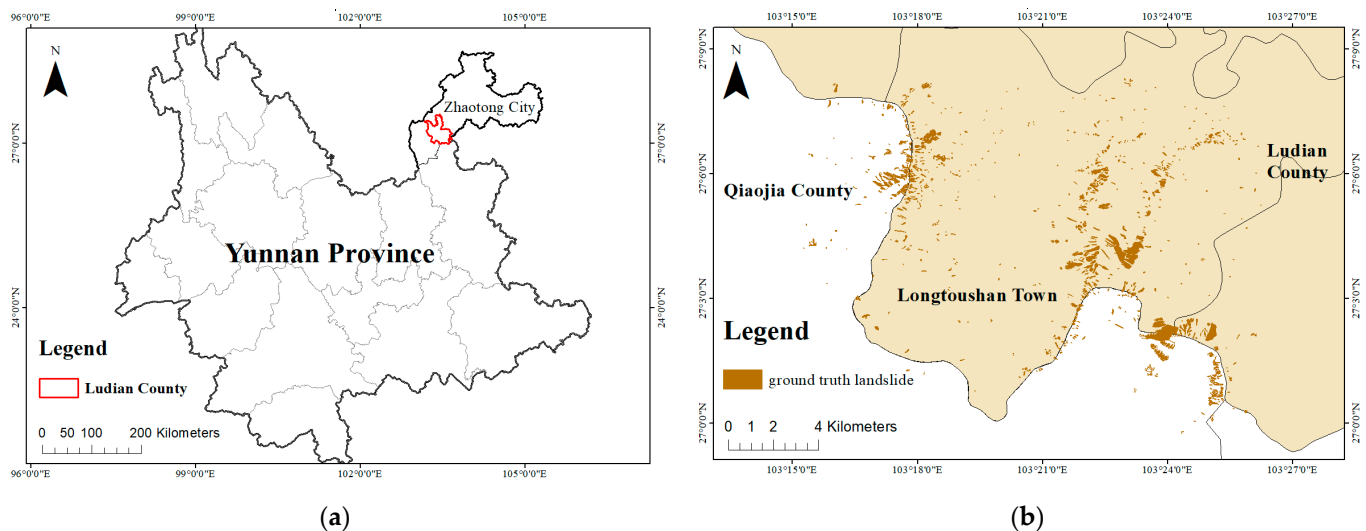


Figure 1. Information of the study area. (a) The location of Ludian county in Yunnan Province; (b) the distribution of landslides in Longtoushan Town, Ludian county.

In this study, the proposed methods combined with Mask R-CNN model are designed to detect landslides that happened in Ludian County. The distribution of Ludian landslides is shown in Figure 1b [31]. The landslides mainly occurred in Longtoushan town of Ludian county, and along the border of Qiaojia county and Ludian county.

2.2. Data Preparation

Considering the image quality and the density of landslides, we chose two areas with blue bounding boxes shown in Figure 2 as training data, and the area with a red bounding box shown in Figure 2 was chosen as test data. The training areas are 43.71 km² in total and have about 400 landslides, while the test area is 21.49 km² and has about 238 landslides.

Post-landslide images used for model training and testing are obtained from the Google Earth platform, taken on 20 August 2014. These images were taken only 17 days after the earthquake, and they also served for other previous studies including visual interpretations [32] and landslide detections [20], thus presumably they can provide sufficient landslide information. Pre-landslide images are also from the Google Earth platform, taken on 30 January 2014. Pre-landslide images are not involved in landslide extraction tasks in this study; they only provide references when labeling landslides. More details of pre- and post-disaster satellite images are shown in Table 1.

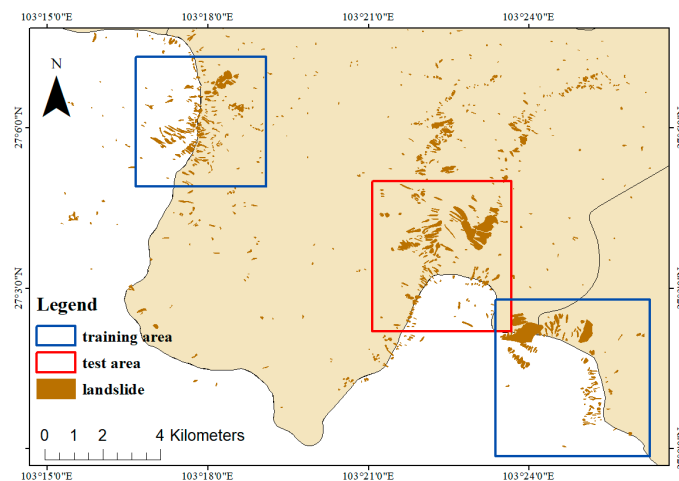


Figure 2. The distribution of test and training area.

Table 1. Detailed information of satellite images.

	Resolution	Source	Satellite	Collected Date
Pre-landslide image	0.27m/pixel	CNES/Airbus	Pleiades PHR1A	30 January 2014
Post-landslide image	0.27m/pixel	Maxar Technologies	GeoEye-01	20 August 2014

The landslide-inducing attributes used in this study are DEM, slope gradient, and distance from river. The DEM data of the study area are from Google Earth platform, with sampling interval of about 8.0 m. Slope data are projected and calculated from the DEM data using ArcGIS 10.2. Both DEM and slope data have been interpolated and resampled so that each pixel in the optical satellite images has its corresponding topographic value. The river data are obtained from Open Street Map, and we made buffers to separate the study area into six levels. The reasons for choosing these topographic data and their relations with landslide distribution will be discussed in Section 3.3.

Due to running memory limits, the large-scale optical satellite images must be split into smaller units. Smaller images can speed up the training process, but it is not always the smaller the better [34]. The size of landslides in Ludian varies widely, as shown in Figure 3a. If the splitting unit is too small, a large-sized landslide will be cut into many blocks, and the features of landslide or background cannot be extracted properly. Considering the above factors, we chose a size of 1024×1024 pixels to be cut out from the original images through the regular grid sampling method. As an example, the splitting results of the test image are shown in Figure 3.

Labels with great quality can be very helpful for model training and result analysis. For lack of field investigations and aerial photographs, this study uses labels made in previous study, and makes revisions based on 3D Google Earth images to ensure the accuracy. From 2014 till now, many researchers have already made investigations on landslides triggered by the Ludian earthquake. The public dataset created in 2014 by Xu et al. [31] (available on <https://www.sciencebase.gov/catalog/item/594428d4e4b062508e32319f>, accessed on 6 January 2022) was made based on visual interpretations using pre- and post-landslide satellite images, aerial photographs and field investigation data. This public dataset has already been cited in many similar studies [32,35,36] for references and comparisons. Thus, it is a reliable dataset providing landslide locations and distributions. Wu et al. [32] pointed out that labels created by Xu et al. are a bit rough due to satellite image resolution, and labels created by Xu et al. contain only earthquake-triggered landslides that old landslides are not included. To improve the labels to be more suitable for this study, some minor revisions are made based on landslide locations provided by Xu et al., 3D Google Earth images, and high-resolution optical satellite images of affected area.

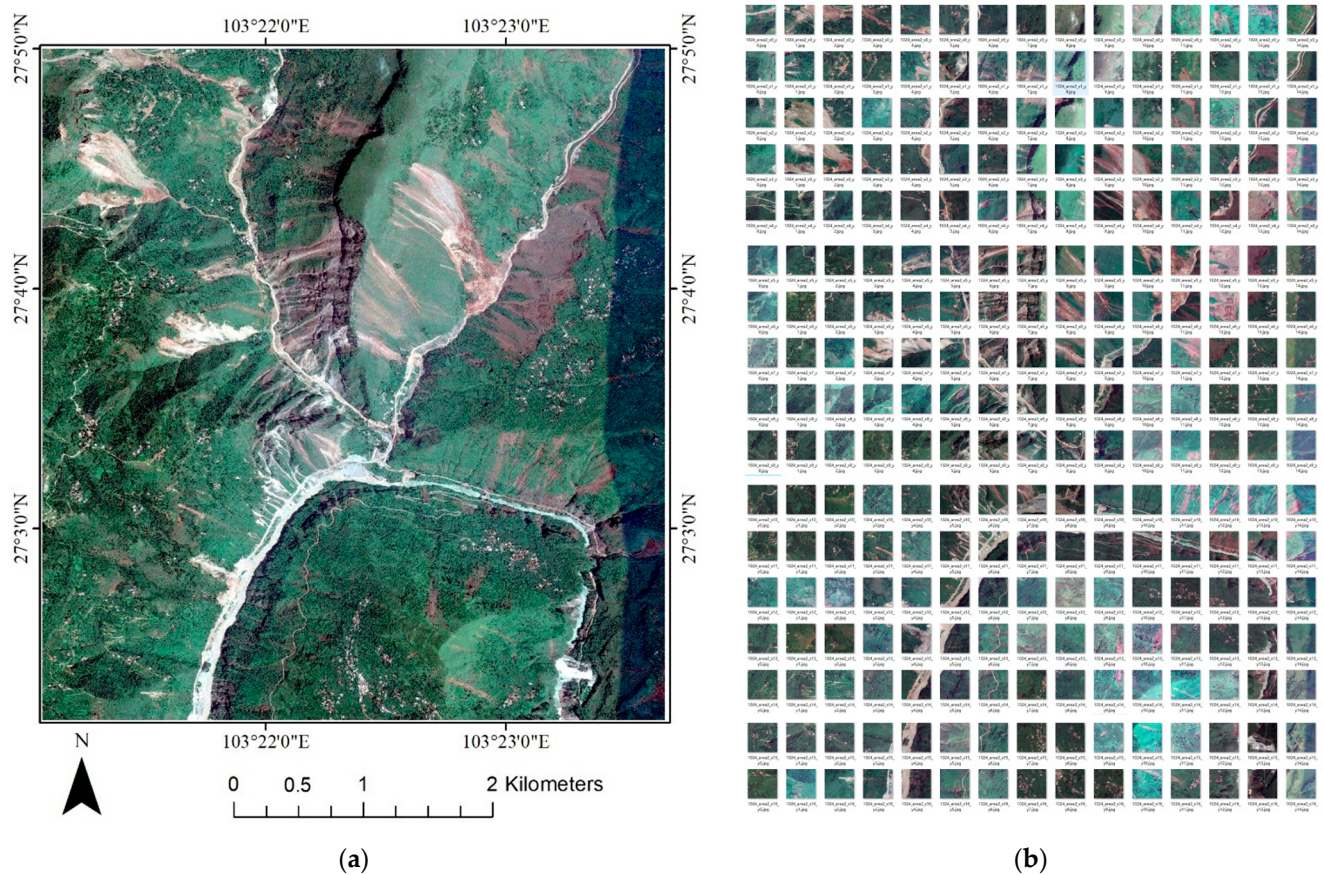


Figure 3. The image splitting process of the test area. (a) The original image of test area; (b) sequential splitting result.

This paper selects two areas as examples for comparing original labels from Xu et al. [31] and our new labels, as shown in Figures 4 and 5.

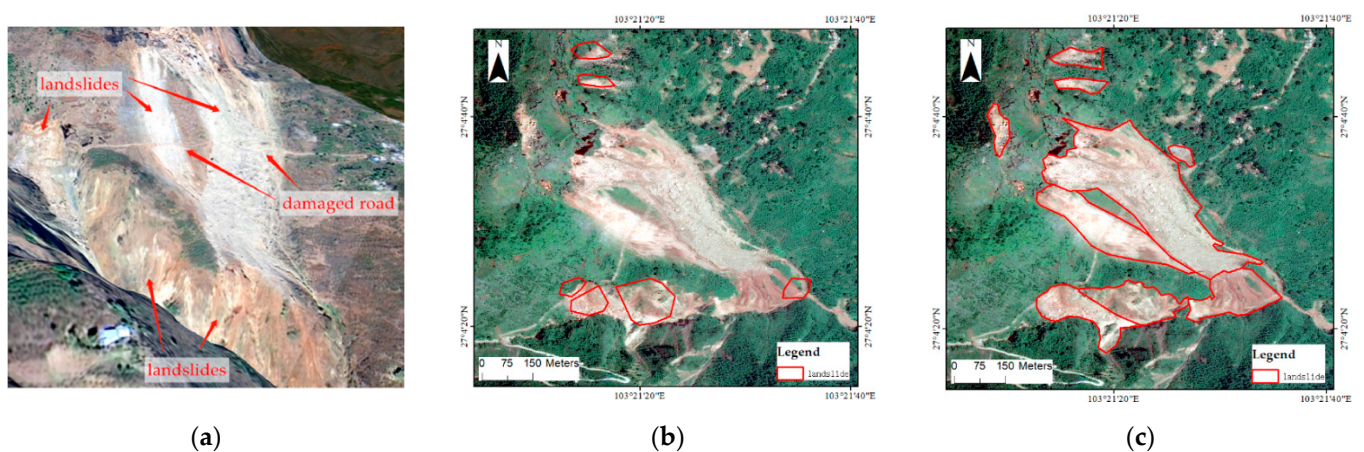


Figure 4. Examples of corrections on old landslide labels. (a) 3D Google Earth image taken on 20 August 2014 with resolution 0.27 m, view to northwest; (b) Google Earth image taken on 20 August 2014 and the labels from Xu et al.; (c) Google Earth image taken on 20 August 2014 and the labels from this study.

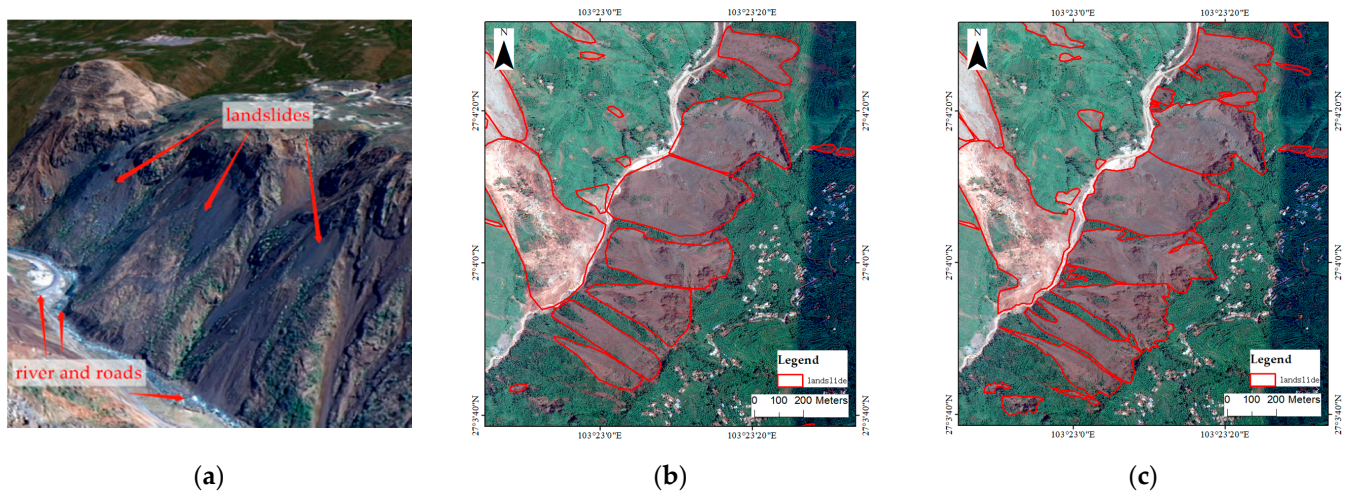


Figure 5. Examples of modifications on landslide boundaries. (a) 3D Google Earth image taken on 20 August 2014 with resolution 0.27 m, view to east; (b) Google Earth image taken on 20 August 2014 and the labels from Xu et al.; (c) Google Earth image taken on 20 August 2014 and the labels from this study.

It can be seen from Figure 4a,b that large areas of old landslides are not labeled, since this study does not only extract landslides triggered by a single event, so we include old landslides in the training and testing dataset. Labels shown in Figure 5b have rough boundaries. Considering that rivers and roads are areas most prone to false extractions, mislabels on these background objects should be avoided as much as possible. Generally, in the new labels, some old landslides and smaller landslides are relabeled, meanwhile the boundaries of landslides are more accurate and detailed.

3. Methods

3.1. Landslide Extraction Framework

Our study aims to build and evaluate landslide extraction methods using background-enhancement method and auxiliary landslide-inducing data. A general demonstration of our work is shown in Figure 6. The whole process mainly includes three parts: data preparation, comparative experiments, and result analysis.

In the data preparation part, we first obtain post-disaster optical satellite images and crop them, label landslides based on open dataset created by Xu et al., and make modifications with reference to 3D satellite images from Google Earth platform and high-resolution pre- and post-disaster images. After these steps, the original satellite samples for training and testing are generated. Then, we select factors having strong correlations with landslide distribution based on quantitative analysis on study area. In this study, DEM, slope, and distance from river are selected for the subsequent experiments, and samples of these additional data are created corresponding to the original satellite samples. Lastly, background-enhancement method is applied to original satellite samples and landslide-inducing data to generate background-enhanced samples.

This study divides the training samples into four types, as shown in Figure 6. The comparative experiments will be conducted based on them, to analyze the impacts brought about by different training data. Furthermore, the background-enhancement method proposed in this study is optimization on data level, so we combine the method with different deep learning models to test its applicability and effectiveness.

Finally, quantitative evaluations on landslide extraction results will be performed based on accuracy, recall, F1 score, and mIoU (mean intersection over union), and we will compare the extraction results and ground truth values in detail, to further analyze the influences brought about by the proposed methods.

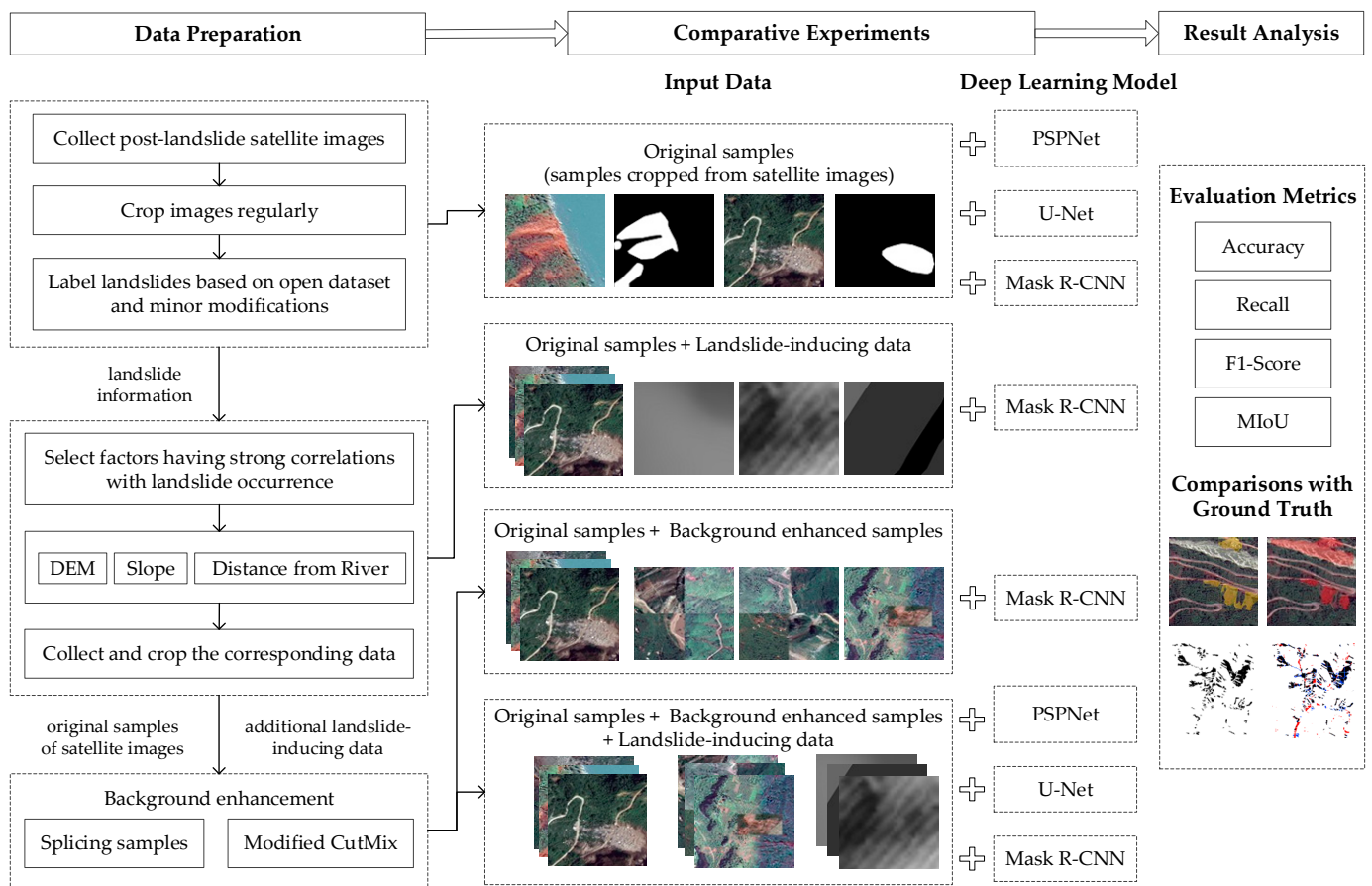


Figure 6. Flowchart of constructing and analyzing landslide extraction models.

3.2. Mask R-CNN Model

Since Girshick et al. [37] proposed the original R-CNN (region-based convolution neural network) model in 2014, many researchers have made contributions to new versions of R-CNN model with various improvements. The object detection model used for our comparative experiments is Mask R-CNN [28] developed by He et al. While some versions of R-CNN, such as fast R-CNN [38] and faster R-CNN [39], only focus on locating objects with bounding boxes, Mask R-CNN adds a branch that outputs a binary mask to further achieve pixel level segmentation of objects. In addition, Mask R-CNN replaced ROI pooling with a new method called ROI align, which can correct the misalignment brought about by ROI pooling [28].

In general, the process of object detection and segmentation using Mask R-CNN includes two main parts: the region proposal and the classification. The main steps are shown in Figure 7.

When an image is input into the training process, the image feature extraction will be performed first in the ResNet101 [27] backbone, which consists of five different convolution modules and output feature layers with different scales. Then, the feature pyramid network (FPN) [40] deals with these feature layers. Each layer will be upsampled and merged with the layer at the next level to obtain a new merged feature layer, which will later be sent into region proposal network (RPN) to generate proposal boxes using sliding windows. After generating the proposal boxes, ROI align cuts out the corresponding area according to the proposal box's position on the feature layer and pools it into a new feature layer with fixed size. Finally, the subsequent classification, regression, and mask generation are carried out. The input convolutional kernel size for classification and regression is 7×7 , and for mask generation is 14×14 .

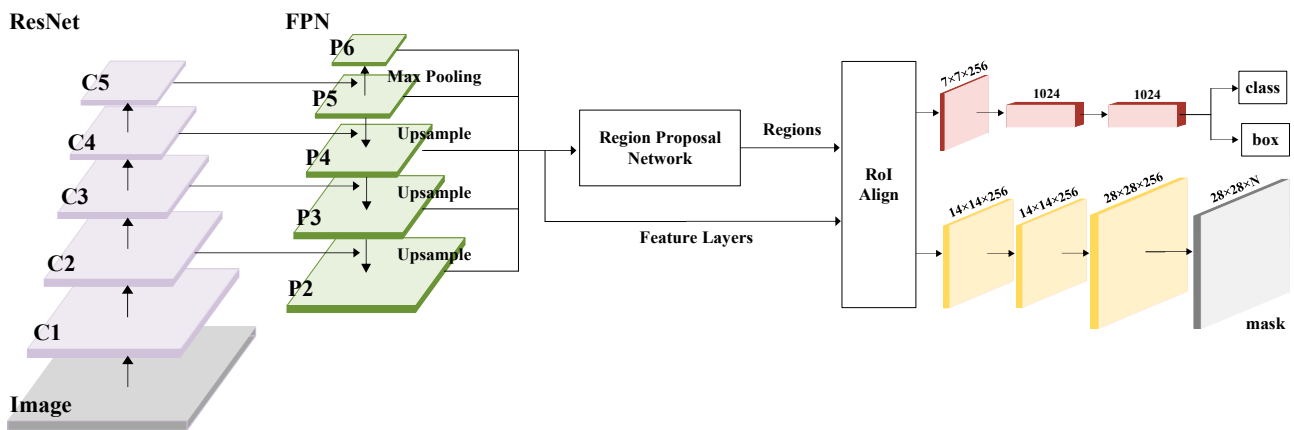


Figure 7. Mask R-CNN model structure.

3.3. Background Enhancement

In order to apply deep learning methods effectively, a labeled training dataset with great quantity and quality must be prepared to ensure the features can be properly learned. Compared with some classic examples of object detection using Mask R-CNN, such as detecting balloons and cells, landslides detection has more complex characters both in background and landslide itself.

Taking landslides in Ludian as an example, the largest landslide has an area of about 345,000 m², while the smallest is only about 12 m² [32]. In addition, seen from these samples cut from images of training area with a size of 1024 × 1024 shown in Figure 8, the color, shape, and texture of landslides also show great diversity. In Figure 8a, there is a landslide of common oval shape, with some small gray rocks and sands covering the road. In Figure 8b, this landslide is just a part of the giant Hongshiyan landslide [31], so the boundary cannot be seen in this block, and it is also covered by many stones. Compared with the former two examples, the texture of landslides shown in Figure 8c,d is rather smooth. The landslides in Figure 8c are covered by reddish brown mud, while those shown in Figure 8d are more similar to debris flows with slender shapes.

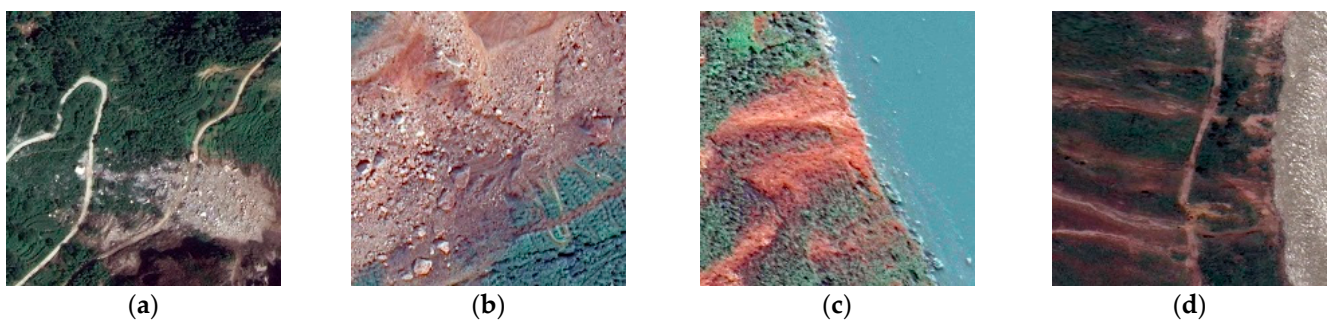


Figure 8. Samples of landslides in Ludian with different characters. (a) A small-scale landslide with gray color and oval shape; (b) A large-scale landslide with crushed stones; (c) Landslides with reddish brown color; (d) Landslides with slender shapes.

To reduce miss and false extraction, the model should learn the texture, shape, color, and other features of landslides and be able to distinguish the differences between landslides and confusing backgrounds. However, in application of landslide detection, the quantity and quality of the training dataset is often unsatisfactory. The majority of landslide samples have simple backgrounds, such as the samples shown in Figure 9. Although these samples have landslides with different sizes, colors, textures, and shapes, these features cannot be properly learned during the training process because landslides are the only

bright areas in these samples, so only very few features are needed for detecting landslides from simple samples.

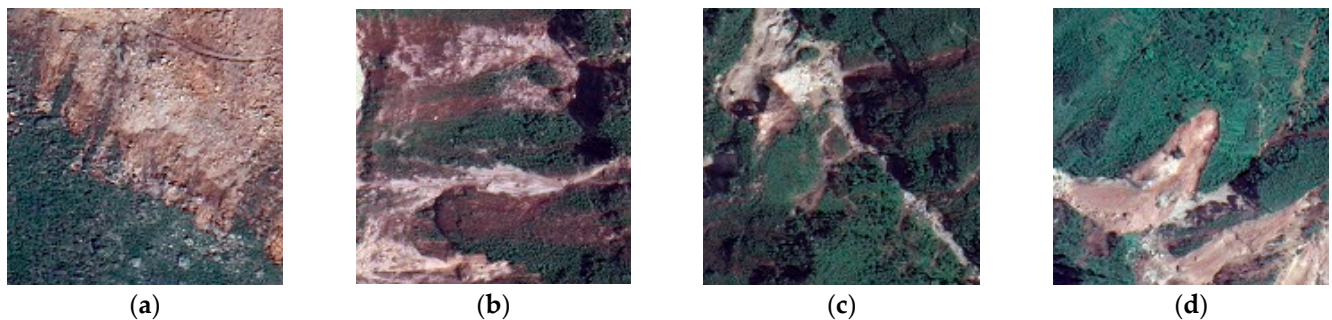


Figure 9. Samples with simple background. (a) A large-scale landslide with green vegetation as background; (b) Three brown tongue-shape landslides with green vegetation as background; (c) A small landslide with green vegetation as background; (d) Tongue-shape landslides also with vegetation as background.

In contrast, the sample shown in Figure 8d has complicated backgrounds, providing comparisons between landslides, river, and roads, but the number of complicated samples in our training dataset is rather small.

To deal with the complex characters of landslides and background and overcome the shortage in training data, we propose a background-enhancement method to create complicated samples, including the following two steps:

- (1) Background enhancement by splicing images.
- (2) Background enhancement by a modified CutMix.

3.3.1. Background Enhancement by Splicing Images

The operations of background enhancement by splicing images are shown in Figure 10. Firstly, one landslide image and three different non-landslide images are randomly chosen from the training dataset. Then they are spliced into a new sample, in which some confusing background objects such as bare land, river, or fields may be added, and so the landslide is surrounded by more complicated background objects. We expect that these samples with splicing operations can help the training process learn more characters of background and improve the ability to detect landslides from confusing backgrounds.

Compared with the simulated hard samples created in the study carried out by Jiang et al. [15], the selection of landslide and non-landslide samples to be spliced in our study is random. Automatic and random selection of samples can reduce the subjective impact brought about by manual work and improve efficiency.

3.3.2. Background Enhancement by a Modified CutMix

The CutMix method [29] proposed by Yun et al. creates a new sample by replacing a random region in an image with pixels from another image to make models pay attention to the entire region instead of some parts that are easy to distinguish, and CutMix method has achieved better performance than similar strategies such as Mixup and Cutout [29]. Furthermore, instead of simply identifying the object category in the images, our study needs to determine the position of landslides, so methods such as Mixup and Cutout are less applicable compared with CutMix.

If only some discriminative characters, such as bright color, are learned for landslide extraction, then the false extraction may happen more frequently when faced with confusing background objects. To learn more detailed features of landslides and differences between landslide and landslide-like background objects, a modified CutMix method is used in this study.

The operation of our modified CutMix is demonstrated in Figure 11.

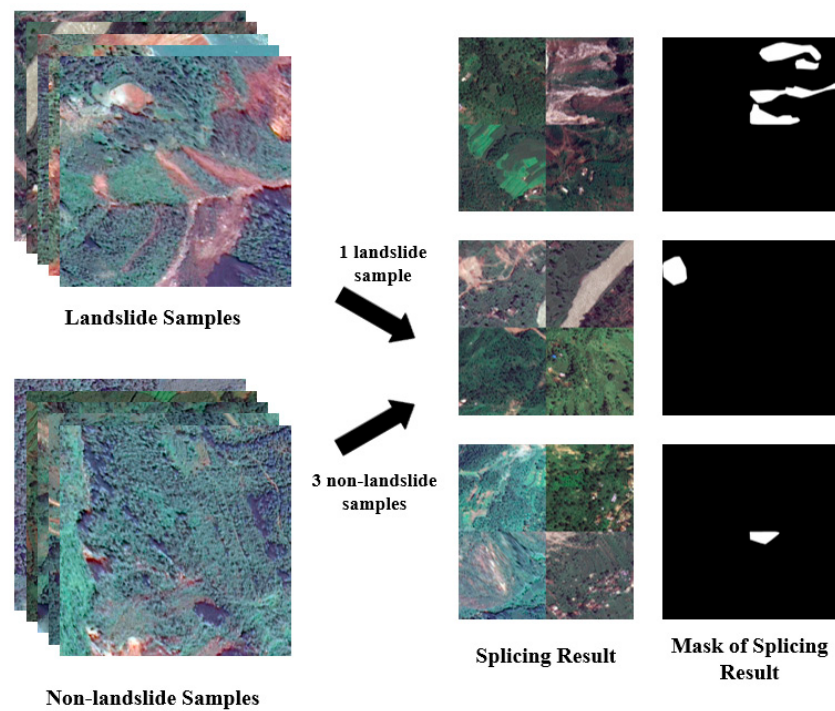


Figure 10. Process of background enhancement by splicing images.

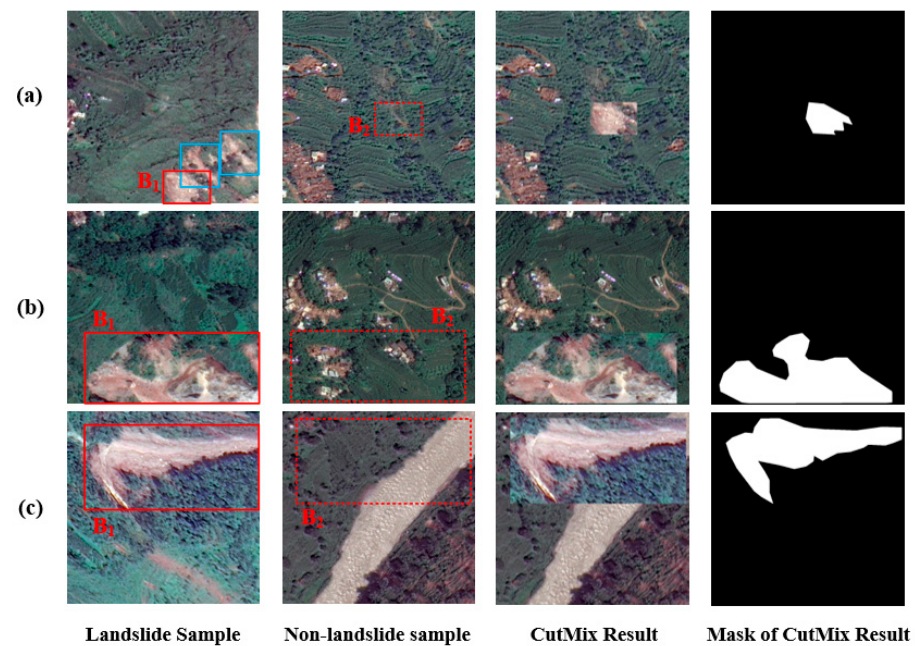


Figure 11. Process of background enhancement by using a modified CutMix. (a) A small landslide was randomly selected among three landslides to fill the part in non-landslide sample; (b) A large-scale landslide was used to fill non-landslide sample; (c) A large-scale landslide was attached to river area in non-landslide sample.

Step 1: One landslide sample (S_1) and one non-landslide sample (S_2) are randomly selected from the training dataset.

Step 2: The area (B_1) to be cut out from S_1 and the area (B_2) to be replaced in S_2 are defined. Unlike the original CutMix method, the areas to cut out and replace are of the same size but possibly with different locations. B_1 is the bounding box of a landslide in S_1 .

If there are multiple landslides in S_1 , only one will be selected (as shown in Figure 11a), and the selection is random. B_2 has the same width and height as B_1 , but is relocated randomly.

Step 3: The offset between B_1 and B_2 caused by the relocation is calculated, the this offset is used to set the label of this newly generated sample.

Some examples of samples generated using CutMix operations are also shown in Figure 11. Different from the result of splicing operations, background objects adjacent to landslide instances are all replaced. As shown in Figure 11b,c, the landslides in the original images can be detected rather easily because they are the only areas with lighter color, but if only judged by very few features, some background objects in the CutMix samples such as river surface, buildings, and bare land may lead to false extraction, so more landslides will be learned from these CutMix samples and the more perceptive the model will be.

3.4. Landslide Inducers

Topography, geology, seismology, meteorology, and hydrology [41–43] are common factors when analyzing the causes of landslide formation, but so far, these kinds of information have not been widely adopted as controlling factors in landslide detection.

This study adds landslide inducers into training data as a supplement because they may offer great help when the model is deciding an area is landslide or not. For example, if an area with landslide-like color, shape, and texture is located in flat terrain, then the possibility for it being actual landslide is rather low. Therefore, based on these auxiliary data, landslide-prone regions can be identified, thus bringing further help in landslide detection. Among all the common inducers, three factors, DEM, slope, and distance from river, are selected, because they are the major landslide-inducing factors and have shown strong correlation with the distribution of landslides in Ludian. Furthermore, the DEM and river system data are easy to obtain, so choosing these factors as auxiliary data can be more conveniently implementable in future research.

The elevation data and the distribution of landslides are demonstrated in Figure 12a. Areas with different elevations usually have different climates, then cause further impact on the occurrence of landslides. As shown in Figure 12b, areas with relatively low elevation, such as mountain and river valley, have more landslides. Landslide area density also shows a decreasing trend as elevation value grows.

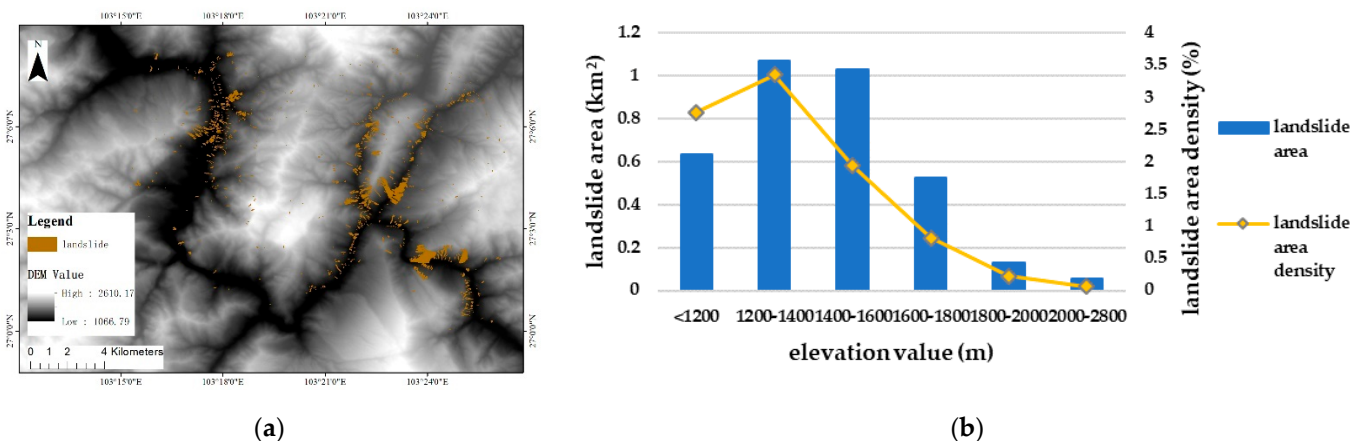


Figure 12. Analysis on relations between landslide distribution and elevation in Ludian. (a) The elevation data of study area; (b) the relations between landslide distribution and elevation.

According to many landslide susceptibility studies, slope is the main factor for landslide formations [44,45]. From Figure 13a, we can see an evident correlation between the distribution of landslides and slope. Most landslides happen in red area with steeper terrain. As shown in Figure 13b, the slope gradient in the study area varies from 0° to 81° . About 80.35% of landslides happened in areas with slope of 20° – 50° , and landslide area

density shows a increasing trend as slope value grows, which is probably because the more significant the gradient, the stronger the gravity, then the more prone the landslide.

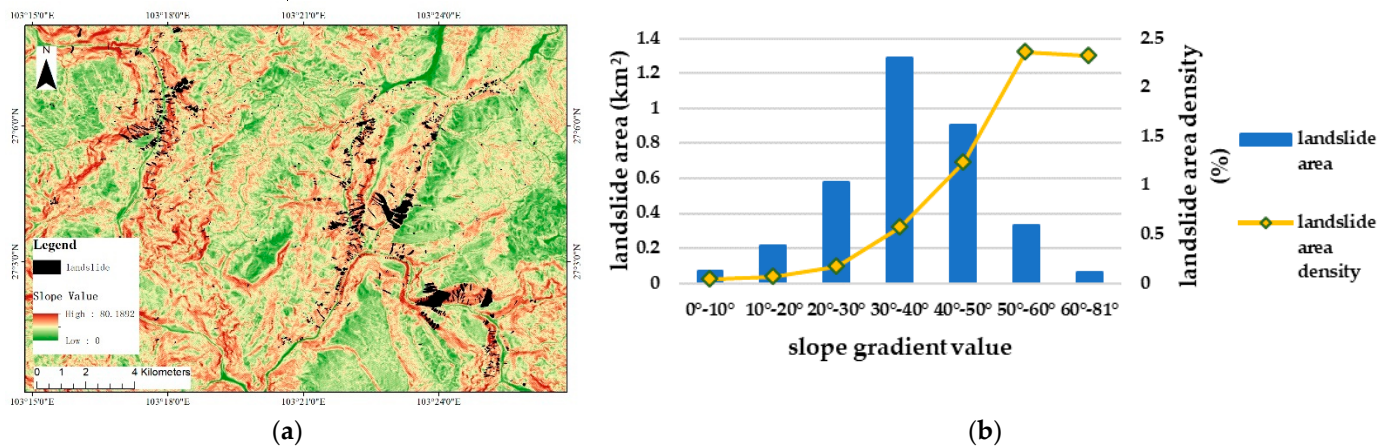


Figure 13. Analysis of relations between landslide distribution and elevation in Ludian. (a) The slope gradient of study area; (b) the relations between landslide distribution and slope gradient.

As mentioned earlier, most landslides are located in the valley area. Normally, the side bank of the river valley is easily washed by the water flow; meanwhile, the frequent change of the river water level will also affect the formation of landslides. Distance from river has already been used in many landslide susceptibility maps, and it has shown a strong correlation with the occurrence of landslides [27].

In the study area, there are Niulan River, Longquan River, and Shaban River, as shown in Figure 14a. The relation between the distribution of landslides and the distance from river is evident: the majority of landslides locate in areas closed to river. When we detected landslides in the test area using the traditional Mask R-CNN model, we found that the emergence of false extractions along the river is relatively frequent. Therefore, in order to ease this problem, we made use of river data from Open Street Map to add auxiliary information. Firstly, we made a 30 m buffer using Open Street Map data, and this buffer is roughly estimated as river surface. Then, we made buffers of 200 m, 200–500 m, 500–1000 m, and 1000–2000 m based on the river surface, and divided the study area into six levels, as shown in Figure 14a. Location with 0 m distance from river is river surface, which is designed specifically to solve the false extractions on river surface, because when an area has a landslide-like bright character, but with 0 m distance from river, then the area is more likely to be river surface instead of landslide, but for areas with distance > 0 m, as shown in Figure 14b, there is a significant trend that the smaller the distance the more likely to landslide.

In view of these relations between topographic factors and the distribution of landslides, we integrate these landslide-inducing factors with color, texture, and shape provided by optical satellite data for comprehensive evaluation, in order to achieve better landslide extraction results.

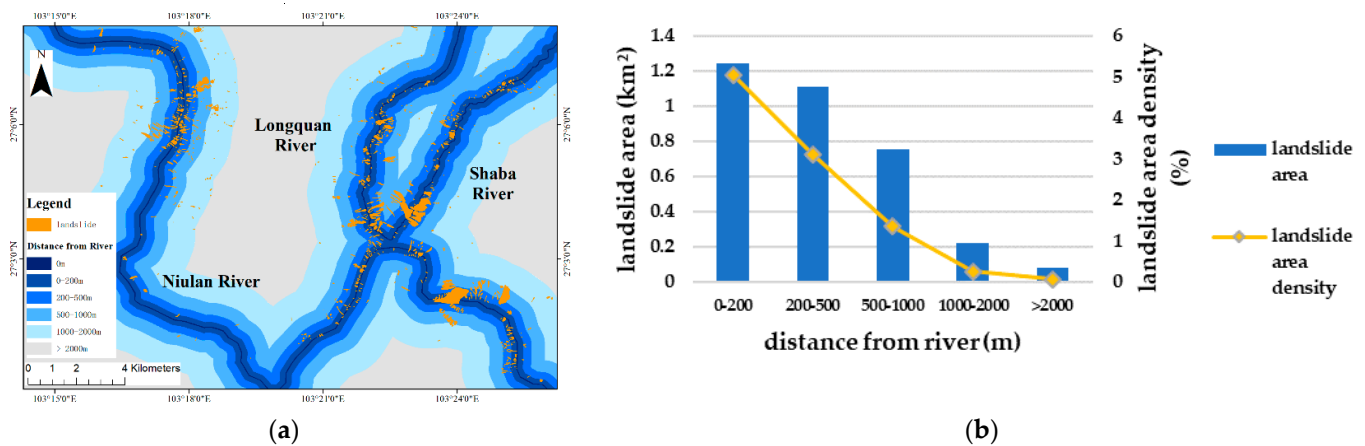


Figure 14. Analysis of relations between landslide distribution and elevation in Ludian. (a) The river data of study area; (b) the relations between landslide distribution and distance from river.

4. Experiment

4.1. Accuracy Evaluation

Precision, recall, F1 score, and mIoU (mean intersection over union) are common measures of accuracy evaluation for object detection and instance segmentation, and they are applied to validate the effectiveness of the methods proposed in this research.

The landslide extraction in this research is a binary classification problem, that is, for each pixel, there are only two cases: landslide or background. Therefore, the validation is based on four kinds of extraction results, which are shown in Table 2, namely, TP (true positive), FP (false positive), TN (true negative), and FN (false negative). TP are the areas correctly extracted as landslide. FP are the background areas incorrectly extracted as a landslide. FN are the landslide areas incorrectly extracted as background. TN are the areas correctly extracted as background.

Table 2. Confusion matrix [46] between predicted value and true value.

True Value \ Predicted Value	Predicted Value		
	Landslide	Background	
Landslide	True Positive (TP)	False Negative (FN)	
Background	False Positive (FP)	True Negative (TN)	

Precision is used to evaluate how many areas extracted as landslides are real landslides. Recall is used to evaluate how many landslides are correctly extracted. Generally speaking, recall will be lower when precision is high, and precision will be lower when a recall is high. To evaluate our model in a more balanced way, F1 score is added. These measures are defined in Equations (1)–(3).

$$\text{precision} = \frac{TP}{TP + FP} \quad (1)$$

$$\text{recall} = \frac{TP}{TP + FN} \quad (2)$$

$$\text{F1 Score} = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (3)$$

The Mask R-CNN model can detect the bounding box of landslides and meanwhile extract the shape of landslides. To compare the extracted shape and ground truth value, we used the mIoU [47] as the accuracy evaluation measure. It calculates the value of two areas'

intersection that divides two areas' union. This measure has been widely used in semantic segmentation and is described in Equation (4).

$$mIoU = \frac{1}{n+1} \times \sum_{i=0}^n \frac{P_{ii}}{\sum_{j=0}^n P_{ij} + \sum_{j=0}^n P_{ji} - P_{ii}} \quad (4)$$

where n is set as 1 in our case, because landslide is the only category. p_{ii} means a pixel of type i predicted as type i . P_{ij} means a pixel of type i predicted as type j . P_{ji} means a pixel of type j predicted as type i .

4.2. Experimental Design

The hardware environment of this research is as follows: graphic card is GeForce GTX 1080 Ti, the processor is Intel(R) Xeon(R) Silver 4210, and memory is 64 GB.

We choose the open source code of matterport Mask R-CNN for comparative experiments (https://github.com/matterport/mask_RCNN, accessed on 1 October 2021). It is built on feature pyramid network and a ResNet101 backbone, using Python 3, Keras, and TensorFlow. The configurations for model training and validation are set as follows: epochs at 20, learning rate at 0.001, weight decay at 0.005, steps per epoch at 1000, gradient clipping at 5.0, and a pretrained Mask R-CNN weighted from MS COCO dataset is obtained to transfer parameter initial values.

Based on this environment and Ludian dataset, the first group of comparative experiments is designed. As shown in Table 3, there are four experiments designed with the same deep learning model but different input training datasets. The original satellite images are cut from two training areas, with an amount of 531. Background-enhanced samples are created based on original samples though background-enhancement operations mentioned in Section 3.3 and have an amount of 600. Corresponding landslide-inducing data for both original and background-enhanced samples were made during data preparation. Experiment I uses only original satellite images as input dataset. Taking it as base experiment, Experiment II adds additional landslide-inducing data, Experiment III adds background-enhanced samples, and Experiment IV adds both landslide-inducing data and background-enhanced samples. These experiments are designed to analyze impacts brought about by different training data, and to validate the effectiveness of the proposed method.

Table 3. Settings of comparative Experiment I.

No.	Deep Learning Model	Training Dataset
I	Mask R-CNN	Original Satellite Images
II	Mask R-CNN	Original Satellite Images + Landslide-inducing Data
III	Mask R-CNN	Original Satellite Images + Background-Enhanced Samples
IV	Mask R-CNN	Original Satellite Images + Background-Enhanced Samples + Landslide-inducing Data

A major character of the background-enhancement method proposed in this study is that it only bring changes on a data level, so, theoretically, it can be applied to deep learning models with different structures. To further test the applicability and effectiveness of the proposed method, we also use U-Net [48] and PSPNet [49] to conduct landslide extraction experiments. Both U-Net and PSPNet are relatively advanced deep learning models in the field of instance segmentation and have fine and stable performance in many studies. They can be reliable for assessing whether the model built in this study has advantage in landslide extraction tasks and for analyzing the impacts on different deep learning models from background-enhancement method. This study builds U-Net and PSPNet based on open source code from Divam Gupta (<https://github.com/divamgupta/image-segmentation-keras>, accessed on 25 April 2022). Comparative experiments are designed by changing input training data, as shown in Table 4.

Table 4. Settings of comparative Experiment II.

No.	Deep Learning Model	Training Dataset
V	U-Net	Original Satellite Images
VI	U-Net	Original Satellite Images + Background-Enhanced Samples + Landslide-inducing Data
VII	PSPNet	Original Satellite Images
VIII	PSPNet	Original Satellite Images + Background-Enhanced Samples + Landslide-inducing Data

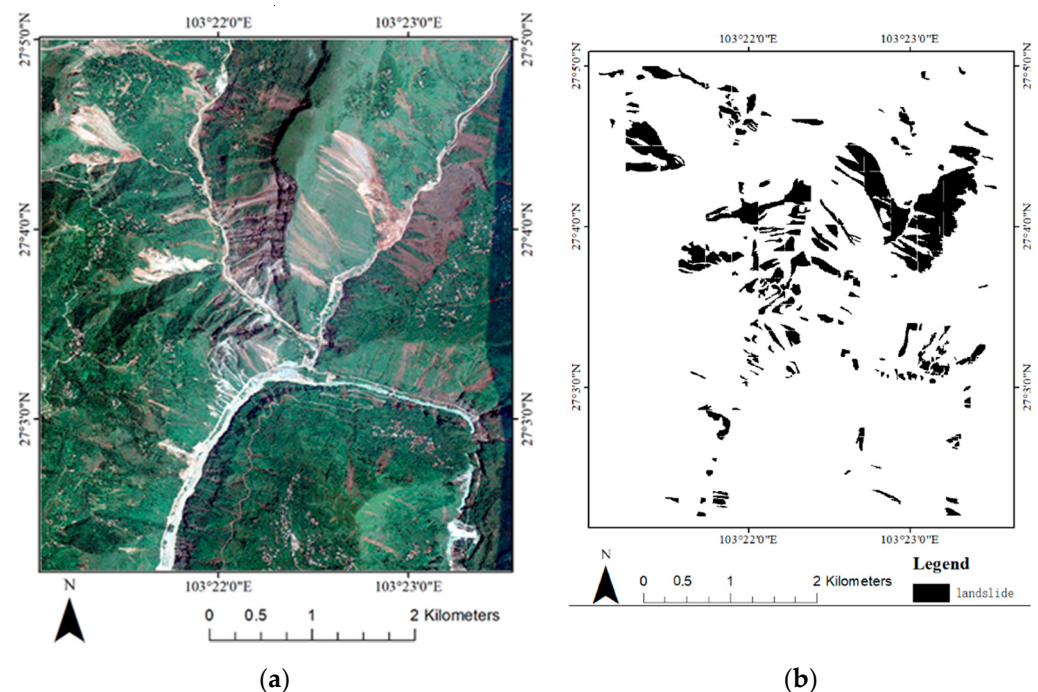
Furthermore, an experiment is also designed to extract landslides using satellite images taken on another date after disaster to test the model's performance on multi-temporal analysis. The detailed information of this post-landslide image is shown in Table 5. The image was taken at 5 May 2015, a long time after the disaster. Some landslide areas have been repaired already, or covered with vegetation, and there is no reliable ground truth value for extraction result evaluations. Therefore, metrics for accuracy evaluation are not calculated, and only some areas with old landslides are used for analysis.

Table 5. Detailed information of satellite images.

	Resolution	Source	Satellite	Collected Date
Post-landslide image	0.27 m/pixel	CNES/Airbus	Pleiades PHR1B	5 May 2015

4.3. Results and Discussions

The original image and ground truth landslide distribution map of the test area is demonstrated in Figure 15, showing a reference for the extraction results. Landslides from the test area were extracted using models trained from four comparative experiments, and the results of four experiments are shown in Figure 16a–d.

**Figure 15.** Information of test area. (a) Original satellite image; (b) ground truth.

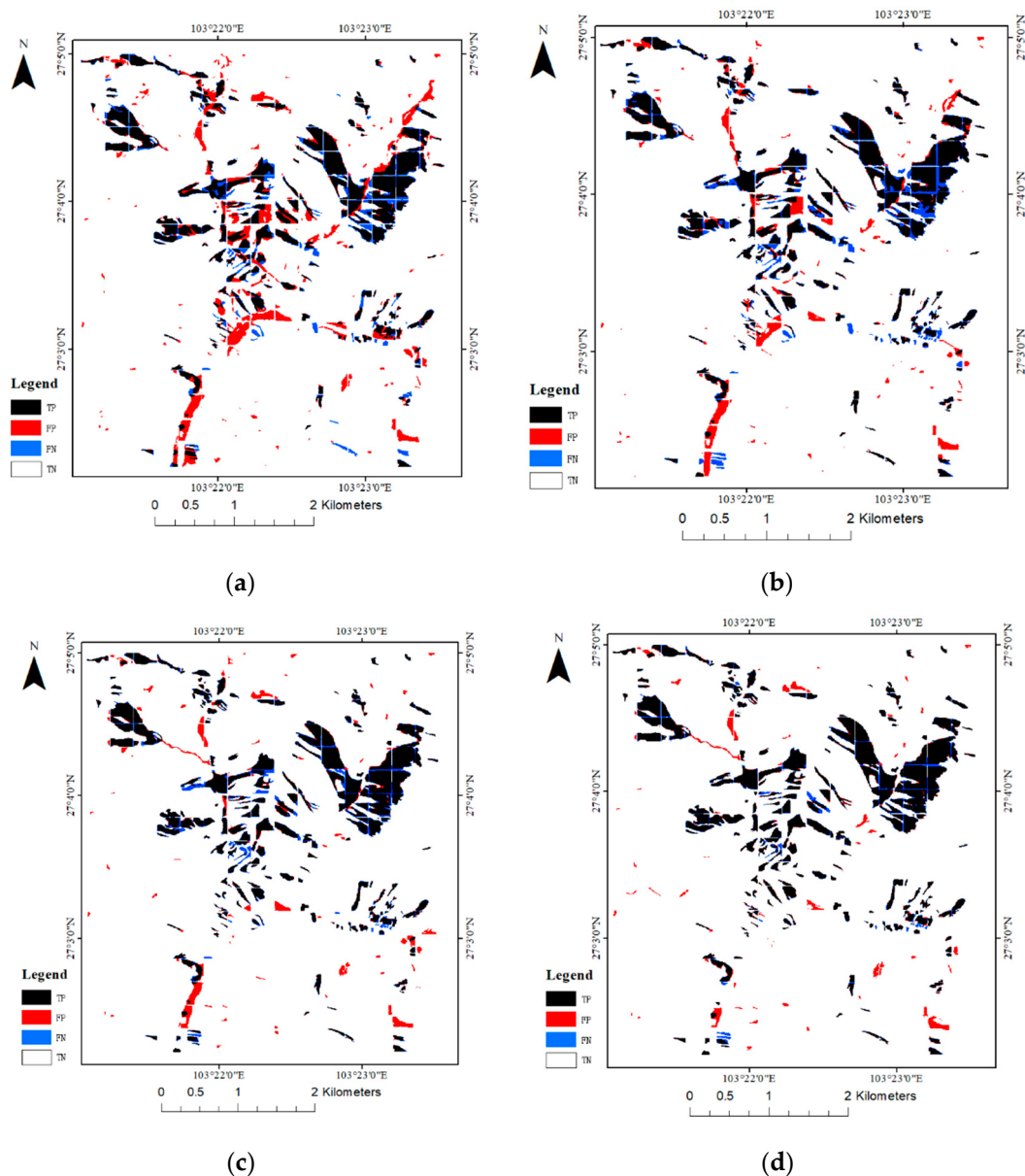


Figure 16. Landslide extraction results using Mask R-CNN and different input data. (a) Experiment I using original satellite images; (b) Experiment II using original satellite images and landslide-inducing data; (c) Experiment III using original satellite images and background-enhanced data; (d) Experiment IV using original satellite images, background-enhanced data, and landslide-inducing data.

Figure 16a shows the result of Experiment I. The model is trained with only original satellite images. We can see that there are a large number of false extractions (red area) along the valley and river, and the missed extractions (blue area) are very obvious. Figure 16b is the extraction result of the model trained with original images and landslide-inducing data. It can be seen that the false extractions along the valley have decreased with the help of landslide-inducing data, and some tiny false extractions caused by buildings have disappeared. However, the missed extractions are not improved. Figure 16c is the result of the model trained with original satellite images and background-enhanced samples. From this result, we can see that the false extractions along the river and valley have disappeared mostly, and the missed extractions along the image blocks located in the northeast area have decreased significantly. Figure 16d is the result of the model trained with original satellite images, landslide-inducing data, and background-enhanced samples. Using both auxiliary landslide-inducing data and background-enhanced samples, Experiment IV has

the best extraction result, on the whole. The areas of false and missed extractions have decreased to a large extent, and the shape of the landslide is more complete compared with the former three experimental results, but the performance in avoiding tiny false extractions caused by buildings is not as satisfying as Experiment II.

To show the changes brought about by the proposed method in a more detailed way, five representative samples were selected from the test dataset. Original satellite images and ground truth labeled data (yellow masks) are shown in Figure 17a–e,f–j, respectively, giving a reference for the comparisons. Comparisons between four comparative experiments are shown in Figure 18. Furthermore, to highlight differences between ground truth and extraction results, boxes with different colors are drawn. Yellow boxes in extraction results indicate false extractions, while white boxes indicate missed extractions. By analyzing the changes of boxes, the impacts brought about by using different training data can be concluded.

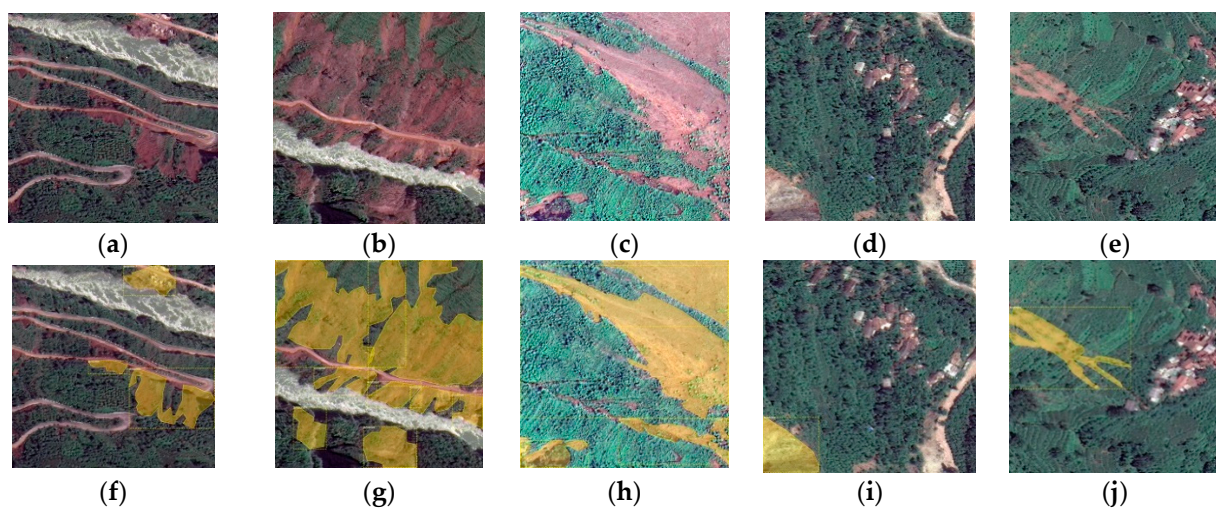


Figure 17. Samples for detailed analysis. (a–e) Original images of 5 samples; (f–j) ground truth labels of 5 samples.

The extraction results are presented as red masks shown in Figure 18. From extraction results of Experiment I shown in Figure 18a–e, we can see that the main factors leading to false extractions are roads, river, and buildings, probably because these objects have high reflectivity and are very similar to landslides in spectral characteristics, so it is rather difficult to distinguish these objects and real landslides through automatic ways. At the same time, some smaller landslides are neglected when multiple landslides exist in one sample; see Figure 17c. After adding landslide-inducing data, as shown in Figure 17f–j, the number of false extractions on river and roads reduced, but missed extractions increased. Figure 17k–o show the extraction results after adding background-enhanced samples. While false extractions reduced, the missed extractions were also less than the results of Experiment III. The extraction results of Experiment IV using both landslide-inducing information and background-enhancement method are shown in Figure 17p–t. Comparatively, it has achieved the best performance. Although missed extractions still exist, the proposed method has avoided majority false extractions on roads, rivers, and buildings. The shapes and boundaries of landslides are also better identified.

The improvements on landslide extraction accuracy brought about by the methods proposed in this study should also be evaluated by quantitative metrics. The precision, recall, F1 score, and mIoU of four comparative experiments are shown in Table 6. Individual impacts brought about by using landslide-inducing data or adding background-enhanced samples are demonstrated by experimental results II and III, respectively, and the experimental result IV shows the performance when using both methods.

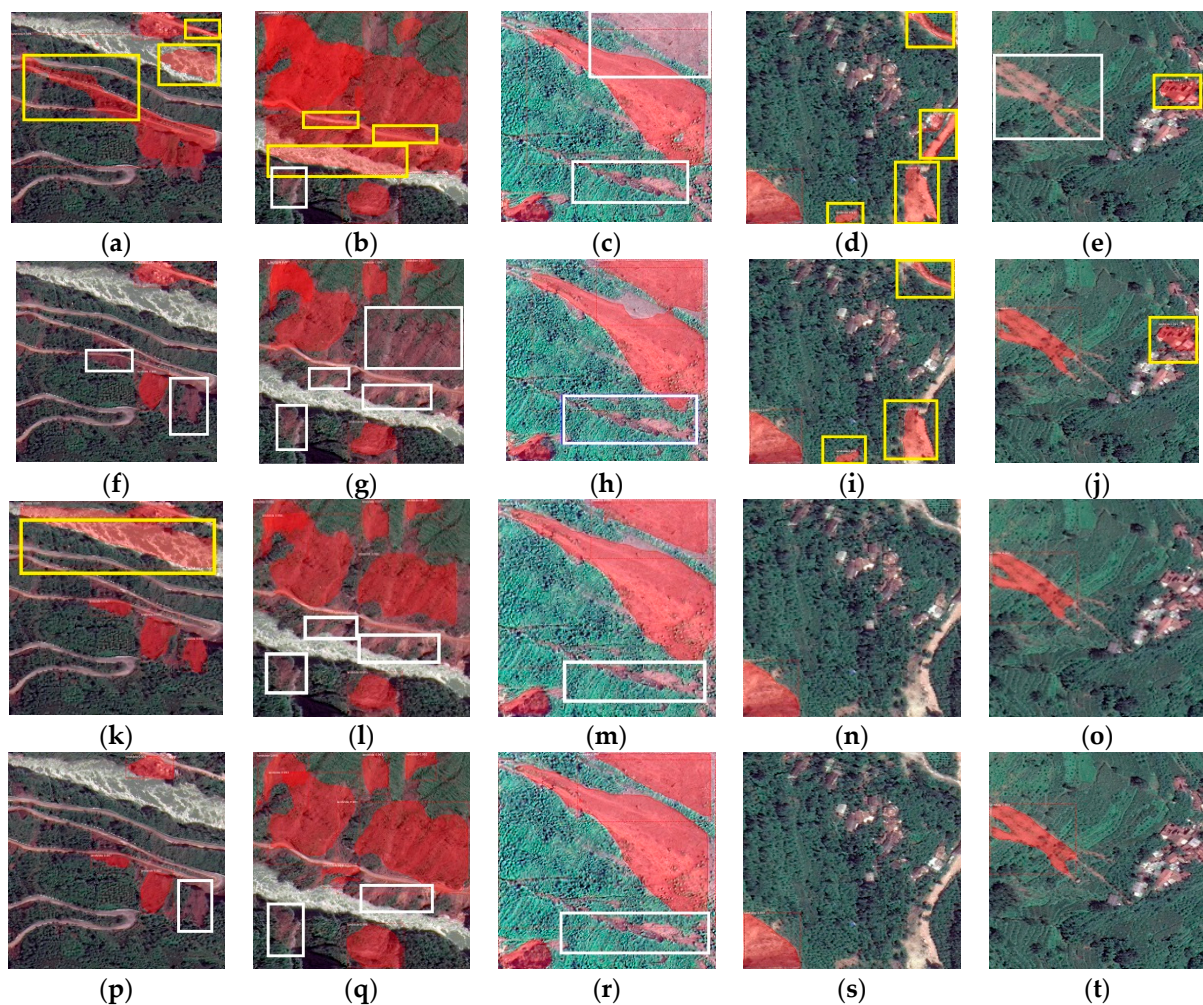


Figure 18. Comparisons of extraction results and ground truth. (a–e) Extraction results of 5 samples from Experiment I; (f–j) extraction results of 5 samples from Experiment II; (k–o) extraction results of 5 samples from Experiment III; (p–t) extraction results of 5 samples from Experiment IV. Yellow box means false extraction, and white box means missed extraction.

Table 6. Comparison of landslide extraction results from Mask R-CNN model trained with different input data.

No.	Model	Input Data	Precision/%	Recall/%	F1 Score/%	mIoU/%
I	Mask R-CNN	Original Satellite Images	67.26	79.31	72.79	75.53
II	Mask R-CNN	Original Satellite Images + Landslide-inducing Data	78.76	78.31	78.53	80.11
III	Mask R-CNN	Original Satellite Images + Background-Enhanced Samples	84.83	84.92	84.87	85.28
IV	Mask R-CNN	Original Satellite Images + Background-Enhanced Samples + Landslide-inducing Data	88.68	89.49	89.08	89.00

Experiment I used Mask R-CNN model trained with original satellite images, and after adjustments on various settings, it achieved precision, recall, F1 score, and mIoU at 67.26%, 79.31%, 72.79%, and 75.53%, respectively. Based on the settings in Experiment I, subsequent experiments were conducted. After adding landslide-inducing data, the precision was improved significantly by 17.10% but the recall was not improved. This is in accordance with previous analysis on extraction maps where false extractions decreased and missed

extractions increased. After adding background-enhanced samples, the precision, recall, F1 score, and mIoU were all improved by 26.12%, 7.07%, 16.60%, and 12.91%, respectively. The improvement brought about by background-enhancement method is much stronger than adding landslide-inducing data. This is probably because the background-enhancement method can also assist in increasing the amount of the training dataset. Finally, the improved model using both landslide-inducing data and background-enhanced samples achieved a great performance, with precision at 88.68%, recall at 89.49%, F1 score at 89.09%, and mIoU at 89.00%. Each parameter is 31.84%, 12.83%, 22.38%, and 17.83% higher than the traditional method using only satellite images as input.

Seen from the comparisons of extraction result maps and the quantitative metrics shown in Table 6, it can be concluded that the background-enhancement method can help models better distinguish the confusing background objects and landslides and can reduce false and missed extractions effectively. Combined with auxiliary landslide-inducing information, the performance can be even better.

To further test the applicability and effectiveness of the proposed method, we also use U-Net and PSPNet for landslide extractions. The second group of comparative experiments are also conducted by changing input training data, and the evaluations of extraction results from different methods are shown in Table 7.

Table 7. Comparison of landslide extraction results from different methods.

No.	Model	Input Data	Precision/%	Recall/%	F1 Score/%	mIoU/%
V	U-Net	Original Satellite Images	52.28	69.95	59.83	66.48
VI	U-Net	Original Satellite Images + Background-Enhanced Samples + Landslide-inducing data	68.29	78.00	72.82	75.59
VII	PSPNet	Original Satellite Images	56.61	65.86	60.89	67.52
VIII	PSPNet	Original Satellite Images + Background-Enhanced Samples + Landslide-inducing data	76.02	66.16	70.74	74.55

Both U-Net and PSPNet have better performance after adding background-enhanced samples and landslide-inducing attributes into training data. The U-Net model's precision, recall, F1 score, and mIoU were improved by 30.62%, 11.51%, 21.71%, and 13.70%, respectively. The PSPNet model's precision, recall, F1 score, and mIoU were improved by 34.29%, 0.46%, 16.18%, and 10.41%, respectively. Increase in precision normally relates to the increase in true prediction and the reduction in false extractions; and increase in recall normally relates to the increase in true prediction and the reduction in missed extractions. Therefore, based on the results, it can be concluded that the background-enhancement method can effectively improve the performance of the U-Net model and reduce false and missed extractions to a certain extent. Although the recall of PSPNet was only slightly improved, the overall accuracy was still much better than Experiment VII considering F1 score and mIoU.

Comparing the results from the above eight experiments, Mask R-CNN trained with both background-enhanced samples and landslide-inducing data has the best performance, on the whole. Based on this model, we obtained satellite images of test area taken at another time after the disaster, to test the landslide extraction performance on multi-temporal analysis. Due to the lack of reliable ground truth and changes on land surface, only some areas with old landslides are used for analysis.

The comparisons of extraction results using images taken on 5 May 2015 and ground truth labeled by visual interpretation are shown in Figure 19. These are four samples with a size of 1024×1024 pixels, in different locations of test area, and have different ground features. For samples shown in Figure 19a,b, the model achieved satisfying results that most landslide areas are correctly extracted, but for the sample shown in Figure 19c, there

is a large amount of missed extractions, and there are also false extractions on buildings in Figure 19h. Possible reasons for the mistakes mentioned above are deficiencies of the landslide extraction model, on one hand, and different spectral characteristics caused by vegetation coverage in different seasons. Fine-tuning may be helpful to achieve better results when using multi-temporal images for landslide extractions.

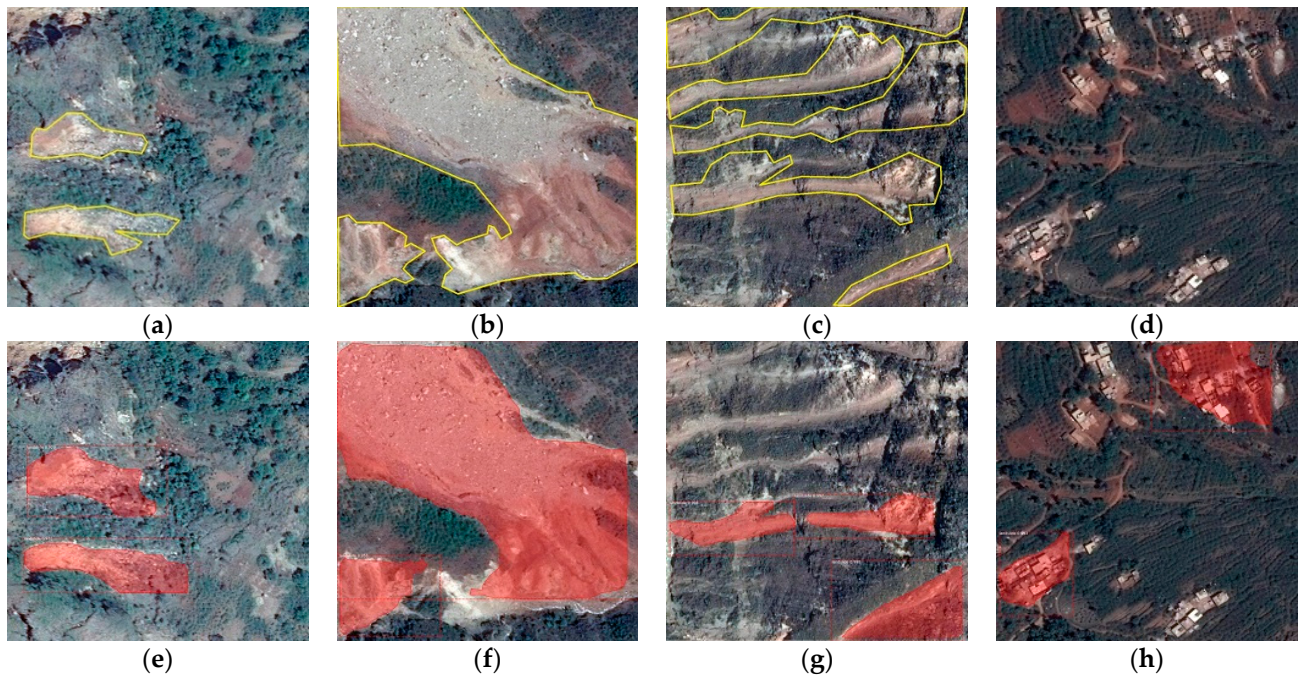


Figure 19. Comparisons of extraction results and ground truth. (a–d) Original satellite image and landslide ground truth (yellow); (e–h) extraction results of images taken in 5 May 2015, using model IV.

5. Conclusions and Prospect

In this study, to deal with the shortage in training data and false extractions on confusing background objects (e.g., river surface, bare land, buildings, roads), a background-enhancement method was proposed, and landslide-inducing factors (DEM, slope gradient, distance from river) were added to input data as auxiliary information. The landslide data of Ludian City, Yunnan Province, and Mask R-CNN model were used to test the feasibility of the method. In addition, U-Net and PSPNet were used to test the applicability of background-enhancement method on different deep learning models.

Through comparisons of landslide extraction results from Mask R-CNN model trained by different input data, it can be concluded that our proposed method can greatly help in reducing false extractions and increasing the accuracy of landslide boundary extraction results. The Mask R-CNN model trained using background-enhanced samples and landslide-inducing data achieved the best performance, on the whole, with precision at 88.68%, recall at 89.49%, F1 score at 89.09%, and mIoU at 89.00%. Compared with the traditional model trained with only satellite images, each metric improved by 31.84%, 12.83%, 22.38%, and 17.83%, respectively. In addition, the background-enhancement method and the use of landslide-inducing factors can bring improvement to different deep learning models theoretically, because the modifications are only on the data level. The results from the second group of comparative experiments using U-Net and PSPNet and different input training samples can support this viewpoint to some extent.

However, the extraction results still have errors, and the training areas and testing areas in this study are rather close geographically. Considering the multi-temporal analysis results in this study, if we want to apply landslide extraction based on the pretrained model to a new area, or the same area but different time, then some adjustments are required.

To reduce the workload of transfer learning, more effort will be placed on improving the adaptability and mobility of the landslide extraction model in future work.

Author Contributions: All authors contributed in a substantial way to the manuscript. Conceptualization, R.Y. and F.Z.; methodology, R.Y., F.Z. and J.X.; software, R.Y. and C.W.; writing—original draft preparation, R.Y.; writing—review and editing, F.Z., J.X. and C.W.; visualization, R.Y.; supervision, F.Z. and J.X.; project administration, F.Z.; funding acquisition, F.Z. and J.X. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key R&D Program of China (2019YFE0127400); KAKENHI (19K20309).

Acknowledgments: The authors would like to thank Chong Xu and others for providing the inventory of landslides triggered by the 2014 Ms 6.5 Ludian earthquake, which helped us improve this work.

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

Abbreviations

The following abbreviations are used in this manuscript:

CNN	Convolutional neural network
DEM	Digital elevation model
mIoU	Mean intersection over union
MS COCO	Microsoft Common Objects in Context
PSPNet	The Pyramid Scene Parsing Network
R-CNN	Region CNN

References

1. 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. [\[CrossRef\]](#)
2. Manconi, A.; Casu, F.; Ardizzone, F.; Bonano, M.; Cardinali, M.; Luca, C.; Gueguen, E.; Marchesini, I.; Parise, M.; Carmela, V.; et al. Brief Communication: Rapid Mapping of Landslide Events: The 3 December 2013 Montescaglioso Landslide, Italy. *Nat. Hazards Earth Syst. Sci.* **2014**, *14*, 1835–1841. [\[CrossRef\]](#)
3. Hölbling, D.; Füreder, P.; Antolini, F.; Cigna, F.; Casagli, N.; Lang, S. A Semi-Automated Object-Based Approach for Landslide Detection Validated by Persistent Scatterer Interferometry Measures and Landslide Inventories. *Remote Sens.* **2012**, *4*, 1310–1336. [\[CrossRef\]](#)
4. 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. [\[CrossRef\]](#)
5. Duro, D.C.; Franklin, S.E.; Dubé, M.G. A Comparison of Pixel-Based and Object-Based Image Analysis with Selected Machine Learning Algorithms for the Classification of Agricultural Landscapes Using SPOT-5 HRG Imagery. *Remote Sens. Environ.* **2012**, *118*, 259–272. [\[CrossRef\]](#)
6. Wang, H.; Zhang, L.; Yin, K.; Luo, H.; Li, J. Landslide Identification Using Machine Learning. *Geosci. Front.* **2021**, *12*, 351–364. [\[CrossRef\]](#)
7. Tien Bui, D.; Shahabi, H.; Shirzadi, A.; Chapi, K.; Alizadeh, M.; Chen, W.; Mohammadi, A.; Ahmad, B.B.; Panahi, M.; Hong, H.; et al. Landslide Detection and Susceptibility Mapping by AIRSAR Data Using Support Vector Machine and Index of Entropy Models in Cameron Highlands, Malaysia. *Remote Sens.* **2018**, *10*, 1527. [\[CrossRef\]](#)
8. Stumpf, A.; Kerle, N. Object-Oriented Mapping of Landslides Using Random Forests. *Remote Sens. Environ.* **2011**, *115*, 2564–2577. [\[CrossRef\]](#)
9. Zhu, X.; Tuia, D.; Mou, L.; Xia, G.-S.; Zhang, L.; Xu, F.; Fraundorfer, F. Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. *IEEE Geosci. Remote Sens. Mag.* **2017**, *5*, 8–36. [\[CrossRef\]](#)
10. Liu, P.; Wei, Y.; Wang, Q.; Xie, J.; Chen, Y.; Li, Z.; Zhou, H. A Research on Landslides Automatic Extraction Model Based on the Improved Mask R-CNN. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 168. [\[CrossRef\]](#)
11. Romero, A.; Gatta, C.; Camps-Valls, G. Unsupervised Deep Feature Extraction for Remote Sensing Image Classification. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 1349–1362. [\[CrossRef\]](#)
12. Nava, L.; Monserrat, O.; Catani, F. Improving Landslide Detection on SAR Data Through Deep Learning. *IEEE Geosci. Remote Sens. Lett.* **2021**, *19*, 1–5. [\[CrossRef\]](#)
13. Ju, Y.; Xu, Q.; Jin, S.; Li, W.; Su, Y.; Dong, X.; Guo, Q. Loess Landslide Detection Using Object Detection Algorithms in Northwest China. *Remote Sens.* **2022**, *14*, 1182. [\[CrossRef\]](#)

14. Mohan, A.; Kumar, B.; Dwivedi, R. Review on Remote Sensing Methods for Landslide Detection Using Machine and Deep Learning. *Trans. Emerg. Telecommun. Technol.* **2021**, *32*, e3998. [\[CrossRef\]](#)
15. Jiang, W.; Xi, J.; Li, Z.; Ding, M.; Yang, L.; Xie, D. Landslide Detection and Segmentation Using Mask R-CNN with Simulated Hard Samples. *Geomat. Inf. Sci. Wuhan Univ.* **2021**. [\[CrossRef\]](#)
16. Yi, Y.; Zhang, W. A New Deep-Learning-Based Approach for Earthquake-Triggered Landslide Detection From Single-Temporal RapidEye Satellite Imagery. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 6166–6176. [\[CrossRef\]](#)
17. Zhu, Q.; Chen, L.; Hu, H.; Xu, B.; Zhang, Y.; Li, H. Deep Fusion of Local and Non-Local Features for Precision Landslide Recognition. *arXiv* **2020**, arXiv:2002.08547.
18. Qi, W.; Wei, M.; Yang, W.; Xu, C.; Ma, C. Automatic Mapping of Landslides by the ResU-Net. *Remote Sens.* **2020**, *12*, 2487. [\[CrossRef\]](#)
19. Ji, S.; Dawen, Y.; Shen, C.; Li, W.; Xu, Q. Landslide Detection from an Open Satellite Imagery and Digital Elevation Model Dataset Using Attention Boosted Convolutional Neural Networks. *Landslides* **2020**, *17*, 1337–1352. [\[CrossRef\]](#)
20. Cheng, L.; Li, J.; Duan, P.; Wang, M. A Small Attentional YOLO Model for Landslide Detection from Satellite Remote Sensing Images. *Landslides* **2021**, *18*, 2751–2765. [\[CrossRef\]](#)
21. Liu, P.; Wei, Y.; Wang, Q.; Chen, Y.; Xie, J. Research on Post-Earthquake Landslide Extraction Algorithm Based on Improved U-Net Model. *Remote Sens.* **2020**, *12*, 894. [\[CrossRef\]](#)
22. Bragagnolo, L.; Rezende, L.R.; da Silva, R.V.; Grzybowski, J.M.V. Convolutional Neural Networks Applied to Semantic Segmentation of Landslide Scars. *Catena* **2021**, *201*, 105189. [\[CrossRef\]](#)
23. Zhang, H.; Cisse, M.; Dauphin, Y.N.; Lopez-Paz, D. Mixup: Beyond Empirical Risk Minimization. *arXiv* **2018**, arXiv:1710.09412.
24. Ghorbanzadeh, O.; Meena, S.; Blaschke, T.; Aryal, J. UAV-Based Slope Failure Detection Using Deep-Learning Convolutional Neural Networks. *Remote Sens.* **2019**, *11*, 2046. [\[CrossRef\]](#)
25. Ahmad, H.; Ningsheng, C.; Rahman, M.; Islam, M.M.; Pourghasemi, H.R.; Hussain, S.F.; Habumugisha, J.M.; Liu, E.; Zheng, H.; Ni, H.; et al. Geohazards Susceptibility Assessment along the Upper Indus Basin Using Four Machine Learning and Statistical Models. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 315. [\[CrossRef\]](#)
26. Azarafza, M.; Azarafza, M.; Akgün, H.; Atkinson, P.M.; Derakhshani, R. Deep Learning-Based Landslide Susceptibility Mapping. *Sci. Rep.* **2021**, *11*, 24112. [\[CrossRef\]](#)
27. Pourghasemi, H.R.; Rahmati, O. Prediction of the Landslide Susceptibility: Which Algorithm, Which Precision? *Catena* **2018**, *162*, 177–192. [\[CrossRef\]](#)
28. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask R-CNN. In Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017.
29. Yun, S.; Han, D.; Chun, S.; Oh, S.J.; Yoo, Y.; Choe, J. CutMix: Regularization Strategy to Train Strong Classifiers With Localizable Features. In Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea, 27 October–2 November 2019.
30. Shi, Z.-M.; Xiong, X.; Peng, M.; Zhang, L.; Xiong, Y.; Chen, H.-X.; Zhu, Y. Risk Assessment and Mitigation for the Hongshiyuan Landslide Dam Triggered by the 2014 Ludian Earthquake in Yunnan, China. *Landslides* **2016**, *14*, 269–285. [\[CrossRef\]](#)
31. Xu, C.; Xu, X.; Lingling, S.; Dou, S.; Wu, S.; Tian, Y.; Li, X. Inventory of Landslides Triggered by the 2014 MS6.5 Ludian Earthquake and Its Implications on Several Earthquake Parameters. *Seismol. Geol.* **2014**, *36*, 1186–1203.
32. Wu, W.; Xu, C.; Wang, X.; Tian, Y.; Deng, F. Landslides Triggered by the 3 August 2014 Ludian (China) Mw 6.2 Earthquake: An Updated Inventory and Analysis of Their Spatial Distribution. *J. Earth Sci.* **2020**, *31*, 853–866. [\[CrossRef\]](#)
33. Chang, Z.; Chen, X.; An, X.; Cui, J. Contributing Factors to the Failure of an Unusually Large Landslide Triggered by the 2014 Ludian, Yunnan, China, Ms = 6.5 Earthquake. *Nat. Hazards Earth Syst. Sci.* **2016**, *16*, 497–507. [\[CrossRef\]](#)
34. Soares, L.P.; Dias, H.C.; Grohmann, C.H. Landslide Segmentation with U-Net: Evaluating Different Sampling Methods and Patch Sizes. *arXiv* **2020**, arXiv:2007.06672.
35. Tian, Y.; Xu, C.; Chen, J.; Hong, H. Spatial Distribution and Susceptibility Analyses of Pre-Earthquake and Coseismic Landslides Related to the 6.5 Earthquake of 2014 in Ludian, Yunan, China. *Geocarto Int.* **2016**, *32*, 978–989. [\[CrossRef\]](#)
36. Zhou, J.; Lu, P.; Hao, M. Landslides Triggered by the 3 August 2014 Ludian Earthquake in China: Geological Properties, Geomorphologic Characteristics and Spatial Distribution Analysis. *Geomat. Nat. Hazards Risk* **2016**, *7*, 1219–1241. [\[CrossRef\]](#)
37. Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. In Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014.
38. Girshick, R. Fast R-CNN. In Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 7–13 December 2015.
39. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 1137–1149. [\[CrossRef\]](#) [\[PubMed\]](#)
40. Lin, T.-Y.; Dollár, P.; Girshick, R.; He, K.; Hariharan, B.; Belongie, S. Feature Pyramid Networks for Object Detection. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017.
41. Tian, Y.; Xu, C.; Xu, X.; Chen, J. Detailed Inventory Mapping and Spatial Analyses to Landslides Induced by the 2013 Ms 6.6 Minxian Earthquake of China. *J. Earth Sci.* **2016**, *27*, 1016–1026. [\[CrossRef\]](#)

42. Xu, C.; Xu, X.; Yao, X.; Dai, F. Three (Nearly) Complete Inventories of Landslides Triggered by the May 12, 2008 Wenchuan Mw 7.9 Earthquake of China and Their Spatial Distribution Statistical Analysis. *Landslides* **2014**, *11*, 441–461. [[CrossRef](#)]
43. Gorum, T.; Fan, X.; van Westen, C.J.; Huang, R.Q.; Xu, Q.; Tang, C.; Wang, G. Distribution Pattern of Earthquake-Induced Landslides Triggered by the 12 May 2008 Wenchuan Earthquake. *Geomorphology* **2011**, *133*, 152–167. [[CrossRef](#)]
44. Pham, B.; Pradhan, B.; Bui, D.; Prakash, I.; Dholakia, M.B. A Comparative Study of Different Machine Learning Methods for Landslide Susceptibility Assessment: A Case Study of Uttarakhand Area (India). *Environ. Modell. Softw.* **2016**, *84*, 240–250. [[CrossRef](#)]
45. Trigila, A.; Iadanza, C.; Esposito, C.; Scarascia-Mugnozza, G. Comparison of Logistic Regression and Random Forests Techniques for Shallow Landslide Susceptibility Assessment in Giampilieri (NE Sicily, Italy). *Geomorphology* **2015**, *249*, 119–136. [[CrossRef](#)]
46. Stehman, S.V. Selecting and Interpreting Measures of Thematic Classification Accuracy. *Remote Sens. Environ.* **1997**, *62*, 77–89. [[CrossRef](#)]
47. Garcia, A.; Orts, S.; Oprea, S.; Villena Martinez, V.; Rodríguez, J. A Review on Deep Learning Techniques Applied to Semantic Segmentation. *arXiv* **2017**, arXiv:1704.06857.
48. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *arXiv* **2015**, arXiv:1505.04597.
49. Zhao, H.; Shi, J.; Qi, X.; Wang, X.; Jia, J. Pyramid Scene Parsing Network. *arXiv* **2017**, arXiv:1612.01105.