



Article DRs-UNet: A Deep Semantic Segmentation Network for the Recognition of Active Landslides from InSAR Imagery in the Three Rivers Region of the Qinghai–Tibet Plateau

Ximing Chen^{1,2,3}, Xin Yao^{1,2,*}, Zhenkai Zhou^{1,2,4}, Yang Liu^{1,2,3}, Chuangchuang Yao^{1,2} and Kaiyu Ren^{1,2}

- ¹ Institute of Geomechanics, Chinese Academy of Geological Sciences, Beijing 100081, China; chenxm@cug.edu.cn (X.C.); zzk114043@cugb.edu.cn (Z.Z.); 2202010444@cug.edu.cn (Y.L.); yaochuangchuang19@mails.ucas.ac.cn (C.Y.); 1202021077@cug.edu.cn (K.R.)
- ² Key Laboratory of Active Tectonics and Geological Safety, Ministry of Natural Resources, Beijing 100081, China
- ³ School of Geography and Information Engineering, China University of Geoscience (Wuhan), Wuhan 430074, China
- ⁴ School of Engineering and Technology, China University of Geosciences (Beijing), Beijing 100083, China
- * Correspondence: yaoxin@mail.cgs.gov.cn

Abstract: At present, Synthetic Aperture Radar Interferometry (InSAR) has been an important technique for active landslides recognition in the geological survey field. However, the traditional interpretation method through human-computer interaction highly relies on expert experience, which is time-consuming and subjective. To solve the problem, this study designed an end-to-end semantic segmentation network, called deep residual shrinkage U-Net (DRs-UNet), to automatically extract potential active landslides in InSAR imagery. The proposed model was inspired by the structure of U-Net and adopted a residual shrinkage building unit (RSBU) as the feature extraction block in its encoder part. The method of this study has three main advantages: (1) The RSBU in the encoder part incorporated with soft thresholding can reduce the influence of noise from InSAR images. (2) The residual connection of the RSBU makes the training of the network easier and accelerates the convergency process. (3) The feature fusion of the corresponding layers between the encoder and decoder effectively improves the classification accuracy. Two widely used networks, U-Net and SegNet, were trained under the same experiment environment to compare with the proposed method. The experiment results in the test set show that our method achieved the best performance; specifically, the F1 score is 1.48% and 4.1% higher than U-Net and SegNet, which indicates a better balance between precision and recall. Additionally, our method has the best *IoU* score of over 90%. Furthermore, we applied our network to a test area located in Zhongxinrong County along Jinsha River where landslides are highly evolved. The quantitative evaluation results prove that our method is effective for the automatic recognition of potential active landslide hazards from InSAR imagery.

Keywords: landslide; semantic segmentation; InSAR; automatic recognition

1. Introduction

The rapid identification of potential active landslides in a wide range is of great significance to prevent destructive geological hazards. Traditionally, landslides recognition is conducted through visual interpretation on optical remote sensing images and verification with field investigation, which is based on interpreters' expert knowledge of and experience with field surveys, so it is time-consuming and costly [1]. Therefore, it is urgent to develop automatic methods for identifying the location and distribution of landslides in practical application. Then a potential active landslides inventory could be established as the database for subsequent prevention efforts and time–series monitoring of a specific single landslide.



Citation: Chen, X.; Yao, X.; Zhou, Z.; Liu, Y.; Yao, C.; Ren, K. DRs-UNet: A Deep Semantic Segmentation Network for the Recognition of Active Landslides from InSAR Imagery in the Three Rivers Region of the Qinghai–Tibet Plateau. *Remote Sens.* **2022**, *14*, 1848. https://doi.org/ 10.3390/rs14081848

Academic Editors: Qiang Xu, Shunping Ji and Wanchang Zhang

Received: 14 March 2022 Accepted: 7 April 2022 Published: 12 April 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/).

In recent years, a wide variety of remote sensors has been used for landslide hazards studies, such as optical remote sensing imagery, SAR (Synthetic Aperture Radar), LiDAR (Light Detection and Ranging), and so on [2]. Among them, the InSAR technology based on SAR data, with its advantages of all-weather, wide range, and high precision, has become one of the important means to study surface deformation disasters [3]. Landslide identification using remote sensing imagery is an economical and efficient method. There have been many studies using optical remote sensing images and InSAR for semi-automatic or automatic landslide extraction [4,5]. Landslide recognition based on remote sensing images can be divided into pixel-based and object-based methods. After the occurrence of a landslide, the difference in spectral characteristics between landslide pixels and nonlandslide pixels will increase significantly, on which the traditional pixel-based landslide recognition methods are based. However, pixel-based methods only make use of the spectral information of individual pixels, lacking consideration of the correlation between adjacent pixels, so pixel-based classification methods usually cannot attain ideal performance [1]. Object-based methods first segment imagery into different objects based on homogeneous pixels and then further classify them using spectral, textural, and spatial contextual features. Compared to pixel-based methods, object-based methods consider more information, so they usually have better performance in classification tasks [6,7]. In terms of landslide extraction research, pixel-based and object-based methods are often combined with traditional machine learning algorithms, such as artificial neural networks (ANN) [8,9], random forest (RF) [10,11], support vector machine (SVM) [12,13], and so on.

With the rapid development of remote sensing technology, masses of accumulated data are available for study purposes. For example, the Sentinel-1 satellite of the European Spatial Agency (ESA) has a 12-day satellite revisit period, generating a massive amount of open-source data every day. Nevertheless, conventional machine learning algorithms need an onerous handcraft feature extraction process, which is generally time-consuming, and they can neither fully extract nor effectively utilize features from such big data. How to make full use of big data and mine useful information becomes a problem demanding a prompt solution in the field of earth science [14]. In recent years, deep learning (DL) has been gradually applied to many disciplines and has become the state-of-the-art method in many cases, such as medical images comprehending, land use and land cover (LULC) classification, road extraction, landslide extraction, and so on [15]. Deep learning is a subset of machine learning (ML), and it is an extension algorithm of artificial neural networks, but it has more complex architecture and neural layers; the characteristics of DL make it better than ML at mining features from big data automatically. The applications of DL in the research of landslides hazards can be divided into three categories: landslide recognition, landslide susceptibility mapping, and landslide displacement prediction [16].

Landslide recognition is a hot research topic recently. In existing studies, high spectral and spatial resolution remote sensing imagery are the primary data for deep learning-based landslide recognition [17–21]. In optical remote sensing images, the spectral and textural features change relatively significantly before and after the occurrence of landslides, making it easier to distinguish them from the environment. Ding et al. applied a six-layer convolutional neural network (CNN) for landslide extraction using GF-1 remote sensing images before and after the landslide occurred and achieved a recall of 72.6% [22]. Omid et al. compared the performance of several conventional machine learning algorithms and convolutional neural networks in landslide recognition, analyzed the influence of the depth of CNN, the size of input data, and topographic factors of the landslide recognition accuracy. The experiment results indicated that when only spectral information was utilized, a four-layer depth CNN with an input image size of 16×16 pixels could achieve the best mean Intersection over Union of 78.26% [8]. Ji et al. produced an open-source landslide dataset with 0.8 m resolution to remedy the problem of lacking enough labeled data for deep learning applications in landslide extraction. The attention mechanism is an important milestone in the development of deep learning model structures. In order to extract distinctive feature representations of landslides from complicated backgrounds, a new 3D SpatialChannel Attention Module (3D SCAM) was proposed. According to the experiments, the proposed 3D SCAM performed better than the SE module (Squeeze-and-Excitation module) [23], BAM (Bottleneck Attention Module) [24], and CBAM (Convolutional Block Attention Module) [25]. The ResNet-50 embedded with 3D SCAM obtained the best results in the test data with an *F*1 score of 96.62% [26]. U-Net [27] was originally proposed for medical image segmentation, but many studies have shown that it is also a robust baseline model for target extraction in remote sensing images. There is a series of studies based on improved U-Net models to automatically extract landslides, and satisfactory results have been obtained [9,28–31].

Landslide recognition can be divided into historical landslide recognition and early landslide recognition [32]. However, the studies mentioned above all focused on historical landslides. Although InSAR has been widely used for active landslide identification, there is a great lack of research on active landslide identification combining deep learning with InSAR [33–35]. Therefore, for this study, we chose areas around the Lancang River and the Jinsha River to explore the feasibility of automatic active landslides recognition from InSAR imagery. With reference to the structure of U-Net, a deep semantic segmentation model called deep residual shrinkage U-Net (DRs-UNet) embedded with a residual shrinkage building unit [36] was proposed. Semantic segmentation is a typical application in computer vision that aims to annotate every pixel within imagery with a specific semantic label. Through semantic segmentation methods, people can acquire both the location and extent of a landslide. For an input of InSAR deformation phase imagery, our pipeline extracts feature maps by eight consecutively stacked RSBU blocks (residual shrinkage building unit blocks). The RSBU-block restrains noise from feature maps by an automatically learned threshold, and its residual connection structure can solve the degradation problem during the training process as the network deepens. Furthermore, two widely applied models, i.e., U-Net, and SegNet [37] were selected to compare with the proposed method. In addition, we applied our method to an independent test area located in the Zhongxinrong County section of the Jinsha River to evaluate the performance.

There have been many studies concerning the automatic extraction of historical landslides in optical remote sensing imagery using deep learning-based methods but few works concentrating on recognition of active landslides. The main goal of this study is to demonstrate that the proposed DRs-UNet could effectively extract active landslides from InSAR deformation phase imagery.

2. Study Area

As shown in Figure 1, the study area is located in the middle and upper reaches of the Lancang River and the middle reaches of the Jinsha River in the southeastern part of the Qinghai–Tibet Plateau. The area is in the Hengduan Mountains where the transition zone of China's first and second terrain grade is, with an average elevation of 4350 m. The wet season starts in June and ends in October, when over 80% of the annual rainfall occurs. The study area lies in the collision zone between the Indian Plate and the Eurasian Plate; with complex geological conditions, strong seismic activities here lead to the frequent occurrence of landslides. Because of ample hydropower resources, a large number of hydroelectric power plants have been built and planned for construction. Potential active landslides pose a serious threat to the security of residents' lives and property, as well as the normal operation of hydropower facilities. A number of studies have been conducted in this area, including early identification of landslides and spatial analysis of their distribution in the "Three Rivers" (Nujiang River, Lancang River, and Jinsha River) regions [38,39] and active landslides inventory mapping along the Jinsha River corridor [40].



Figure 1. Geographical location map of the research area. The research area is comprised of a study area (cyan frame) and an independent test area (red frame). Red points represent active landslides through manual interpretation.

3. Data Preparation

3.1. InSAR Data

In this study, 30 SAR images of Sentinel-1 (https://scihub.copernicus.eu/dhus/#/ home (accessed on 1 May 2020)) for each of the ascending and descending orbits were used for D-InSAR (Differential Interferometry Synthetic Aperture Radar) processing. We built D-InSAR imagery pairs at 12-day and 24-day intervals of both ascending and descending tracks. Then, the InSAR phase imagery was averaged over the whole period, which was expected to weaken the interference of random noise [41]. In addition, high-resolution optical remote sensing images, geological maps, and topographic data were selected to comprehensively interpret active landslides. Based on the interpreting results, the dataset was produced to train landslide extraction models.

InSAR technology can realize large-scale, high-precision, and all-weather surface deformation measurement and has become the main technology of surface deformation hazards research. D-InSAR is developed based on InSAR; two or more SAR images spanning the deformation period are used for interferometric processing to obtain the initial interferogram consisting of atmospheric phase, topographic phase, deformation phase, etc. Then, the phase contributions other than the surface deformation phase are separated, and, finally, the surface deformation phase is converted to the deformation displacement in the line-of-sight direction to obtain the displacement during the two imaging moments [42,43]. The SAR data used for this study were provided by ESA's open-source Sentinel-1 satellite, and the relevant parameters of the data are shown in Table 1. D-InSAR was conducted using the GAMMA Software; the process flow is presented in Figure 2.

SAR Sensor	Waveband (cm)	Direction	Spatial Resolution (m)	Incidence Angle (°)	The Heading Angle (°)	Number of Scenes	Polarization	Temporal Coverage
Sentinel-1	C (5.63 cm)	Ascending	5 by 20	38.5°	-12.6	30	VV	2019.3–2020.3
Sentinel-1	C (5.63 cm)	Descending	5 by 20	40.6°	129.6	30	VV	2019.3–2020.3

Table 1. Parameters of radar image data used in the study.



Figure 2. D-InSAR process flow.

Given that the purpose of this study was to rapidly and comprehensively identify the location and extent of active landslides, accurate calculation of displacement is not necessary, so interferograms were not converted to displacement velocity maps. Figure 3 shows the D-InSAR processing results of ascending and descending tracks in a sub-region of the study area, respectively. It is worth noting that some phase patterns caused by glacier movement and snowmelt, mainly in high mountains, are very similar to those caused by landslides displacement, such as the phase patches marked by black frames in Figure 3, which are mainly distributed above the snowline.



Figure 3. Example of D-InSAR results. Black boxes in (**a**) ascending and (**b**) descending maps indicate the noise phase caused by glacier movement and snowmelt.

3.2. Data Preprocessing

The first part of Figure 5 represents the data preprocessing. First of all, the active landslides were manually interpreted and outlined on the InSAR interferograms using the Environmental Systems Research Institute's (ESRI) ArcGIS software to obtain binary masks, as shown in Figure 4b, in which white and black pixels belong to landslides and no-landslides, respectively. It is worth noting that the raw input data in this study are three-channel fusion imagery of the D-InSAR phase and SAR intensity maps. Due to the limitation of computer memory conditions, the size of input data could not be too large, so the data needed to be cut into small patches. Furthermore, considering that the performance of the deep learning model depended heavily on the size of the training dataset [44], he data were cut with a 50% overlap to enlarge the dataset. Figure 4 demonstrates the process of image splitting.



Figure 4. Data production. (a) InSAR imagery, (b) landslides binary map, (c) image patches.

A total of 1712 images of 128×128 pixels were collected. Then, the samples were normalized using the Z-score method so that each sample had a mean pixel value of 0 and a variance of 1. The objective of normalization was to facilitate the training process of the deep learning models. Finally, according to the ratio of 7:2:1, the dataset was divided into the training, validation, and test sets. Note that the training and validation sets were used to train the model and choose optimal model parameters, respectively, while the test set was to quantitatively evaluate the performance of a model.

4. Method

Figure 5 shows the flowchart of this study. The first stage was data preprocessing; with the help of geological experts, active landslides were labeled using InSAR imagery combined with multi-source remote sensing data and geological data. Then, the annotated data were sliced into image patches and normalized and were further split into three data

sets, i.e., training, validation, and test sets. Next, models were constructed and trained. At the stage of accuracy evaluation, the trained models were quantitatively evaluated by precision, recall, *F*1 score, and *IoU* (Intersection over Union). Finally, in practical application, the deformation phase caused by snow melt and glacier movement is similar to that which is caused by landslide displacement. Therefore, an elevation threshold should be determined according to some factors, e.g., snowline, ELA (equilibrium-line altitude) and so on, of a specific area to mask out useless objects and optimize the recognition results.



Figure 5. Flowchart of automatic extraction of active landslides from InSAR imagery.

4.1. Deep Learning and Semantic Segmentation

The task of semantic segmentation is to annotate every pixel in the image with a semantic label. In remote sensing and earth observation, semantic segmentation has been widely applied to land cover and land use classification, change detection, landslide recognition, etc. [9,22,26,45–47]. For instance, the method of the extraction of landslides from remote sensing imagery based on the semantic segmentation method is to label pixels belonging to landslides and non-landslides differently (e.g., "1" for landslides, "0" for non-landslides). Long et al. [48] proposed one of the first deep learning works for semantic image segmentation, using a fully convolutional network (FCN) including only convolution layers. FCN replaces the dense connection layers in the traditional CNN model with convolutional layers so that it not only accepts input images of arbitrary size but also greatly reduces model parameters, and the segmented image of the same size as the input image is output after upsampling by deconvolution layers. Since FCN implements an endto-end approach to training models that can be applied to dense prediction tasks for images of arbitrary size, it has become a cornerstone for the development of subsequent deep learning semantic segmentation models [49]. In recent years, with the development of deep learning semantic segmentation techniques, the encoder-decoder structure has become the main subject of research [44]. The design of the encoder-decoder structure compensates for the defect that FCN does not take into account global contextual information. The encoder-decoder network is composed of two parts: an encoder extracting potential feature representations hierarchically from raw images through stacked convolution layers and a decoder-predicting probability map at the pixel level [50]. The encoding and decoding parts are generally symmetrical so that the information can flow from the encoder to the corresponding layer in the decoder for better integration of high-level and low-level features. SegNet, U-Net, HRNet [51], and LinkNet [52] all take such a structure. In terms of deep learning-based landslide recognition, U-Net and ResNet [53] are the most commonly used baselines, and a series of related studies have shown that they are easy to train and have high robustness [54].

4.2. Proposed DRs-UNet

The deep residual shrinkage network (DRSN) is a variant of ResNet, which was originally proposed by Zhao et al. [36] for fault diagnosis of mechanical transmission systems. Conventional CNNs suffer from the vanishing gradients problem as the network deepens, making the training process difficult. To solve this problem, He et al. [53] proposed the ResNet, composed of a series of residual modules with shortcut connections. As shown in Figure 6, the output feature maps of the residual module are connected to the input feature maps through an identity shortcut. Suppose the input feature of the residual module is x_l , the output feature of the residual branch is $F(x_l)$, and the final output feature x_{l+1} equals the element-wise addition of x_l and $F(x_l)$: $F(x_l) + x_l$. The forward propagation can be expressed as Equation (1):

$$x_{l+1} = F(x, w_l) + x_l$$
 (1)

where w_l is the weight and $F(x, w_l)$ is the residual mapping to be learned by the network. There are two forms of residual modules: the plain residual module containing two convolution layers, which were used in ResNet-18 and ResNet-34, as shown in Figure 6a, and the Bottleneck Residual Block, including three convolution layers, as Figure 6b represents, which deeper networks adopt, i.e., ResNet-50 and ResNet-152. In this study, the backbone of our proposed method builds on the plain residual block.



Figure 6. Residual block, (a) plain residual block, (b) bottleneck residual block.

Although ResNet can effectively solve the network degradation problem, the learning ability of the model is greatly weakened when dealing with signals of high noise. Soft thresholding is a key step in many traditional denoising methods. In general, the raw signal is transformed to a domain in which the near-zero numbers are unimportant, and then soft thresholding is applied to convert the near-zero features to zeros [36]. The function of soft thresholding can be expressed as follows:

$$y = \begin{cases} x - \tau, x > \tau \\ 0, -\tau \le x \le \tau \\ x + \tau, x < -\tau \end{cases}$$
(2)

where *y* is the output feature maps, *x* is the input feature maps, and τ is the threshold. As represented in Figure 7c, soft thresholding is embedded in the RSBU-block. In the residual branch of the plain residual block, the output feature map from two stacked convolution layers is converted to absolute values and then passed to a GAP (Global Average Pooling) layer to obtain feature |x|. Next, the feature |x| is put into a two-layer fully connected layer that is subjected to batch normalization (BN) and linear rectification function operations. Then, the feature is scaled to (0, 1) by a sigmoid function to obtain the feature α . Finally, the threshold t is calculated through the element-wise multiplication of α and |x|, and then

soft thresholding is applied to the feature *x*. The final output of the RSBU-block is the element-wise summation of the raw input feature and the feature *x* that has been denoised by the soft thresholding. The calculation of the threshold in the RSBU-block is similar to the attention calculation method of the SENet [23], where the global features obtained by the GAP are used to learn the weights between different channels, and then the weights are used to multiply the input feature maps to obtain the threshold. Using the learned threshold, the features are filtered using soft thresholding (Equation (2)) to suppress noise and redundant information. The threshold is automatically learned via gradient descent algorithm. The derivative of soft thresholding can be expressed as Equation (3):

$$\frac{\partial y}{\partial x} = \begin{cases} 1, \ x > \tau \\ 0, -\tau \le x \le \tau \\ 1, x < -\tau \end{cases}$$
(3)



Figure 7. Schematic of the proposed DRs-UNet: (**a**) overall structure of DRs-UNet; (**b**) initial block; (**c**) inner core structure of RSBU-block; (**d**,**e**) inner core structure of transposed convolution block (TCB) and the final convolution block.

There are two main difficulties in identifying active landslides from InSAR images using CNNs. First, the phase patches caused by the deformation of some non-active landslides, such as snowmelt, are very similar to those of active landslides. Second, subject to atmospheric delay, vegetation coverage, and spatial-temporal decoherence, a large amount of impulse noise is generated when performing large-scale InSAR phase calculations [55]. These may reduce the overall precision of landslide extraction. To address the problem, an encoder-decoder network, DRs-UNet, is designed with reference to the model structure of U-Net to achieve end-to-end active landslide extraction. DRs-UNet contains an encoder and a decoder. The encoding part downsamples the input data and

extracts potential feature representations through a series of stacked convolutional layers and pooling layers. The decoder restores the resolution of the feature maps to finally obtain the segmentation map.

Figure 7a demonstrates that the backbone of the encoder part is based on a deep residual shrinkage network (DRSN), which has been demonstrated to be applicable to the process of highly noised images by several studies [56–59]. The DRSN in this study was rebuilt from the ResNet-18. Specifically, the encoder consists of an initial module and eight stacked RSBU-blocks. The initial module consists of a convolutional layer with a kernel size of 2 and a pooling layer whose output feature maps are sequentially subjected to a BN operation and a ReLU nonlinear activation function. Figure 7c represents the inner core structure of the RSBU-block whose forward propagation was elaborated earlier in this section. The RSBU-block is improved through residual modules using the soft thresholding method, which can not only effectively prevent the vanishing and exploding gradients problem caused by the deepening of the CNN and accelerate the training and convergence of the model but also can attenuate the influence of noise from features to the network and reduce the redundant information. Thus, the classification accuracy is improved. The encoder downsamples features using four blocks so that the size of the feature maps is reduced to 1/8 of the original images. In the decoding stage, the features extracted by the encoder are first fed into four transpose convolution blocks (TCB) to be upsampled. As shown in Figure 7d, the TCB consists of a transposed convolutional layer with a kernel size of 3 and a stride of 2 and then a convolutional layer with kernel size and stride of 3 and 1, respectively. As shown in Figure 7a, the features of each RSBU-block are connected to those of the corresponding TCB by skip connection (i.e., the feature maps of larger size are cropped and then concatenated to the smaller size features in the channel dimension). During the convolution and pooling operations in the encoding stage, a large amount of spatial information was lost, so the lower layers (layers near the input image) of the CNN retained more spatial information. By fusing low-level and high-level features through skip connection, the accuracy of segmentation can be improved. All the convolution layers in the decoding path are subjected to a BN and ReLU layer. After passing through four repeated TCBs, the feature map progressively increases dimensionally, and the number of channels decreases gradually. Finally, there is a convolutional layer with both a kernel size and stride of 1, then a sigmoid layer is used to predict the probability map with the same spatial size as the original input image. The closer the pixel value is to 1, the more likely the pixel is to be a landslide feature, and a pixel value of 0 means a non-landslide feature.

5. Results and Analysis

5.1. Experiment Settings and Evaluation Criteria

The experiments were implemented in a computer with the configuration of Intel Core i7-11800H CPU (2.30 GHz) and NVIDIA GeForce RTX 3060. In this article, Pytorch (GPU version) was adopted as the deep learning framework. CUDA 11.0 and CuDNN 8.2 were selected for GPU parallel calculation and acceleration, respectively. The Adam optimizer was used to update and optimize the parameters. The initial learning rate of training was set to 1×10^{-3} for better training of the model. The learning rate was dynamically adjusted as the number of training epochs increased, which is presented in Figure 8. The learning rate decreases per 10 epochs by a ratio of 0.5.

The BCE Loss (Binary Cross-Entropy Loss) was chosen as the loss function for training, which is defined as follows:

$$loss = -\frac{1}{N} \sum_{i=1}^{N} y_i \cdot \log a + (1 - y_i) \cdot \log(1 - a)$$
(4)

$$a = p(y_i = 1|x_i) \tag{5}$$

where *N* denotes the number of classification categories, and *a* represents the probability that the label category of x_i is $y_i = 1$. During the training process, the batch size and

the number of epochs were set to 16 and 80, respectively. The default 0.5 was chosen as the threshold to classify the predicted probability maps into the landslide and non-landslide targets.





The confusion matrix for the binary classification task in this paper is shown in Table 2. In this article, TP (true positive) denotes correctly classified landslide targets. TN (true negative) indicates that the model correctly predicts the non-landslide target. FP (false positive) denotes non-landslide targets that are incorrectly predicted as landslides. FN (false negative) indicates that the model incorrectly classified the landslide target.

Table 2. Confusion matrix for landslides extraction task.

		Prediction		
		Landslide	Non-Landslide	
Ground Truth	Landslide Non-landslide	TP FP	FN TN	

Several quantitative evaluation metrics, i.e., precision, recall, *F*1 score, and Intersection over Union (*IoU*) were used to evaluate the performance of different models. Precision represents the proportion of the samples that are correctly classified within the samples predicted to be positive. Recall represents the correct pixels over the labels. The *F*1 score is the harmonic mean of precision and recall, which is a more balanced metric that takes into account both precision and recall. *IoU* represents the intersection over the union of the ground truth and the model prediction. Precision, recall, *F*1 score, and *IoU* are expressed as follows:

$$Precision = \frac{TP}{TP + FP} \tag{6}$$

$$Recall = \frac{TP}{TP + FN}$$
(7)

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}$$
(8)

$$IoU = \frac{TP}{TP + FP + FN}$$
(9)

5.2. Experiment Results

The U-Net and Vgg16-based [60] SegNet, which are widely used in semantic segmentation tasks, were trained under the same experimental settings to compare with the designed DRs-UNet model. Figure 9a-c summarizes the training and validation loss curves of DRs-Unet, U-Net, and SegNet, respectively, and Figure 9d-f shows the F1 score with the change of epochs. It can be seen that the loss and F1 score curves of all models become smooth after about 60 epochs, and the models reach convergence as the epochs continue to increase. It is worth noting that the loss values of DRs-UNet are lower than those of the other two models after convergence. The probable reason is that the RSBU-block could better solve the vanishing gradient problem. Furthermore, the F1 score curves of the proposed method during both training and validation are consistently higher than those of other comparative models as the epoch increases, which indicates that DRs-UNet has a better balance between precision and recall and achieves the best accuracy.



Figure 9. The loss and f1 score of DRs-UNet, U-Net, and SegNet from left to right, respectively; (**a**–**c**) loss curve during training and validation, (**d**–**f**) F1 curve during training and validation.

Table 3 shows the results of the quantitative evaluation of the three models on the test dataset. Given that the topography is an important factor that affects the stability of the ground surface in the mountainous areas, for instance, slope can reflect the steepness of the hillsides in the study area, while curvature can reflect the topography and the complexity of the terrain. Generally speaking, landslides are more likely to occur in areas with greater slope [61–63]. Therefore, we also trained the DRs-UNet using DEM, Slope (Slope), and Curvature together with InSAR imagery separately to analyze how topographic factors affect the accuracy. It can be seen from Table 3 that the DRs-UNet model had the highest score in all metrics, with precision, recall, F1 score, and *IoU* of 96.07%, 96.12%, 96.08%, and 92.48%, respectively. Among them, the *IoU* of U-Net and SegNet were both below 90% on the test set. U-Net performed slightly better than SegNet, which may be attributed to the skip connection of U-Net that incorporates the low-level and high-level features so that the decoding path has more spatial information to improve the segmentation performance.

Model	Data	Precision (%)	Recall (%)	F1 Score (%)	IoU (%)
DRs-UNet	InSAR imagery	96.07	96.12	96.08	92.48
	InSAR imagery + DEM	88.48	95.26	91.70	84.72
	InSAR imagery + Slope	95.18	96.79	95.97	92.27
	InSAR imagery + Curvatures	95.83	97.40	96.61	93.46
U-Net	InSAR imagery	94.02	95.22	94.60	89.77
SegNet	InSAR imagery	90.77	93.24	91.98	85.20

Table 3. Evaluation of the active landslides extraction results on the test dataset.

Note: The values in bold indicate the highest score of the corresponding metrics.

However, the results show that after adding the DEM for training, the overall accuracy of the DRs-UNet model decreased, which may be due to the active landslides in this area having a weak relation to the elevation. The slope factor did not obviously change the overall accuracy; however, the curvature improved the overall performance slightly, with an increase in recall, *F*1 score, and *loU* of 1.28%, 0.53%, and 0.98%, respectively. However, the improvement of the accuracy after adding Curvatures was not significant. In addition, when the model was practically applied for landslide extraction on a large scale, adding topographic factors made the preparation of input data more onerous. Therefore, we chose to apply the model trained only by InSAR imagery in the subsequent work. Figure 10 shows some prediction results of the three models on the test set.



Figure 10. Recognition results of some InSAR patches from the test set; (**a**) input InSAR imagery, (**b**) corresponding ground truth, (**c**–**e**) recognition results of DRs-UNet, U-Net, and SegNet, respectively.

5.3. Results of the Zhongxinrong Test Area

The Zhongxinrong section of the Jinsha River was selected as an independent test area to evaluate different models' performance. In this paper, the time of data acquisition in this area was consistent with the data used for model training. The test area is characterized by strong slope erosion due to the downcutting effect of the river, resulting in a dense distribution of landslides. Due to the limitation of computer memory, it is not possible to input large-scale images directly into the network for inference. Therefore, the sliding window method was adopted. The sliding window extracted 128×128 pixel patches with an overlap ratio locally, and then the patches were fed into the network in mini batches. Finally, the output binary classification maps ("1" for landslide targets, "0" for non-landslide targets) were mosaicked according to their positions on the original map. Figure 11a–c shows the active landslide extraction results of DRs-UNet, U-Net, and SegNet on the ascending track of the area, respectively. To present the recognition results in detail and compare them with recently published results, the detection results of the pre-study [39] were superimposed on the final map (green dots in Figure 11d).



Figure 11. Results of the Zhongxinrong County section of the Jinsha River. The yellow patches in (**a**–**c**) represent the landslides extraction of DRs-UNet, U-Net, and SegNet, respectively. Red polygons in (**a**–**d**) indicate ground truth. Green dots superimposed on (**d**) represent the recognition results of the pre-study [39].

It can be seen from rectangle 1 that the DRs-UNet model works best for the extraction of the landslides that are characterized by weak signals. In addition, the landslide highlighted by rectangle 2 was only recognized by DRs-UNet. Due to the effect of decoherence caused by high vegetation coverage, the deformation signals of the landslide in rectangle 2 are discontinuous with impulse noise (Figure 12). Since DRs-UNet embeds soft thresholding in the RSBU-block, which can automatically learn the threshold, the features below the threshold will be directly set to 0, thus reducing the impact of redundant information and noise on the extraction accuracy. For this reason, DRs-UNet can recognize active landslides with high vegetation coverage from InSAR imagery with higher accuracy. However, the model has some false positive results, as shown in rectangle 3.



Figure 12. The landslide highlighted by rectangle 2 in Figure 11 in (**a**) InSAR imagery and (**b**) optical remote sensing imagery.

To compare the quantitative evaluation criteria of different models in this test area, the *F*1 score and *IoU* were calculated based on the confusion matrix of the extracted results. From the results shown in Table 4, it is known that DRs-UNet has the highest *IoU* and *F*1 score, 79.26% and 88.31%, respectively. Additionally, compared to U-Net and SegNet, DRs-UNet is a more lightweight model with only 17.19 million parameters (65.6 MB) and 19.1 G FLOPs, but it achieved the highest accuracy.

Model	F1 Score (%)	IoU (%)	Parameters (Million)	Size (MB)	FLOPs (G)
DRs-UNet	88.31	79.26	17.19	65.60	19.12
U-Net	81.86	72.18	31.03	118.40	54.68
SegNet	80.97	69.5	29.45	112.32	40.10

Table 4. Quantitative assessment, complexity of different models in the test area.

Note: The values in bold indicate the highest score of the corresponding metrics.

6. Discussions

As the means of collecting data become easier, a large amount of accumulated data is available to be exploited. For instance, the revisit period of a single Sentinel-1 satellite is only 12 days, which can be shortened to 6 days by a tandem formation, producing massive data every day. Using deep learning methods, such big data can be fully exploited to automatically, quickly, and adequately extract ground surface information contained therein. In this study, SAR imagery from Sentinel-1 is used as the raw data to prepare the dataset. A deep semantic segmentation model called DRs-UNet was proposed for the early identification of potential active landslides. The proposed method was compared to U-Net and SegNet, and the experiment results indicated that our method had the highest overall accuracy. Specifically, as presented in Table 3, the *F*1 score of DRs-UNet is 1.48% and 4.1% higher than those of the other two models, and the *IoU* is 2.71% and 7.28% higher.

Figure 10 shows some recognition results of the three models on the test set, and it can be seen that the proposed method extracts more accurately in terms of landslide signals that are relatively weak, as well as small-scale landslides. Moreover, according to the results shown in the sixth row of Figure 10, the method proposed in this study can extract a more complete potential active landslide surrounded by high pepper-and-salt noise, while the results of other two methods are incomplete; the SegNet cannot even identify the landslide.

Furthermore, we applied our method to the Zhongxinrong county section of the Jinsha River. This area is characterized by a strong downcutting effect caused by rivers. The elevation and slope of this region are 2300~3000 m and 20~50 degrees, respectively. The typical landslide in this area are colluvium slopes whose lithology are mainly composed of ophiolite, mudstone, sandstone, siltstone, coal streak, tuff, etc. The *IoU* in this area is 79.26%; it is 7.08% and 9.76% higher than those of the other two comparisons, which demonstrates the feasibility of using our method for the rapid identification of potential active landslide hazards.

Inaccurate samples can affect the training of the network. When labeling samples, because of the similar patterns of some local deformation signals, the labeled target may be an area with surface subsidence, deformation caused by cut fill, etc. but not necessarily a landslide. In addition, visual interpretation relies on expert experience, and the recognition results are rather subjective. For example, even a skillful interpreter is unable to accurately delineate a potential active landslide just from InSAR imagery. Geological, topographic, and optical remote sensing data are usually referenced together with InSAR imagery. Therefore, in practical application, the extent of potential landslides extracted using the proposed method will be somewhat different from the accurate extent. The goal of early identification of potential landslides is to recognize as many as possible, which means a consequence of high false positives is acceptable. The proposed method can improve efficiency for the identification of potential active landslides. Interpreters can filter undesirable targets based on the automatic extraction results, which can immensely reduce costs. Furthermore, time-series monitoring can be conducted on a single landslide subsequently, which can avoid atmospheric effects on large-scale InSAR computation, and it is important for the prevention and management of destructive landslides.

The early identification of active landslide hazards aims to discover landslides that threaten the safety of human lives and property [64]. The phase patterns caused by the melt and movement of glaciers and snow are very similar to those of landslides, which is the main factor that causes false positive results, mostly when applied to large scale areas. In China, the southwest region is the most seriously endangered by landslide hazards. For instance, there are the Hengduan Mountains, located in the northeast of the Qinghai–Tibet plateau where there are complex terrain, large terrain undulation, and widely distributed glaciers. The snow cover is mainly distributed in high-altitude areas, and the areas above the snowline are covered by snow all year round. In 2019, the snow cover fraction (SCF) above a 3000 m elevation is about 50%; because human activity is rare, the snow cover over a 3000 m elevation is stable [65]. The average snow cover days (SCD) exceed 30 per year in the Hengduan Mountains, and the ELA (equilibrium-line altitude) is 5132 m [66,67]. Therefore, when using the proposed method for large-scale landslides identification in this area, the local snow line, ELA, etc., can be taken into account to choose a suitable elevation threshold to filter the extraction results.

Because of the special side-looking imaging geometry of SAR satellites, different observation angles will yield different InSAR results. Furthermore, in mountainous areas with steep terrain, the phenomenon of layover is very likely to occur [68]. On the one hand, a landslide may be detected either on AT (ascending track) or DT (descending track) but cannot be detected on both tracks. Therefore, in order to avoid the omission of potential landslides, the recognition results of AT and DT are fused. On the other hand, sometimes a landslide may be extracted both on AT and DT; when the extractions of AT and DT are intersected, we count these objects as the same deformation source. Specifically, as shown in Figure 13, if an element in the extraction result of the AT intersects with one or more



elements in the DT, these elements are counted as the same deformation source, and the one with the largest area element is retained as the fusion result of the AT and DT.

Figure 13. The example of fusing extractions of AT and DT: (**a**) superposition of AT (red) and DT (green) recognition results, (**b**) the fusion result.

7. Conclusions

In this study, an improved deep semantic segmentation model was proposed to automatically identify potential active landslides. Given that the InSAR imagery contains high noise, it can cause a lot of redundant information and influence the learning efficiency of the convolution layers. In order to address this problem, the RSBU-block was embedded in the proposed network. On the one hand, DRs-UNet can effectively improve the discriminative feature learning ability from features with high noise, and the residual connection can promote the convergency during the training process. On the other hand, the skip connection fuses the high- and low-level features and restores the spatial information lost during the decoding path, which can effectively promote the classification performance. Two networks, i.e., U-Net and SegNet, were adopted as comparisons. To verify the efficiency of the proposed model, besides quantitative evaluation on the test set, an independent area known to have high landslide density was used to test the model. According to the statistic evaluations, DRs-UNet has higher overall accuracy and robustness, and also shows significant advantages in terms of parameter size and computational efficiency.

Recently, in the field of geological hazards, deep learning-based methods have been state-of-the-art in many cases, which can immensely improve the automation and accuracy of geological disasters identification. There have been many studies using deep learning methods to automatically extract historical landslides. However, the automatic recognition of potential active landslides based on deep learning is still lacking. This study is expected to be a reference scheme for the early identification of active landslides based on deep learning.

Author Contributions: Conceptualization, X.Y.; methodology, X.Y. and X.C.; software, X.C.; validation, X.Y. and X.C; formal analysis, X.Y. and X.C.; investigation, K.R., C.Y., Y.L. and X.C.; resources, X.Y. and Z.Z.; original draft preparation, X.C. and X.Y.; writing, X.C. and X.Y.; writing—review and editing, X.Y., X.C., Z.Z., K.R., Y.L. and C.Y.; supervision, X.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the China Three Gorges Corporation (YMJ(XLD)(19)110), China Geology Survey Project (DD20221738-2), National Key R&D Program of China (2018YFC1505002), National Science Foundation of China (41672359).

Data Availability Statement: The SAR data are available at https://scihub.copernicus.eu/dhus/#/ home (accessed on 1 May 2020).

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

References

- 1. 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]
- 2. Zhao, C.; Lu, Z. Remote sensing of landslides—A review. Remote Sens. 2018, 10, 279. [CrossRef]
- 3. Zhu, Y.; Yao, X.; Yao, L.; Yao, C. Detection and characterization of active landslides with multisource SAR data and remote sensing in western Guizhou, China. *Nat. Hazards* **2022**, *111*, 973–994. [CrossRef]
- 4. Amatya, P.; Kirschbaum, D.; Stanley, T.; Tanyas, H. Landslide mapping using object-based image analysis and open source tools. *Eng. Geol.* **2021**, *282*, 106000. [CrossRef]
- 5. Guzzetti, F.; Mondini, A.C.; Cardinali, M.; Fiorucci, F.; Santangelo, M.; Chang, K.-T. Landslide inventory maps: New tools for an old problem. *Earth-Sci. Rev.* 2021, 112, 42–66. [CrossRef]
- 6. Keyport, R.N.; Oommen, T.; Martha, T.R.; Sajinkumar, K.; Gierke, J.S. A comparative analysis of pixel-and object-based detection of landslides from very high-resolution images. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *64*, 1–11. [CrossRef]
- Lu, P.; Stumpf, A.; Kerle, N.; Casagli, N. Object-oriented change detection for landslide rapid mapping. *IEEE Geosci. Remote Sens.* Lett. 2011, 8, 701–705. [CrossRef]
- 8. 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]
- 9. 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]
- 10. Chen, T.; Trinder, J.C.; Niu, R. Object-Oriented Landslide Mapping Using ZY-3 Satellite Imagery, Random Forest and Mathematical Morphology, for the Three-Gorges Reservoir, China. *Remote Sens.* **2017**, *9*, 333. [CrossRef]
- 11. Wang, H.; Zhang, L.; Yin, K.; Luo, H.; Li, J. Landslide identification using machine learning. *Geosci. Front.* **2021**, *12*, 351–364. [CrossRef]
- Yao, X.; Tham, L.; Dai, F. Landslide susceptibility mapping based on support vector machine: A case study on natural slopes of Hong Kong, China. *Geomorphology* 2008, 101, 572–582. [CrossRef]
- Wang, Z.; Brenning, A. Active-learning approaches for landslide mapping using support vector machines. *Remote Sens.* 2021, 13, 2588. [CrossRef]
- 14. Mohan, A.; Singh, A.K.; Kumar, B.; Dwivedi, R. Review on remote sensing methods for landslide detection using machine and deep learning. *Trans. Emerg. Telecommun. Technol.* **2020**, *32*, e3998. [CrossRef]
- 15. Sun, Z.; Sandoval, L.; Crystal-Ornelas, R.; Mousavi, S.M.; Wang, J.; Lin, C.; Ma, X. A review of Earth Artificial Intelligence. *Comput. Geosci.* 2022, *159*, 105034. [CrossRef]
- 16. Ma, Z.; Mei, G.; Piccialli, F. Machine learning for landslides prevention: A survey. *Neural Comput. Appl.* **2021**, 33, 10881–10907. [CrossRef]
- 17. Cai, H.; Chen, T.; Niu, R.; Plaza, A.J. Landslide Detection Using Densely Connected Convolutional Networks and Environmental Conditions. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 5235–5247. [CrossRef]
- Ghorbanzadeh, O.; Shahabi, H.; Crivellari, A.; Homayouni, S.; Blaschke, T.; Ghamisi, P. Landslide detection using deep learning and object-based image analysis. *Landslides* 2022, 19, 929–939. [CrossRef]
- 19. Li, H.; He, Y.; Xu, Q.; Deng, J.; Li, W.; Wei, Y. Detection and segmentation of loess landslides via satellite images: A two-phase framework. *Landslides* 2022, *19*, 673–686. [CrossRef]
- 20. 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]
- 21. Yu, B.; Chen, F.; Xu, C.; Wang, L.; Wang, N. Matrix SegNet: A Practical Deep Learning Framework for Landslide Mapping from Images of Different Areas with Different Spatial Resolutions. *Remote Sens.* **2021**, *13*, 3158. [CrossRef]
- 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.
- Hu, J.; Shen, L.; Sun, G. Squeeze-and-excitation networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018.
- 24. Park, J.; Woo, S.; Lee, J.-Y.; Kweon, I.S. Bam: Bottleneck attention module. arXiv 2018, arXiv:1807.06514.
- 25. Woo, S.; Park, J.; Lee, J.-Y.; Kweon, I.S. CBAM: Convolutional Block Attention Module. arXiv 2018, arXiv:1807.06521.
- 26. Ji, S.; Yu, D.; 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]
- 27. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015.
- 28. Bragagnolo, L.; Rezende, L.; da Silva, R.; Grzybowski, J. Convolutional neural networks applied to semantic segmentation of landslide scars. *CATENA* **2021**, 201, 105189. [CrossRef]
- 29. 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]
- Qi, W.; Wei, M.; Yang, W.; Xu, C.; Ma, C. Automatic Mapping of Landslides by the ResU-Net. *Remote Sens.* 2020, 12, 2487. [CrossRef]

- 31. Zhang, P.; Xu, C.; Ma, S.; Shao, X.; Tian, Y.; Wen, B. Automatic Extraction of Seismic Landslides in Large Areas with Complex Environments Based on Deep Learning: An Example of the 2018 Iburi Earthquake, Japan. *Remote Sens.* **2020**, *12*, 3992. [CrossRef]
- 32. Ju, Y.; Xu, Q.; Jin, S.; Li, W.; Guo, Q. Automatic Object Detection of Loess Landslide Based on Deep Learning. *Geomat. Inf. Sci. Wuhan Univ.* **2020**, *45*, 1747–1755. [CrossRef]
- Kamiyama, J.; Noro, T.; Sakagami, M.; Suzuki, Y.; Yoshikawa, K.; Hikosaka, S.; Hirata, I. Detection of Landslide Candidate Interference Fringes in DInSAR Imagery Using Deep Learning. *Recall* 2018, 90, 94–95.
- 34. Zheng, X.; He, G.; Wang, S.; Wang, Y.; Wang, G.; Yang, Z.; Wang, N. Comparison of Machine Learning Methods for Potential Active Landslide Hazards Identification with Multi-Source Data. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 253. [CrossRef]
- 35. Zhu, X.; Montazeri, S.; Ali, M.; Hua, Y.; Wang, Y.; Mou, L.; Bamler, R. Deep learning meets SAR: Concepts, models, pitfalls, and perspectives. *IEEE Geosci. Remote Sens. Mag.* (*GRSM*) **2021**, *9*, 143–172. [CrossRef]
- Zhao, M.; Zhong, S.; Fu, X.; Tang, B.; Pecht, M. Deep residual shrinkage networks for fault diagnosis. *IEEE Trans. Ind. Inform.* 2019, 16, 4681–4690. [CrossRef]
- Badrinarayanan, V.; Kendall, A.; Cipolla, R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* 2017, 39, 2481–2495. [CrossRef] [PubMed]
- Dai, F.; Deng, J. Development characteristics of landslide hazards in three-rivers basin of southeast Tibetan Plateau. *Adv. Eng. Sci.* 2020, 52, 3–15.
- Yao, X.; Deng, J.; Liu, X.; Zhou, Z.; Yao, J.; Dai, F.; Li, L. Primary recognition of active landslides and development rule analysis for Pan Three-river-parallel Territory of Tibet Plateau. *Adv. Eng. Sci.* 2020, 52, 16–37.
- Liu, X.; Zhao, C.; Zhang, Q.; Lu, Z.; Li, Z.; Yang, C.; Liu, C. Integration of Sentinel-1 and ALOS/PALSAR-2 SAR datasets for mapping active landslides along the Jinsha River corridor, China. *Eng. Geol.* 2021, 284, 106033. [CrossRef]
- 41. Yao, X.; Chen, Y.; Liu, D.; Zhou, Z.; Liesenberg, V.; Junior, J.M.; Li, J. Average-DInSAR method for unstable escarpments detection induced by underground coal mining. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *103*, 102489. [CrossRef]
- Gabriel, A.K.; Goldstein, R.M.; Zebker, H.A. Mapping small elevation changes over large areas: Differential radar interferometry. J. Geophys. Res. Solid Earth 1989, 94, 9183–9191. [CrossRef]
- 43. Zhu, J.; Li, Z.; Hu, J. Research Progress and Methods of InSAR for Deformation Monitoring. *Acta Geod. Cartogr. Sin.* **2017**, *46*, 1717–1733.
- Ma, L.; Liu, Y.; Zhang, X.; Ye, Y.; Yin, G.; Johnson, B.A. Deep learning in remote sensing applications: A meta-analysis and review. ISPRS J. Photogramm. Remote Sens. 2019, 152, 166–177. [CrossRef]
- 45. Camps-Valls, G.; Tuia, D.; Zhu, X.X.; Reichstein, M. Deep Learning for the Earth Sciences: A Comprehensive Approach to Remote Sensing, Climate Science and Geosciences; John Wiley & Sons: Hoboken, NJ, USA, 2021.
- Pedrayes, O.D.; Lema, D.G.; García, D.F.; Usamentiaga, R.; Alonso, Á. Evaluation of semantic segmentation methods for land use with spectral imaging using sentinel-2 and pnoa imagery. *Remote Sens.* 2021, 13, 2292. [CrossRef]
- Song, K.; Cui, F.; Jiang, J. An Efficient Lightweight Neural Network for Remote Sensing Image Change Detection. *Remote Sens.* 2021, 13, 5152. [CrossRef]
- Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015.
- Garcia-Garcia, A.; Orts-Escolano, S.; Oprea, S.; Villena-Martinez, V.; Garcia-Rodriguez, J. A review on deep learning techniques applied to semantic segmentation. *arXiv* 2017, arXiv:1704.06857.
- 50. Minaee, S.; Boykov, Y.Y.; Porikli, F.; Plaza, A.J.; Kehtarnavaz, N.; Terzopoulos, D. Image Segmentation Using Deep Learning: A Survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, 1. [CrossRef]
- 51. Yuan, Y.; Chen, X.; Wang, J. Object-contextual representations for semantic segmentation. In Proceedings of the European Conference on Computer Vision, Glasgow, UK, 23–28 August 2020.
- Chaurasia, A.; Culurciello, E. Linknet: Exploiting encoder representations for efficient semantic segmentation. In Proceedings of the 2017 IEEE Visual Communications and Image Processing (VCIP), Saint Petersburg, FL, USA, 10–13 December 2017.
- He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016.
- 54. Ma, Z.; Mei, G. Deep learning for geological hazards analysis: Data, models, applications, and opportunities. *Earth-Sci. Rev.* **2021**, 223, 103858. [CrossRef]
- 55. Moreira, A.; Prats-Iraola, P.; Younis, M.; Krieger, G.; Hajnsek, I.; Papathanassiou, K.P. A tutorial on synthetic aperture radar. *IEEE Geosci. Remote Sens. Mag.* 2013, 1, 6–43. [CrossRef]
- Li, Y.; Chen, H. Image recognition based on deep residual shrinkage Network. In Proceedings of the 2021 International Conference on Artificial Intelligence and Electromechanical Automation (AIEA), Nanjing, China, 14–16 May 2021.
- 57. Lin, N.; Chen, G.; Zhou, Q.; Liu, C. Dilated Residual Shrinkage Network for SAR Image Despeckling. In Proceedings of the 2021 IEEE 6th International Conference on Signal and Image Processing (ICSIP), Suzhou, China, 20–22 July 2022.
- Shi, B.; Zhang, Q.; Wang, D.; Li, Y. Synthetic Aperture Radar SAR Image Target Recognition Algorithm Based on Attention Mechanism. *IEEE Access* 2021, *9*, 140512–140524. [CrossRef]
- 59. Wu, P.; Cui, Z.; Gan, Z.; Liu, F. Two-Stage Attention Network for hyperspectral image classification. *Int. J. Remote Sens.* 2021, 42, 9249–9284. [CrossRef]
- 60. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. arXiv 2014, arXiv:1409.1556.

- 61. Dai, F.; Lee, C.; Li, J.; Xu, Z. Assessment of landslide susceptibility on the natural terrain of Lantau Island, Hong Kong. *Environ. Geol.* **2001**, *40*, 381–391.
- 62. Varnes, D.J. Landslide Hazard Zonation: A Review of Principles and Practice; International Association for Engineering Geology: Paris, France, 1984; p. 63.
- Wang, Y.; Fang, Z.; Hong, H. Comparison of convolutional neural networks for landslide susceptibility mapping in Yanshan County, China. *Sci. Total Environ.* 2019, 666, 975–993. [CrossRef] [PubMed]
- 64. Xu, Q.; Lu, H.; Li, W.; Dong, X.; Guo, C. Types of Potential Landslide and Corresponding Identification Technologies. *Geomat. Inf. Sci. Wuhan Univ.* **2022**, *47*, 377–387.
- 65. Zou, Y.; Sun, P.; Zhang, Q.; Ma, Z.; Lü, Y.; Bian, Y.; Liu, R. Analysis on spatial-temporal variation of snow cover and its influencing factors in the Hengduan Mountains from 2001 to 2019. *J. Glaciol. Geocryol.* **2021**, *43*, 1641–1658. [CrossRef]
- 66. Che, T.; Hao, X.; Dai, L.; Li, H.; Huang, X.; Xiao, L. Snow cover variation and its impacts over the Qinghai-Tibet Plateau. *Bull. Chin. Acad. Sci.* **2019**, *34*, 1247–1253.
- 67. Zhang, X.; Wang, X.; Liu, S.; Guo, W.; Wei, J. Altitude structure characteristics of the glaciers in China based on the Second Chinese Glacier Inventory. *Acta Geogr. Sin.* **2017**, *72*, 397–406.
- Eineder, M. Efficient simulation of SAR interferograms of large areas and of rugged terrain. *IEEE Trans. Geosci. Remote Sens.* 2003, 41, 1415–1427. [CrossRef]