



Article

A New Deep Learning Neural Network Model for the Identification of InSAR Anomalous Deformation Areas

Tian Zhang ^{1,2} , Wanchang Zhang ^{1,*} , Dan Cao ^{1,2}, Yaning Yi ³ and Xuan Wu ^{1,2}

¹ Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China; zhangtian@radi.ac.cn (T.Z.); caodan@radi.ac.cn (D.C.); wuxuan19@mails.ucas.ac.cn (X.W.)

² College of Earth and Planetary Sciences, University of Chinese Academy of Sciences, Beijing 100049, China

³ National Institute of Natural Hazards, Ministry of Emergency Management of China, Beijing 100085, China; yanningyi@ninhm.ac.cn

* Correspondence: zhangwc@radi.ac.cn; Tel.: +86-10-8217-8131

Abstract: The identification and early warning of potential landslides can effectively reduce the number of casualties and the amount of property loss. At present, interferometric synthetic aperture radar (InSAR) is considered one of the mainstream methods for the large-scale identification and detection of potential landslides, and it can obtain long-term time-series surface deformation data. However, the method of identifying anomalous deformation areas using InSAR data is still mainly manual delineation, which is time-consuming, labor-consuming, and has no generally accepted criterion. In this study, a two-stage detection deep learning network (InSARNet) is proposed and used to detect anomalous deformation areas in Maoxian County, Sichuan Province. Compared with the most commonly used detection models, it is demonstrated that the InSARNet has a better performance in the detection of anomalous deformation in mountainous areas, and all of the quantitative evaluation indexes are higher for InSARNet than for the other models. After the anomalous deformation areas are identified using the proposed model, the possible relationship between the anomalous deformation areas and potential landslides is investigated. Finally, the fact that the automatic and rapid identification of potential landslides is the inevitable trend of future development is discussed.

Keywords: deep learning; InSAR; landslides; object detection; surface deformation



Citation: Zhang, T.; Zhang, W.; Cao, D.; Yi, Y.; Wu, X. A New Deep Learning Neural Network Model for the Identification of InSAR Anomalous Deformation Areas. *Remote Sens.* **2022**, *14*, 2690. <https://doi.org/10.3390/rs14112690>

Academic Editors: Massimo Fabris, Mario Floris and José Fernández

Received: 3 April 2022

Accepted: 2 June 2022

Published: 3 June 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 are a natural phenomenon during which soil and rock slide down the slope as a whole or as separate masses under the action of gravity [1]. Landslides are driven by river erosion [2], groundwater activity [3], rainfall [4,5], earthquakes [6–8], and human activities [9,10], and they cause large numbers of injuries and deaths each year all over the world. According to the Global Fatal Landslide Database (GFLD), although the number of landslides is correlated with periodic extreme weather events, the annual global losses of life and property are still large compared to other disasters [11]. According to the Statistics of the China Geological Disasters Bulletin, the average direct economic losses caused by disasters in China in the past 10 years were 4 billion yuan per year. However, fortunately, 948 geological disasters were successfully predicted in 2019, direct economic losses of 830 million yuan were avoided, and more than 24,000 people's lives were protected [12]. Therefore, effective methods of identifying potential landslides and providing early warning are still urgently needed.

In 2019, Xu et al. [13] proposed an integrated space-air-ground multi-source monitoring system for early detection, i.e., the three-step investigation system. It includes a general investigation stage, which involves scanning potential geological hazards within a wide spatial range; a detailed investigation stage, which includes determining the geological hazard risk sections within a local range; and a verification stage, which includes in-situ

confirmation. The traditional surface deformation detection method is mainly on-site detection. GPS has the advantages of simple operation, small error accumulation and the ability to obtain 3D absolute deformation information of the surface in real time. Therefore, it is widely used in deformation monitoring such as slope, dam and surface settlement [14]. Borehole inclinometers are another typical method for measuring landslide phenomena [15]. It can timely obtain the position, development speed and development direction of landslide displacement surface by measuring the variation value of the inclination angle at different depths of the inclinometer pipe through the sensor [16]. In addition, time domain reflectometry (TDR) optical fiber sensing technology [17,18] and RGB-D sensors [19,20] are also commonly used deformation monitoring methods. However, the above methods need to deploy many instruments on site for collection, which is more suitable for detection in a certain area. At the initial stage of landslide identification, in the face of a large range and complex terrain, these methods will face the problems of high cost and low efficiency, which makes it difficult to realize effective monitoring. Interferometric synthetic aperture radar (InSAR) is one of the most appropriate methods for use in the general investigation stage due to its advantages of wide coverage, high spatial resolution, and low comprehensive cost [21]. Differential InSAR (D-InSAR) was originally used for landslide monitoring, but in applications, especially in mountainous areas with large topographic relief, the application effect of spaceborne InSAR is often restricted by geometric distortion, spatiotemporal decoherence, and atmospheric disturbance, resulting in unsatisfactory results [22]. Subsequently, time series InSAR techniques, such as persistent scatterer InSAR (PS-InSAR) [23], corner reflector InSAR (CR-InSAR) [24], and small baseline subset InSAR (SBAS-InSAR) [25], weakened the influences of the interfering factors, more accurately restored the real deformation of the surface, and identified potential landslides. However, at present, the main method of identifying anomalous deformation areas based on InSAR data is manually delineated [26,27], which is time-consuming, labor-consuming, and has no commonly accepted criterion [28]. Thus, an automatic or semi-automatic method of identifying anomalous deformation areas that can improve the identification efficiency and avoid the omissions caused by manual identification is needed.

The automatic and semi-automatic extraction technology of landslides originated in the early 21st century. In the early stage, the main method was landslide sustainability mapping. According to whether the model takes into account the internal physical and mechanical mechanism of landslide, the landslide sustainability mapping model can be divided into a deterministic model and a non-deterministic model [29]. The deterministic model is based on the mechanical mechanism and physical process of slope failure, and uses the stability state of slope as the evaluation index. Deterministic models have high accuracy, including the Sinmap model [30], the TRIGRS model [31], etc. Its advantage is that it can quantitatively calculate the slope stability, but the deterministic model should have detailed mechanical and physical parameters as the model input, and the model is very sensitive to these parameters. It is more suitable for the landslide research model in small-area homogeneous areas, while for large-area areas, the model parameters are not easy to obtain, the calculation is complex and the cost is high. Through the statistical analysis of historical disaster information, the non-deterministic model establishes the mathematical relationship between geological disasters and influencing factors, and applies this relationship model to similar geological environment areas. Common non-deterministic methods include the logistic regression model [32], SVM [33], neural network [34], and other machine learning methods [35]. Applying multi-source data to change detection is also one of the effective methods to identify potential landslides [36]. InSAR surface deformation data is one of the important factors in these models [37].

With the rapid development of artificial intelligence technology, a series of model methods represented by deep learning algorithms have attracted considerable attention in the field of remote sensing. Deep learning algorithms have higher accuracy, faster operation speeds and less computational space [38]. At present, the potential landslide identification model based on InSAR deformation data and deep learning models is still

in the exploratory stage, while the extraction of potential landslide according to surface deformation is urgently needed by local governments to control landslide. Therefore, this study aims to establish a model to realize the rapid identification of landslides in a large area through more advanced and faster deep learning algorithms.

In the last few decades, deep learning (DL) architectures have become one of the most rapidly developing technical methods in the computer vision field. The concept of deep learning was first proposed in 2006. Hinton et al. proposed stacking layer by layer unsupervised pre-training models to build a deep neural network model [39]. The layer by layer pre-training strategy solves the difficult problem of neural network parameter training and expands the application scope of neural network. Since then, deep learning has ushered in a period of rapid development. The outbreak of deep learning began in 2012. Krizhevsky et al. proposed the deep convolution neural network alexnet [40] in the Imagenet international image classification competition, and finally won the competition with an overwhelming advantage of more than 10 percentage points. After that, CNN has become the research focus of deep learning and has been widely used in the field of computer vision. Many excellent models have been proposed, such as VGGNet [41], ResNet [42], Fast R-CNN [43], DeepLab [44], etc. and they have been widely applied in the fields of image classification [45], segmentation [46], and detection [47]. The purpose of a detection model is to determine where and what the object is, which is very consistent with the requirements for identifying anomalous deformation areas. Image object detectors are usually divided into two categories. The first category includes two-stage detectors [48], in which the detection is segmented. First, candidate object bounding boxes are proposed through a regional proposal network, and then the features are extracted for each proposed object bounding box to enable classification and bounding box regression. The advantage of a two-stage detector is the higher positioning and object recognition accuracies. The second type includes one-stage detectors, which skip the step of extracting the prediction frame and directly extract it from the image instead. One-stage detectors have significantly better detection speeds, and their advantage in efficiency makes it possible to achieve real-time detection [42,49,50]. However, to avoid major casualties caused by miss identification, accuracy is a more important factor in landslide identification, and it is worth losing a little operation time to ensure accuracy. Therefore, a two-stage model is a more suitable potential landslide identification model.

In this study, a new detection model (InSARNet) was developed to detect anomalous deformation areas. The SBAS-InSAR deformation results for Maoxian County were used as the samples. The InSARNet was compared with several models, and its unique advantages in identifying anomalous deformation in mountain areas were investigated. After the anomalous deformation areas were identified using the model, the suspected potential landslide was calibrated to provide a basis for subsequent judgment.

2. Study Area and Materials

2.1. Description of the Study Area

Maoxian County, which is located on the eastern margin of the Qinghai-Tibet Plateau in the western Sichuan Basin and has an area of 3903.28 km², was selected as the study area (31°25′–32°16′N, 102°56′–104°10′E). The study area contains high/extremely high mountains and deep valleys (Figure 1). The terrain of Maoxian County is high in the northwest and low in the southeast. The altitude fluctuates greatly. The highest point is located on the Wannian snow mountain in the west, with a height of 5230 m. The average altitude in Maoxian County is 1580 m. The regional landforms mostly include the Minshan Mountain Range of the Qionglai Mountain system, and the southeastern border is the tail section of the Longmen Mountain system [51]. The study area has a subtropical monsoon climate, with an average annual precipitation of 484.1 mm over the region [52]. The river systems in Maoxian County include the Minjiang River system and the Tuojiang River system, and there are more than 170 rivers and 25 lakes in the study area [53].

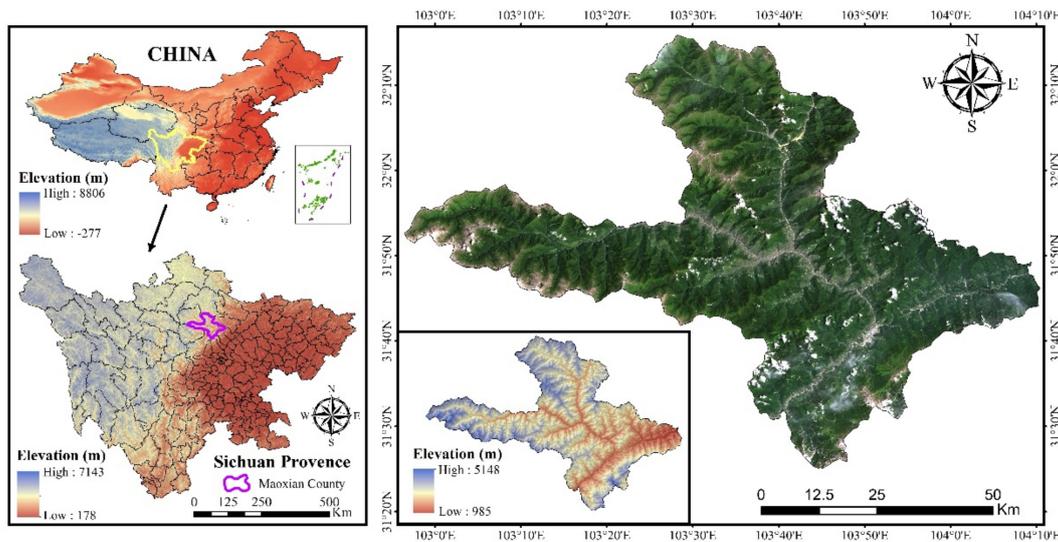


Figure 1. Geographic location of the study area.

Affected by the 1933 Diexi Earthquake and 2008 Wenchuan Earthquake, once heavy rainfall occurs, there is a high risk of geological hazards such as landslides, rock falls, and dammed lakes [54,55]. Therefore, disaster prevention based on remote sensing identification is urgently needed in Maoxian Country.

2.2. Acquisition of Surface Deformation Data

The SAR image data used in this study were acquired by the sentinel-1A satellite, which was launched by the European Space Agency (ESA) in 2014. In this study, 46 ascending single look complex (SLC) images and 37 descending SLC images of Mao County acquired from January 2015 to July 2017 were selected. The specific image date and track direction are shown in the Appendix A (Table A1). The imaging mode was the interference wide (IW) mode, and the polarization mode was vertical-vertical (VV) polarization. The corresponding precise orbit data published by the ESA and the Shuttle Radar Topography Mission (SRTM) digital elevation model (DEM) with a resolution of 30 m were used to correct for the image orbit error and terrain phase.

SBAS-InSAR was used to produce the InSAR time series deformation results. The Sarscape 5.2 from the ENVI software platform of Swiss SARMAP company was used. The first step of SBAS-InSAR processing is to generate a connection plot. A long time baseline will lead to a large number of incoherence and affect the accuracy of the results. Therefore, considering the characteristics of the data itself and the characteristics of the region, while ensuring the number of image pairs in the small baseline set, we reduced the impact of seasonal changes on the interference processed images. In this study, the maximum time baseline for ascending orbit data is 90 days. In the descending orbit data, the maximum time baseline for orbit lowering is 180 days in order to ensure the number of image pairs in the set because the data from September 2016 to February 2017 cannot be obtained. The corresponding time-position plot and time-baseline plot of the ascending and descending orbit are shown in Figures 2 and 3.

The second step of SBAS-InSAR is the interference workflow. The interference pairs are processed by differential interference to remove the error caused by the flatten effect, and the random noise is eliminated by filtering. The filtering method selected in the study is the Goldstein filtering method. Further, the filtered interferogram is phase unwrapped, so that the phases in the processed interferogram are between one cycle. In order to show the effect of interference and unwrapping, two image pairs in orbit lifting and descending are selected for display and description. Figure 4 is the interference result diagram of 20170425–20170507, and Figure 5 is the interference result diagram of 20170219–20170303, in which (a), (b) and (c) are the overall interferogram, coherence coefficient diagram and unwrapping

results of Maoxian County, and (d), (e) and (f) are the partial amplified interferogram, coherence coefficient diagram and unwrapping results. The coherence coefficient map is represented by gray scale, and dark color indicates that the coherence of the two images is weak, while light color indicates that the correlation between the two images is strong. The after filtering interferogram can determine the interference effect of the two images through the color continuity. The better the interference effect, the more continuous the color distribution on the image. Figures 4 and 5 confirm that the unwrapping effect in this study is correct. Then, 120 ground control points (GCPs) were selected to estimate and remove the residual phases.

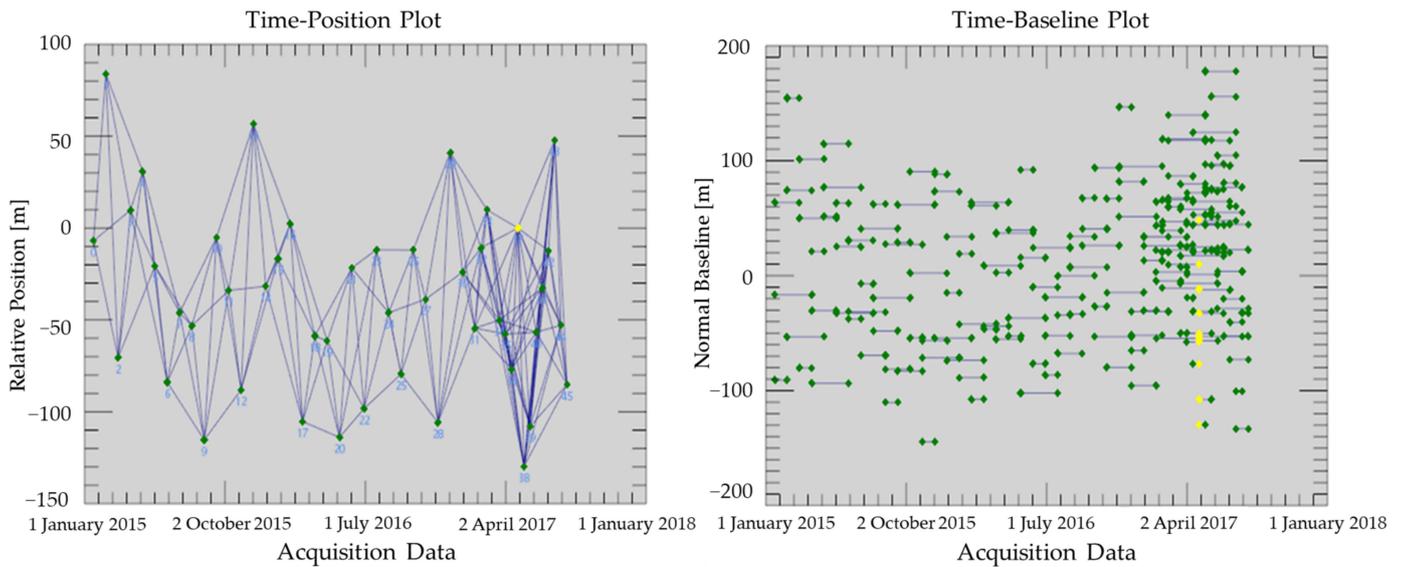


Figure 2. Time-position plot and time-baseline plot of ascending orbit data.

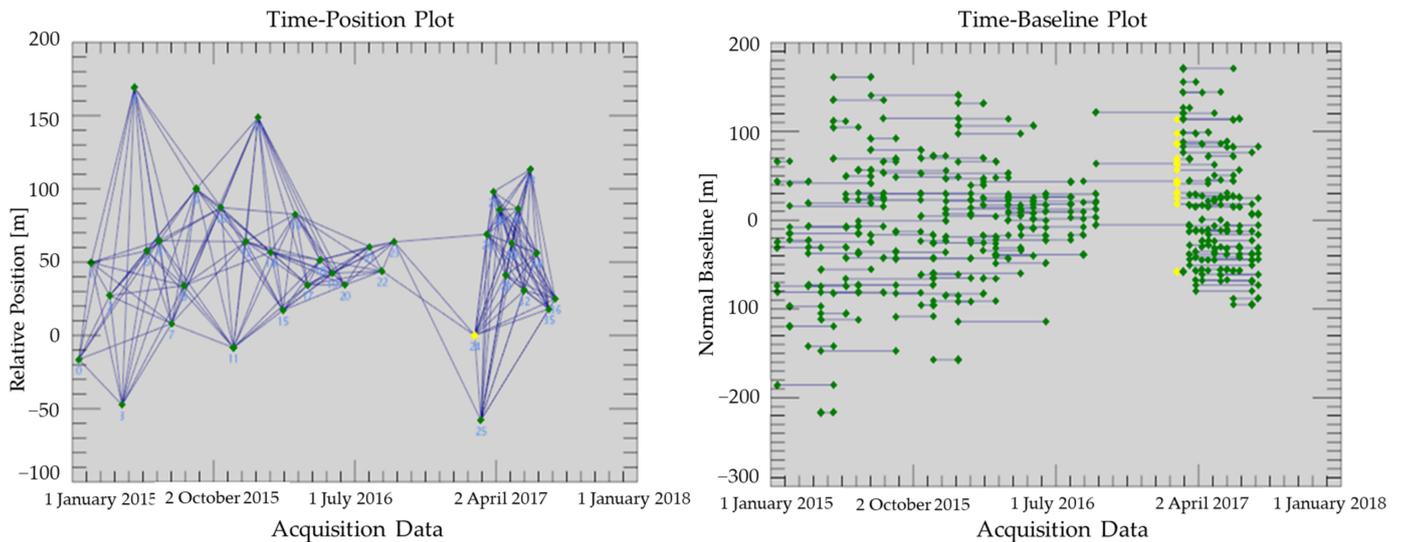


Figure 3. Time-position plot and time-baseline plot of descending orbit data.

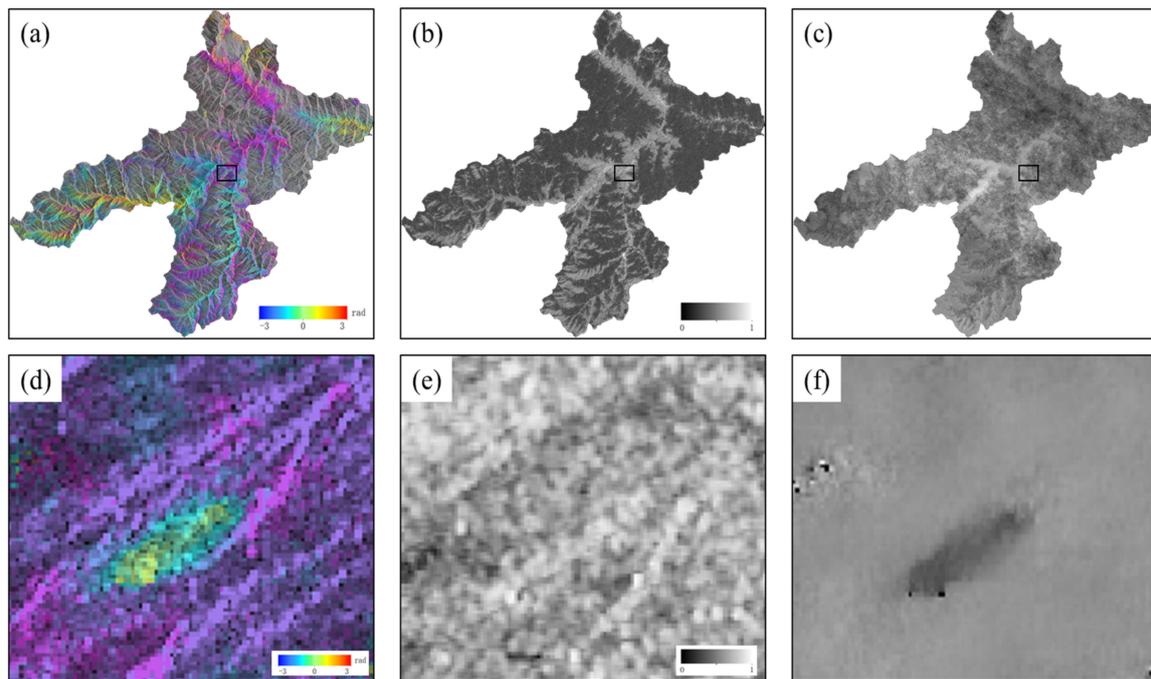


Figure 4. Interferogram and partial amplified diagram of 20170425-20170507 ascending orbit data. (a) is the overall interferogram of Maoxian county, (b) is the coherence coefficient diagram of Maoxian county, (c) is the unwrapping results of Maoxian County, (d) is the partial amplified interferogram, (e) is the partial amplified coherence coefficient diagram, and (f) is the partial amplified unwrapping results.

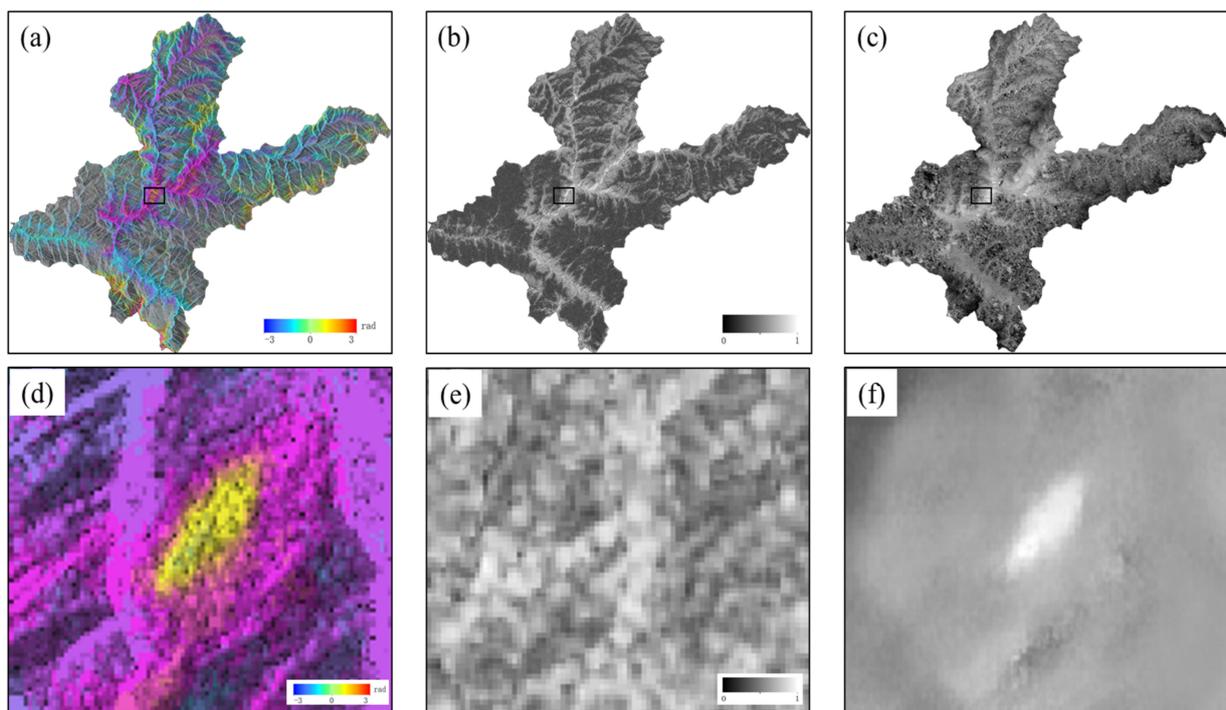


Figure 5. Interferogram and partial amplified diagram of 20170219-20170303 descending orbit data. Interferogram and partial amplified diagram of 20170425-20170507 ascending orbit data. (a) is the overall interferogram of Maoxian county, (b) is the coherence coefficient diagram of Maoxian county, (c) is the unwrapping results of Maoxian County, (d) is the partial amplified interferogram, (e) is the partial amplified coherence coefficient diagram, and (f) is the partial amplified unwrapping results.

Finally, the deformation sequence is extracted by two inversions. In the first inversion, the residual phase in the overall phase is separated by the secondary unwrapping of the minimum cost flow method, and the deformation rate is preliminarily estimated. The second inversion eliminates the influence of atmospheric phase and noise phase in the overall phase by filtering. The linear model with the highest robustness was used to estimate the deformation rate and the residual phase of each image pair. The deformation results from the ascending orbit dataset and the descending orbit dataset are shown in Figure 6.

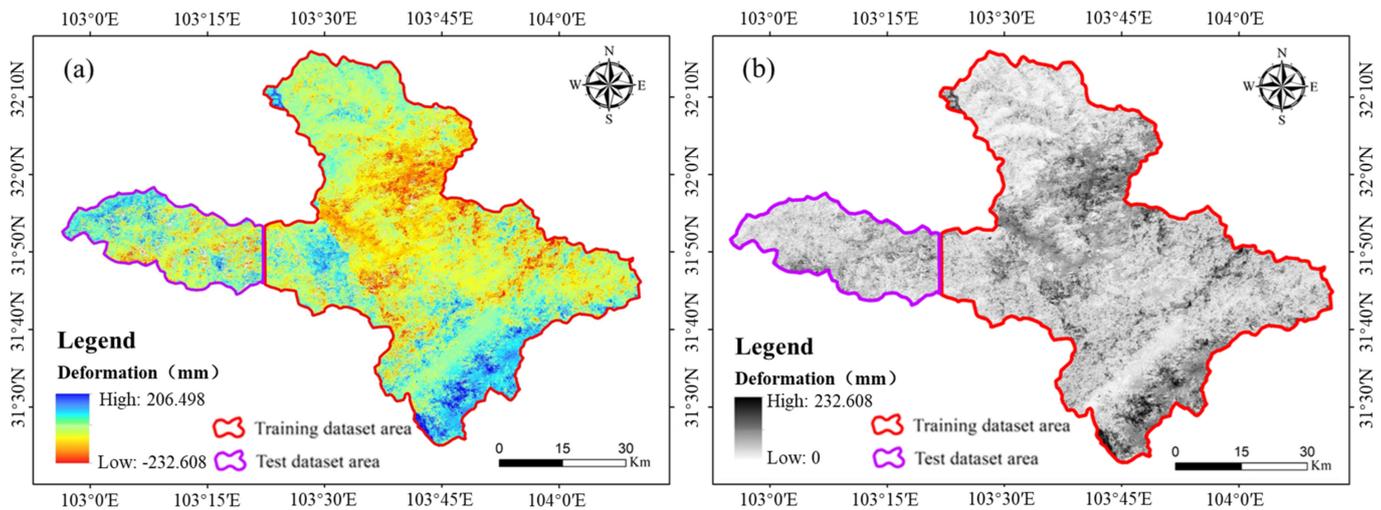


Figure 6. SBAS-InSAR deformation results obtained from the ascending orbit dataset (a) and the descending orbit dataset (b).

Through the line of sight cumulative deformation extracted by the ascending and descending orbit data, it can be seen that the overall distribution of the high deformation area is similar, but there are differences in some areas. There are three reasons for the difference: (1) Differences in satellite data acquisition. The observation direction of the sentinel sensor is the right viewing direction, so the data of ascending orbit and descending orbit are different in the observation direction. At the same time, with the differences of satellite incident angle, route angle and acquisition time, the same deformation shows different results in ascending orbit and descending orbit results. (2) The study area is a mountainous area with complex terrain, geometric distortions such as overlap and perspective shrinkage will occur in the hillside and areas with large topographic relief, resulting in inconsistent results of ascending orbit and descending orbit in such areas. (3) For the East–West slope in the study area, the projection of the same deformation in the line of sight direction of ascending and descending orbit is opposite, which also leads to the difference in the deformation results of ascending and descending orbit. In summary, it is not accurate to identify the potential landslide directly through the single line of sight deformation data obtained from ascending orbit data or descending orbit data. The deformation result of the single line of sight may have a large deviation from the real surface deformation. Therefore, it is necessary to obtain accurate and real slope surface shape variables through further analysis and processing.

The InSAR deformation results are the one-dimensional deformation in the line-of-sight direction. However, the slope direction is diverse, and the actual surface deformation direction does not coincide with the line of sight of the radar instrument, which leads to reduced availability of the results and the inability to meet the requirements for landslide deformation monitoring. Interpolation was used to register and fuse the InSAR deformation results obtained from the ascending orbit dataset and the descending orbit dataset on the temporal and spatial scales. The technology flowchart is shown in Figure 7.

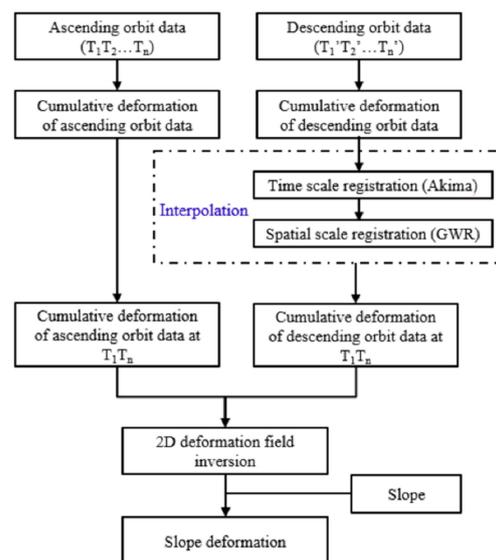


Figure 7. The technology flowchart of slope deformation extraction by ascending and descending orbit data fusion.

In order to retain more real measurement data, this study takes the ascending orbit deformation with more data as the main data and the descending orbit deformation with less data as the auxiliary data. The continuous displacement sequence is obtained by using Akima interpolation on the time scale of the descending orbit deformation, and the data with the same time as the ascending orbit deformation data is extracted from it to realize the registration on the time scale. Further, the synonym points in the ascending and descending orbit deformation image of after time registration are extracted, and the vertical deformation difference of the synonym points is obtained. Taking the longitude, latitude and aspect as factors, the difference is regressed by geographical weighted regression (GWR) method, and the corresponding point difference value of the overall study area is estimated. The estimated interpolation and the descending orbit deformation are added to realize the registration on the spatial scale of ascending and descending orbit deformation. Finally, the corrected $d_{T_1 T_2}$ and $d_{T'_1 T'_2}$ and the 2D deformation are obtained by the spatial geometric relationship (Equation (1)).

$$\begin{cases} d_{T_1 T_2} = d_V \cos \theta - d_E \cos(\varphi - \frac{3}{2}\pi) \sin \theta \\ d_{T'_1 T'_2} = d_V \cos \theta' - d_E \cos(\varphi' - \frac{3}{2}\pi) \sin \theta' \end{cases} \quad (1)$$

According to Equation (2), the deformation results were transformed into the slope direction according to the spatial geometric relationship (Figure 8).

$$d_{Slope} = d_V / \sin(S) + d_E / \cos(S) \sin(A - \pi/2), \quad (2)$$

where, S is the surface slope, A is the surface slope direction, d_V is the vertical deformation component, and d_E is the East–West deformation component.

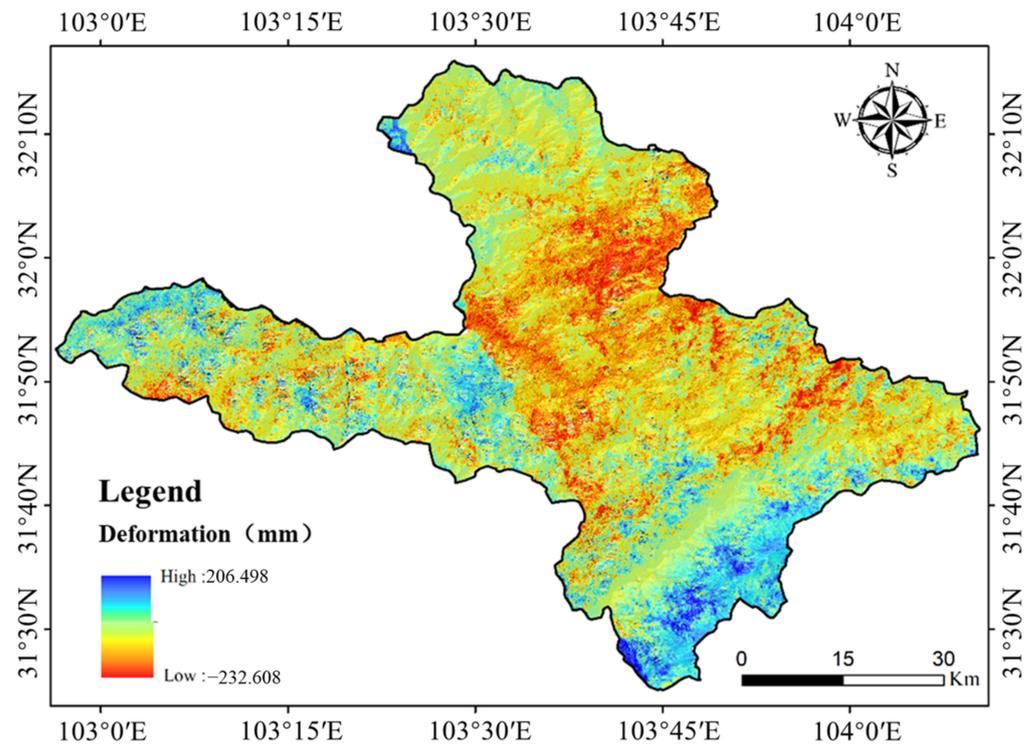


Figure 8. Deformation of slope direction in the study area.

3. Methods

3.1. Framework of Model

To achieve the high-precision identification and detection efficiency required for potential landslide detection, based on the characteristics of InSAR deformation data, a two-level detection network (InSARNet) was constructed in this study. InSARNet uses a mask region-based convolutional neural network (RCNN) as the basic framework, uses a new convolution mode involution operator to improve the operation efficiency, and uses deformable region of interest (ROI) pooling to improve the ability to monitor small-scale targets. The structure of InSARNet includes the input, backbone, neck, head, and output. The specific structure is shown in Figure 9.

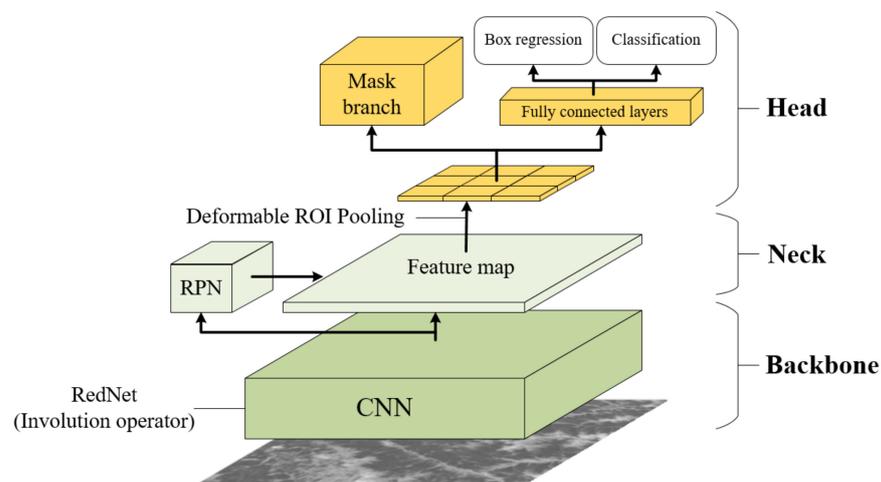


Figure 9. Architecture of InSARNet. (CNN: convolutional neural network, RPN: region proposal network).

3.2. Backbone

As the basis and core of the target detection model, the main task of the backbone network is to extract the characteristics of the input data and output the characteristics of the response for the subsequent parts. By introducing a residual module, ResNet eliminates the gradient disappearance and gradient explosion that occur in traditional neural networks through the deepening of the network, making it one of the most commonly used and efficient backbone CNN models.

In a CNN, the convolution layer is the main structure used for the feature extraction. In the current widely used convolution algorithm, to reduce the number of parameters, the same convolution kernel is used to scan one channel, and different convolution kernels are selected in different channels. As is shown in Figure 10, the convolution can be expressed as follows:

$$Y_{i,j,k} = \sum_{c=1}^{C_i} \sum_{(u,v) \in \Delta_K} \mathcal{F}_{k,c,u+[K/2], v+[K/2], [kG/2]} X_{i+u,j+v,c'} \tag{3}$$

$$\Delta_K = \left[-\left\lfloor \frac{K}{2} \right\rfloor, \dots, \left\lfloor \frac{K}{2} \right\rfloor \right] \times \left[-\left\lfloor \frac{K}{2} \right\rfloor, \dots, \left\lfloor \frac{K}{2} \right\rfloor \right], \tag{4}$$

where C_0 and C_i are the numbers of output and input channels, $X \in \mathbb{R}^{H \times W \times C_i}$ is the input data, $Y \in \mathbb{R}^{H \times W \times C_0}$ is the output data, K is the kernel size, and $\mathcal{F} \in \mathbb{R}^{C_0 \times C_i \times K \times K}$ is the convolution kernel. Its parameter sharing effectively reduces the number of parameters. However, the disadvantages of spatial-agnostic and channel-specific methods are also obvious. The number of channels is usually large. Therefore, to limit the scale of the parameters and the computation, the value of K is often small, which limits the ability of the convolution operation to capture long-distance relationships at one time.

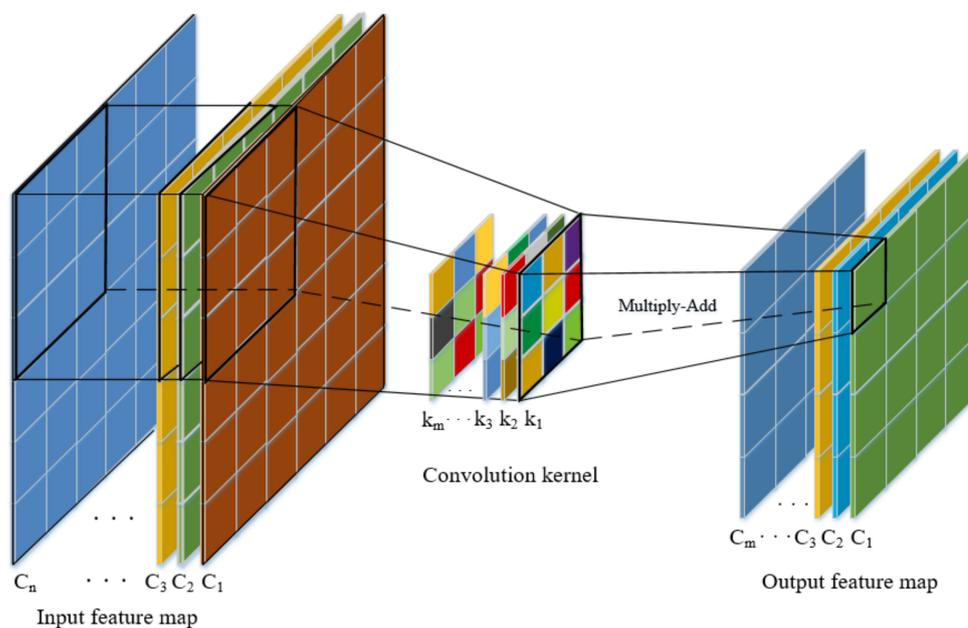


Figure 10. Schematic diagram of the conventional convolution (C is feature channel, k is convolution kernel).

To improve the detection ability and operation efficiency of the model, we selected a new operator-based network (RedNet) as the backbone network. The main difference between traditional operators and RedNet is that the 3×3 convolution in ResNet is replaced by involution. The involution operator is a new type of neural network operator, which was proposed by Li et al. in 2021 [56]. Compared with traditional convolution operators, it is more efficient because it has fewer parameters and less computation. The involution operator is designed from the opposite perspective of a convolution operator. The corresponding involution kernel is generated based on the input feature map in the

spatial dimension, and the kernel is shared in the channel dimension. The operation of the involution is as follows:

$$Y_{i,j,k} = \sum_{(u,v) \in \Delta_K} \mathcal{H}_{i,j,u+[K/2], v+[K/2], [kG/C]} X_{i+u,j+v,k}, \tag{5}$$

where $\mathcal{H} \in \mathbb{R}^{H \times W \times K \times K \times G}$ is the involution kernel. Instead of using a fixed weight matrix as the learnable parameter, the involution operator generates the corresponding operator for the input characteristic graph. As is shown in Figure 11, for a point feature vector in the input feature map, the involution kernel is generated through channel transformation expansion, and then, it is multiplied and summed with the neighborhood of the input feature map to calculate the final output feature map.

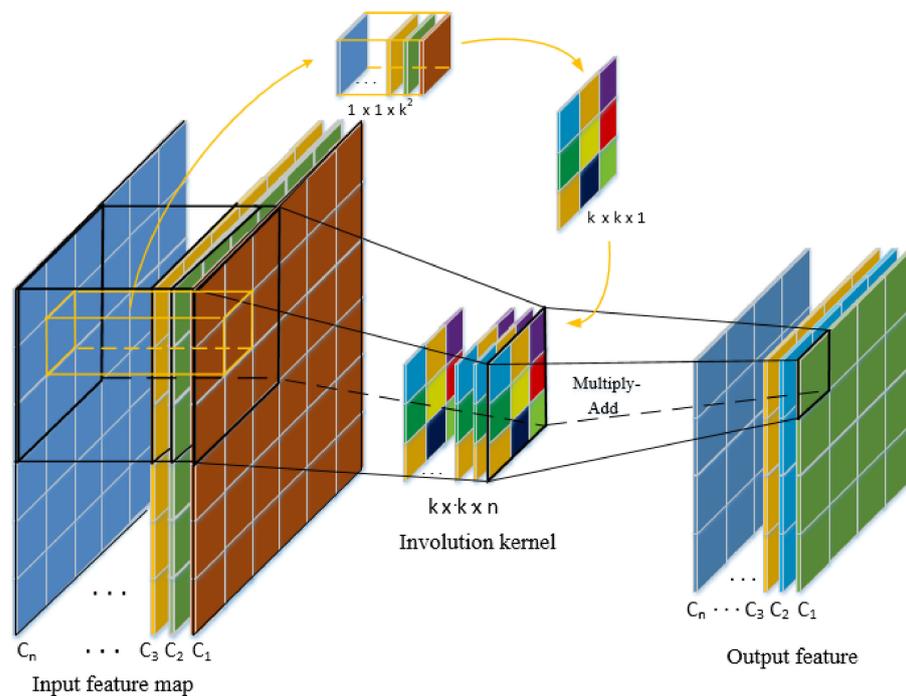


Figure 11. Schematic diagram of the involution (C is feature channel, k is involution kernel).

3.3. Neck

As the link between the backbone network and the head network, the neck network can obtain more complex features by adding the network layer of the feature map in different stages. During the traditional target detection process, after the feature information is extracted by the backbone network, it is directly input into the head network for classification and target frame prediction. This is feasible for the detection of objects with large targets. However, for the detection of objects with small targets, the feature mapping of the head network directly divides the coordinates by the step size. However, when the convolution pool is deep, the mapping is extremely small or it disappears. The InSAR anomalous deformation areas have different sizes and shapes. Therefore, it is necessary to use the feature pyramid network as the neck network to meet the needs for multi-scale detection.

The feature pyramid network (FPN) was proposed in 2017 [36]. By changing the network connection method, the FPN network can greatly improve the ability of the network to monitor small target objects without changing the number of calculations of the original model. A schematic diagram of the FPN network structure is shown in Figure 12.

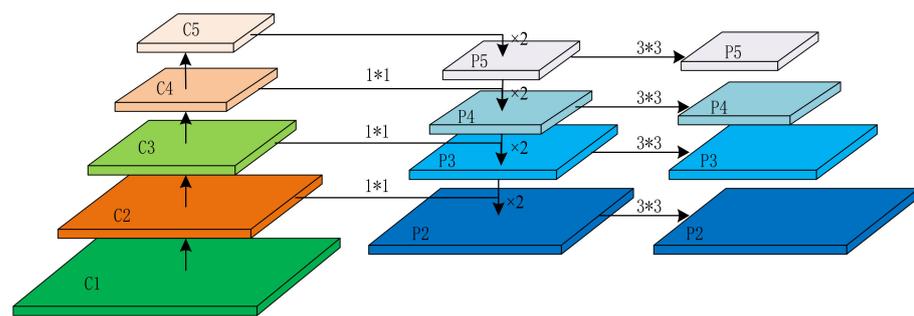


Figure 12. Schematic diagram of the feature pyramid network (FPN) network structure (C is feature channel, P is the channel through lower connection and horizontal connection).

3.4. Head

As the output framework, the purpose of the head network is to predict the type and location of the object detection. The head network can be divided into two parts according to its purpose. One part classifies the input feature map and predicts the target candidate box regression, and the other part generates the mask for the pixel-level segmentation prediction. Before the two parts of the prediction are conducted, the feature map needs to be extracted using a region proposal network (RPN) to extract the target frame, extract the features of the different feature layers, and then carry out their respective convolution operations to obtain the final results. The overall framework of the head network is shown in Figure 13.

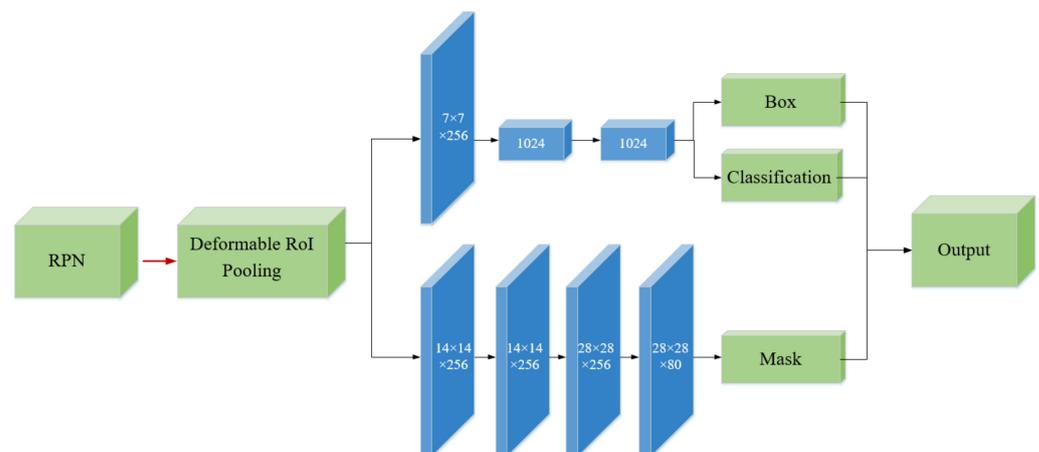


Figure 13. Schematic diagram of the FPN network structure.

In the initial two-stage detection framework, after the RPN is applied, the role of the ROI pooling is to pool the corresponding area into a fixed-size feature map in the feature map according to the coordinates of the preselected box to conduct the subsequent classification and bounding box regression operations. Since the position of the preselected box is calculated through model regression, and the coordinates are usually floating-point numbers, the ROI pooling needs to be quantized twice to fix the size. First, the candidate box boundary is quantized into integer point coordinate values. Then, the quantized boundary region is evenly divided and the boundary of each unit is quantized. However, the quantized candidate frame has a certain deviation from the initial regression position, which affects the accuracy of the detection or segmentation. Due to this disadvantage of ROI pooling, the Mask-RCNN network was used to improve this defect by proposing an ROI alignment [37]. The ROI alignment cancels the quantization operation and uses the bilinear interpolation method to obtain the image values of the pixels, the coordinates of which are floating-point numbers, to transform the entire feature aggregation process into a continuous operation. The geometric structure detected using the ROI alignment is fixed,

and it is difficult to conduct geometric transformation of the object's size, attitude, and angle. To dynamically adjust the receptive field, the deformable ROI pooling from the deformable convolutional network [38] was introduced into InSARNet. The operation of the deformable ROI pooling is

$$y(i, j) = \sum_{p \in \text{bin}(i,j)} x(p_0 + p + \Delta p_{ij}) / n_{ij}, \quad (6)$$

where $\text{bin}(i,j)$ is the total number of pixels in the grid at coordinates (i,j) , $p_0 + p$ is the coordinate of the sampling point, and Δp_{ij} is the offset generated from the feature mapping.

As is shown in Figure 14, the main principle of deformable ROI pooling is to add a displacement variable in the convolution sampling layer, which is learned from the data. After the offset, it is equivalent to the scalable change in each block of the convolution kernel to expand the range of the receptive field.

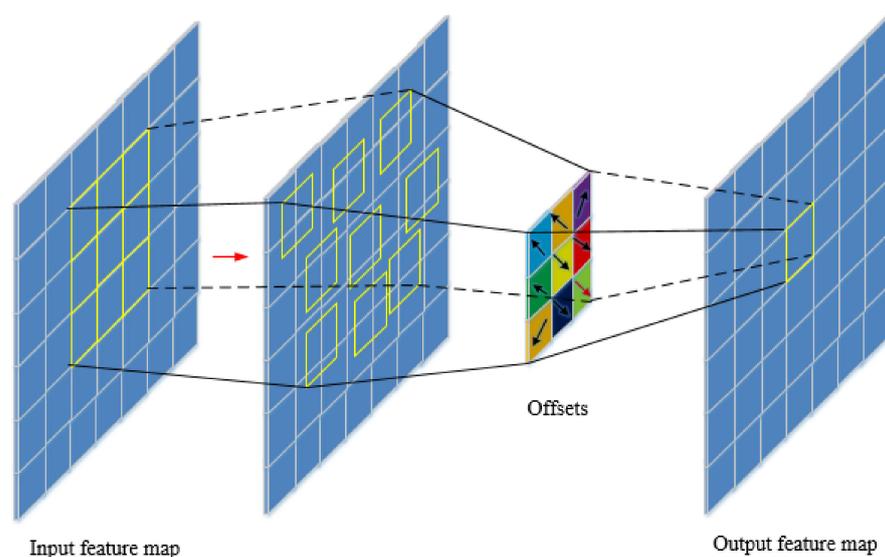


Figure 14. Schematic diagram of the deformable regions of interest (ROI) pooling.

3.5. Loss Function

The loss function is used to test the gap between the predicted value and the real data. The optimization process of the neural network is the process of minimizing the loss function. The loss function L used in this study consists of three parts: the classification loss L_{cls} , positioning frame regression loss L_{box} , and mask loss L_{mask} . The L_{cls} is determined using the L1 loss function, and L_{box} and L_{mask} are determined using the cross-entropy function.

4. Experiment and Results

4.1. Datasets

In this study, the InSAR deformation rate in Maoxian was used to construct the anomalous deformation area dataset, which was divided into east and west parts as the training area and test area, respectively (Figure 15). The training area was 3089.6 km², and the image size was 8477 × 4279 pixels. The test area on the west side was 819.7 km², and the image size was 4279 × 2302 pixels. To highlight the changes in the surface deformation, the deformation image was transformed into a gray-scale image. Since the positive and negative signs only represent the directions of the displacement relative to the satellite, and the focus of this study was the size of the shape variable, the absolute value of the image shape variable was calculated during the conversion to the gray-scale image (Figure 15b). The surface deformation resolution calculated from the Sentinel-1 data is 20 m, the deformation result is in the form of grid data, and each grid corresponds to 32-bit floating-point data.

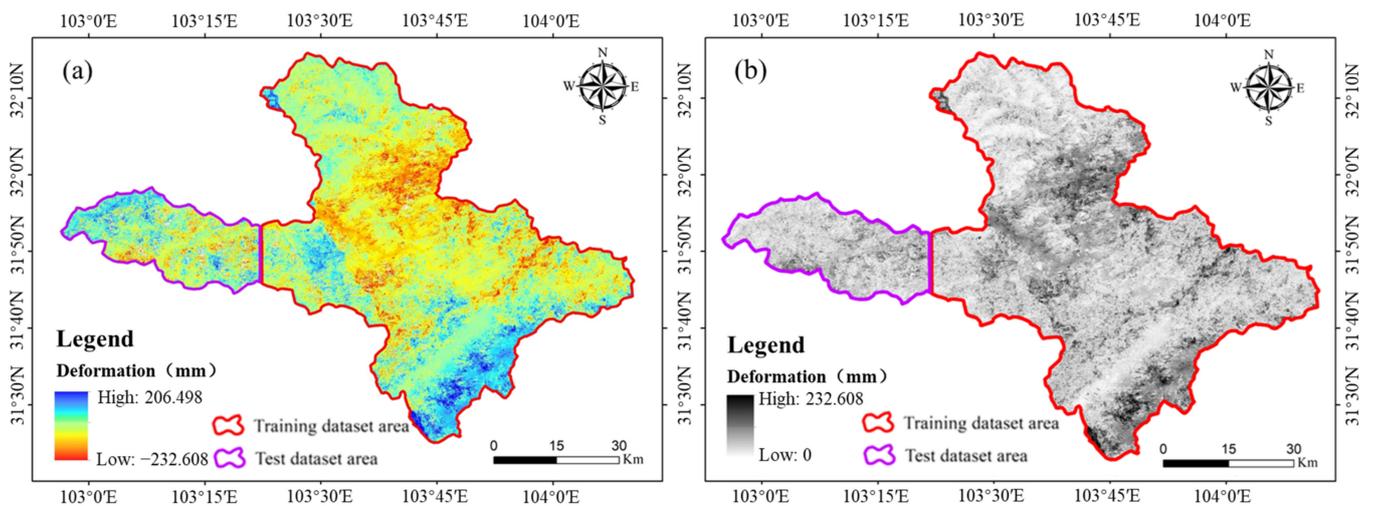


Figure 15. Division of training area and test area (a), and gray-scale conversion (b).

In this study, it is determined that the anomalous deformation area needs to meet the following three conditions: (1) The internal deformation difference in the anomalous deformation area shall not exceed 5 mm. (2) The difference between internal and external deformation at the boundary of anomalous deformation area is greater than 5 mm. (3) The anomalous deformation area is a closed polygon, and the number of internal grids is not less than 16. The distributions of the 316 training samples and 103 test samples are shown in Figure 16.

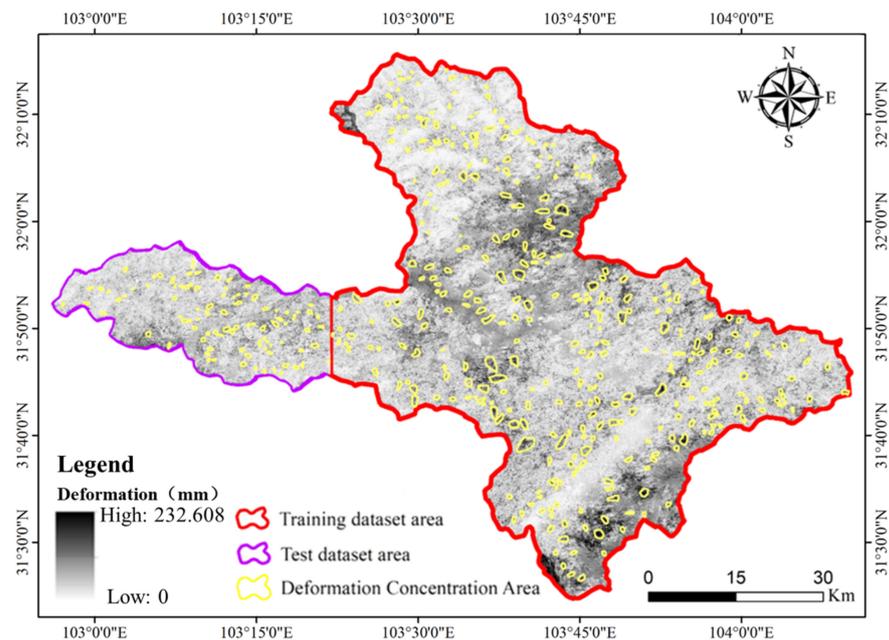


Figure 16. Delineation of anomalous deformation areas in the study area.

Through the statistical analysis, the size of the dense area was concentrated in the 56×56 – 128×128 interval, and the largest intensive deformation area was 256×116 , so the size of a single training image was set to 256×256 , which ensured that all of the anomalous deformation areas were fully reflected in a single image and reduced the computer performance requirements. Based on this, the data were enhanced via rotating and flipping (Figure 17), and 1896 training samples and 618 test samples were obtained. Then, labelme software was used to label the anomalous deformation area samples image

by image, and to convert them into a coco format dataset. Finally, 2514 dataset samples and their corresponding label files were obtained.

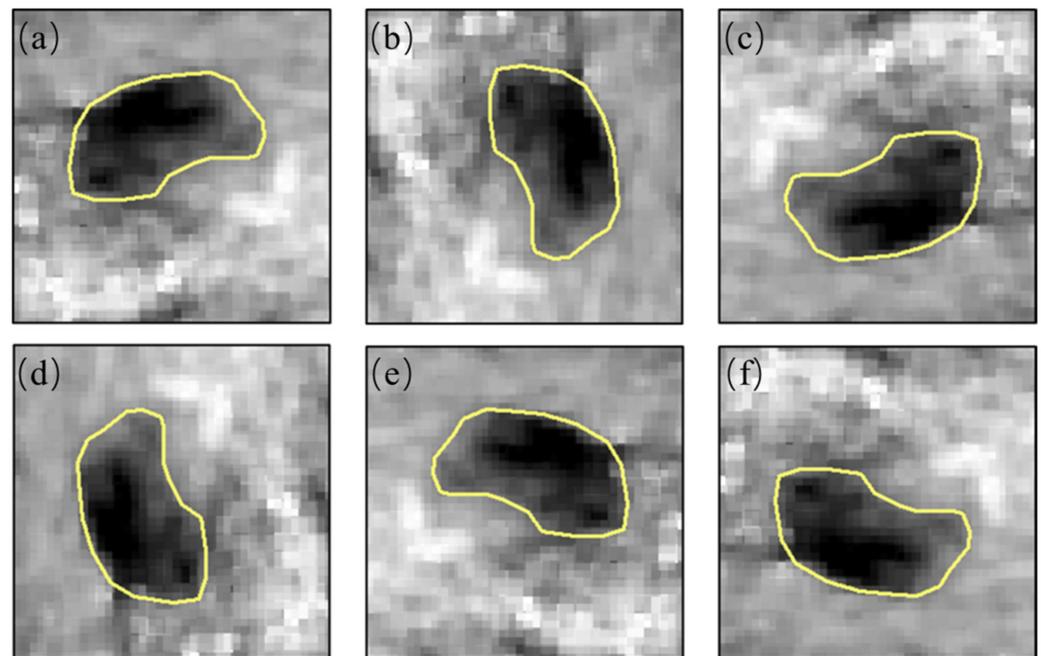


Figure 17. Schematic diagram of the data augmentation. (a) Original image, (b) Rotated 90° , (c) Rotated 180° , (d) Rotated 270° , (e) Flip horizontally, and (f) Flip vertical.

4.2. Experimental Setup

The experiments were implemented using the PyTorch 1.9 software on an NVIDIA GeForce RTX 2080Ti GPU. The model code was written in Python 3.6, the computer environment was a Linux system, and the graphics memory was 64 GB. The training parameters are listed in Table 1.

Table 1. Experimental parameter settings of the network model.

Parameter	Value
Optimizer	SGD
Validation data scale	0.25
Epoch	30
Iteration	7110
Batch size	8
Initial learning rate	0.001
Learning rate decay interval	10
Learning rate attenuation	0.1

After 30 training epochs, InSARNet converged. The training and verification curves are shown in Figure 18.

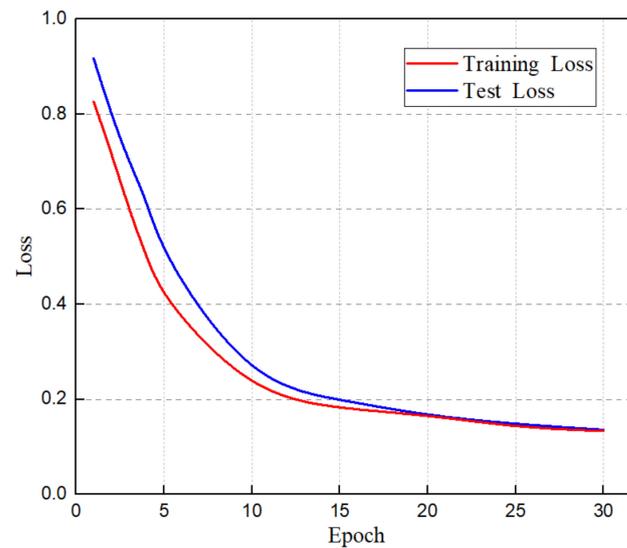


Figure 18. Loss curves for the training and testing of the InSARNet model.

4.3. Evaluation Index

In InSAR anomalous deformation area detection, the calibration samples can be divided into two categories: anomalous deformation areas (True) and non-anomalous deformation areas (False). The detection results can also be divided into two categories: the areas determined as anomalous deformation areas by the classifier (Positive) and the areas determined as non-anomalous deformation areas by the classifier (Negative). We used a combination of two initials to represent the number of samples in the different cases. For example, TP represents the number of samples with anomalous deformation areas determined through both manual calibrations and using the classifier. In this experiment, the sample data were used to verify the accuracy from the perspective of five common verification indexes: precision, recall, F1 score coefficient, overall accuracy (OA), and kappa coefficient. The indexes were calculated using the following equations:

$$\text{Precision} = TP / (TP + FP), \quad (7)$$

$$\text{Recall} = TP / (TP + FN), \quad (8)$$

$$F1 = 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall}), \quad (9)$$

$$OA = (TP + TN) / (TP + TN + FP + FN), \quad (10)$$

$$\text{Kappa} = \frac{OA - \frac{[(TP+TN) \times (TP+FP) + (FP+FN) \times (TN+FN)]}{n \times n}}{1 - \frac{[(TP+TN) \times (TP+FP) + (FP+FN) \times (TN+FN)]}{n \times n}}. \quad (11)$$

The accuracy, recall, and F1 score reflect the recognition effect of the model in the dense deformation areas, while the overall accuracy and kappa coefficient are overall evaluation indicators, which more comprehensively reflect the extraction effect of the model in the dense deformation areas.

4.4. Experimental Results

In this study, three rectangular areas (1024×1024 grids, about 11 km) in the study area were selected for analysis to verify the recognition accuracy of InSARNet. One of the verification areas was located in the training area and the other two were located in the test area (Figure 19).

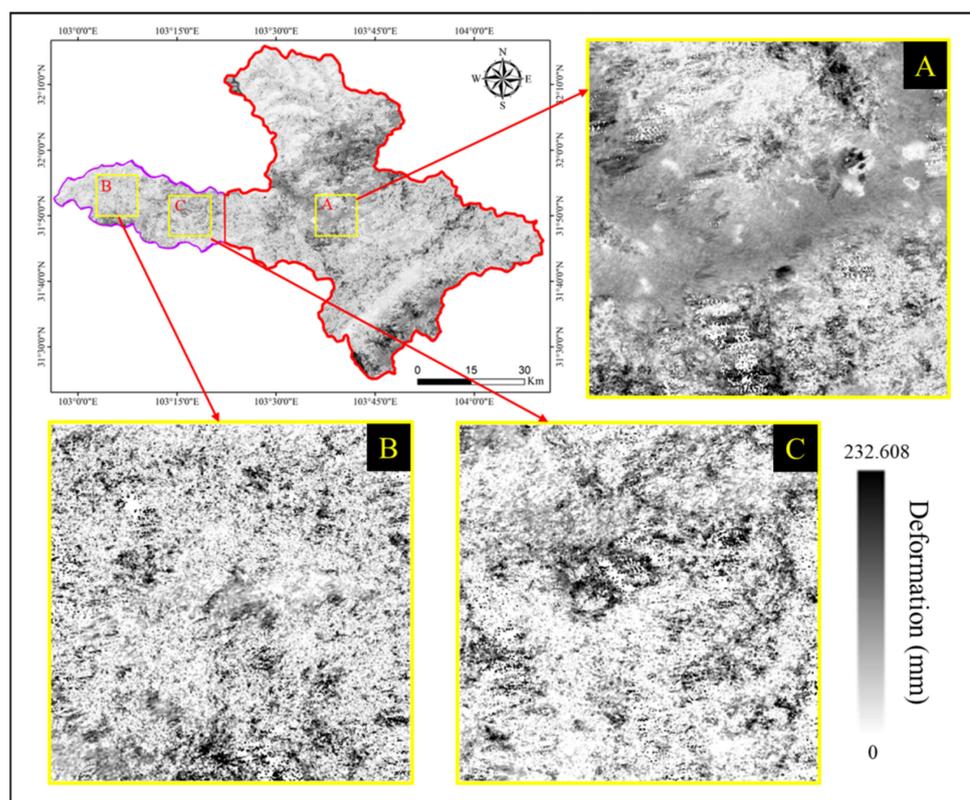


Figure 19. Locations of testing areas. (A) InSAR result of test area A, (B) InSAR result of test area B, and (C) InSAR result of test area C.

4.4.1. Comparison of Modules

To test the effect of introducing an involution operator and deformable ROI pooling into the model, we tested and compared the traditional Mask RCNN, the Mask RCNN with only an involution operator, the Mask RCNN with only deformable ROI pooling, and InSARNet with both an involution operator and deformable ROI pooling. These four models were labeled model I, model II, model III, and model IV, respectively. The corresponding network framework is shown in Table 2. The same training environment, training strategy, and parameters were used for the tests, and the training test was carried out using the same dataset.

Table 2. Detection models with different module combinations.

Model	Backbone	Neck	Head
I	ResNet-50	¹ FPN	² RPN + ROI Align
II	ResNet-50	¹ FPN	RPN + Deformable ROI pooling
III	RedNet-50	¹ FPN	² RPN + ROI Align
IV	RedNet-50	¹ FPN	² RPN + Deformable ROI pooling

¹ FPN: feature pyramid network. ² RPN: region proposal network.

Area A is located in the middle of Maoxian County and has a relatively flat terrain. The InSAR results show that the anomalous deformation areas are obvious and the noise is smaller. The recognition results of the four models in test area A are shown in Figure 20. It can be seen that the four models effectively identified most of the anomalous deformation areas, but model I and model III had higher missed detection rates. The introduction of deformable ROI pooling improved the ability of model II and model IV to monitor small targets, but model IV had fewer false detection areas than model II. Therefore,

comprehensively, for area A, model IV controlled both the missed detection and false detection rates and had a better detection effect.

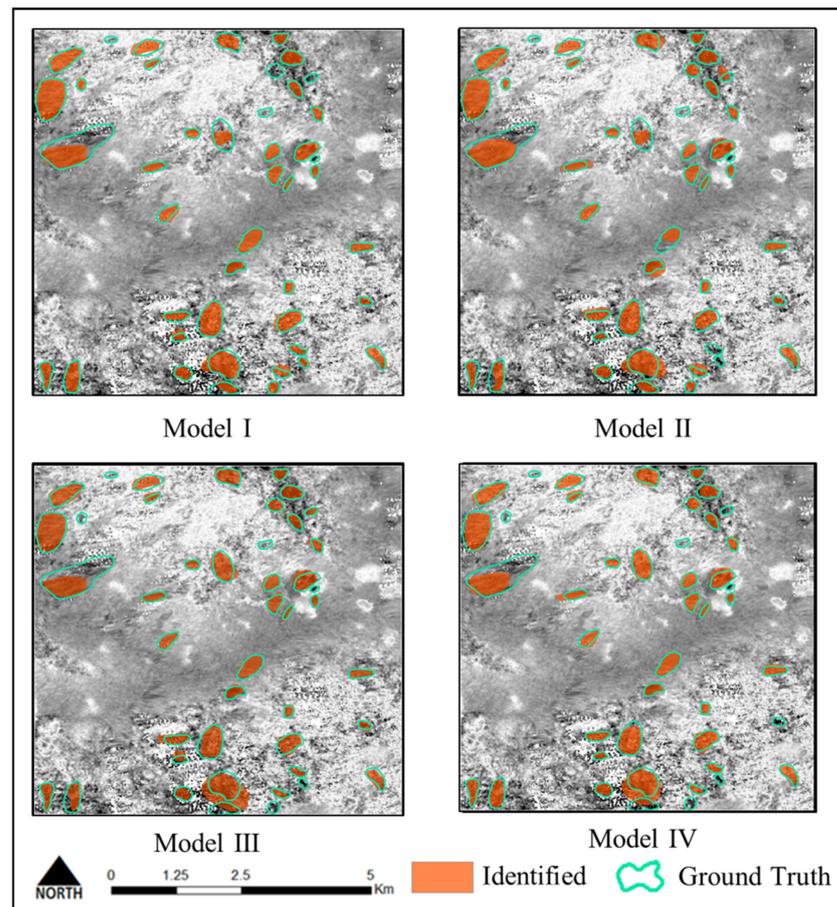


Figure 20. Comparison of the different modules for test area A.

This is also reflected in the quantitative evaluation indicators of the four models (Table 3). The four models had good detection effects in test area A. For the monitoring accuracy, InSARNet (model IV) had the highest accuracy, exceeding 0.92, which is nearly 2% higher than that of model II, which had the second-highest accuracy. The recall rates of the four models exceeded 0.9. The effect of InSARNet was better than those of the other three models, and its recall rate exceeded 0.93, that is, 1% better than those of the other models. The F1 coefficient and kappa coefficient of InSARNet reached 0.93 and 0.86, respectively, and were higher than those of the other three models. The overall accuracies of the four models exceeded 0.90, indicating that the four models effectively extracted the target and background. Among them, the effect of InSARNet was slightly better than those of the other three models.

Table 3. Comparison of the accuracies of the models with different module combinations in test area A.

Model	Precision	Recall	F1	Kappa	¹ OA
I	0.8984	0.9133	0.9058	0.8074	0.9050
II	0.9076	0.9167	0.9121	0.8219	0.9117
III	0.8935	0.9233	0.9082	0.8081	0.9067
IV	0.9274	0.9367	0.9320	0.8621	0.9317

¹ OA: Overall Accuracy.

Area B was located in the western part of Maoxian County. Compared with area A, the situation of test area B was more complex. There are more anomalous deformation areas with a wide distribution range and different sizes, and the noise is also significantly higher than that in area A. The detection results of the four models are shown in Figure 21. Although most of the anomalous deformation areas were identified by the four models, many false detections occurred in models I and III. Model II and model IV had good detail recognition abilities. The introduction of the deformable ROI pooling improved the detail recognition in small areas. In addition, model II also had many small areas of missed detection. Although InSARNet also had a certain degree of missed detection in these areas, compared with the other three models, it had fewer missed detection areas, and the boundary of the monitored area was more consistent with the real values, indicating that the introduction of the involution operator also improved the model's detection ability.

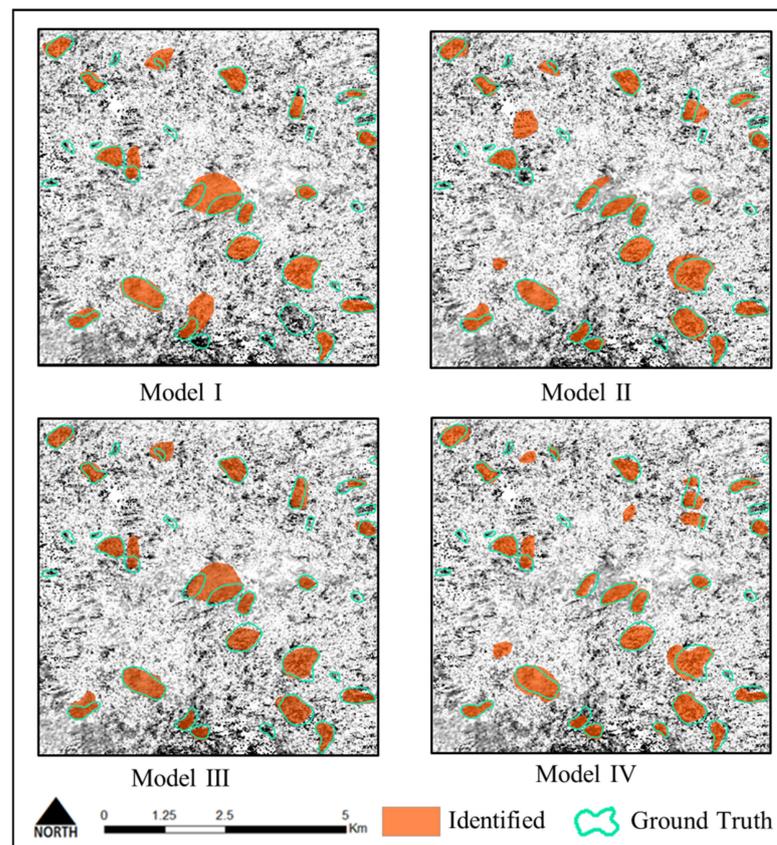


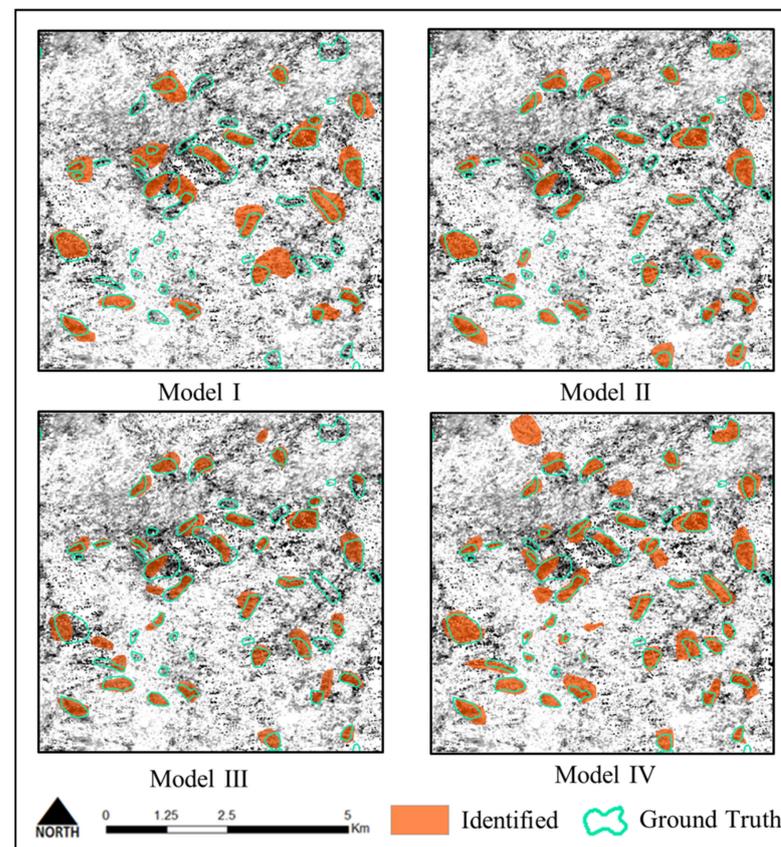
Figure 21. Comparison of the different modules in test area B.

The quantitative evaluation indexes for the four models in test area B are presented in Table 4. Compared with area A, the values of all of the indicators were lower, but the four models still successfully achieved object detection. Among the four models, InSARNet still had the highest precision (0.89), followed by model II (0.86). Compared with the other three networks, the detection accuracy of InSARNet was nearly 5% better on average. The recall rates of all of the models exceeded 0.8, and the recall rate of InSARNet reached 0.92, which was the highest of the four models. Similarly, for the F1 values and kappa coefficients, the values of the InSARNet model were also significantly higher than those of the other three models. The overall accuracies of the four models also exceeded 0.80, indicating that the four models effectively extracted the target and background. The comprehensive results show that for the more complex situation, the simultaneous introduction of an involution operator and deformable ROI pooling resulted in a better performance.

Table 4. Comparison of the accuracies of the models with different module combinations in test area B.

Model	Precision	Recall	F1	Kappa	OA
I	0.8289	0.8400	0.8344	0.6637	0.8333
II	0.8696	0.8667	0.8681	0.7373	0.8683
III	0.8466	0.8833	0.8646	0.7144	0.8617
IV	0.8990	0.9200	0.9094	0.8131	0.9083

The identification results of the four models for test area C are shown in Figure 22. Area C was located to the east of area B, and area C contained scattered anomalous deformation areas and had more noise. Model I had more false detection areas. Although model II and model III produced better results, they still had many missing detection areas. Model IV exhibited an excellent effect in small target detection, but there was also a certain amount of false detection areas.

**Figure 22.** Comparison of different modules in test area C.

The quantitative evaluation indicators of the four models in area C are presented in Table 5. The detection effects of the four models decreased with increasing image interference. For the accuracy, the InSARNet still had the highest accuracy (0.88), which was nearly 5% higher than the second-highest accuracy (model II), but the accuracy of model I was only 0.76. For the recall rate, the effect of InSARNet was also significantly better than those of the other three models. Its recall rate exceeded 0.85, which was 5% higher than that of model III, which had the second-highest recall rate. In particular, the kappa coefficient of InSARNet exceeded 0.75, which was much higher than those of the other three models. The overall accuracies of most of the models exceeded 0.80, except for model I (0.76). This shows that although the four models completed the basic task of

detection, the detection effect gradually varied with the complexity of the situation, and InSARNet exhibited unique advantages in the area with more noise.

Table 5. Comparison of the accuracies of the models with different module combinations in test area C.

Model	Precision	Recall	F1	Kappa	OA
I	0.7616	0.7667	0.7641	0.5250	0.7633
II	0.8386	0.7967	0.8171	0.6544	0.8217
III	0.8013	0.8200	0.8105	0.6111	0.8083
IV	0.8822	0.8733	0.8777	0.7585	0.8783

For object detection, the accuracy of the detection needs to be guaranteed, but the performance and actual needs of the computer, the operation speed of the model, and the space it occupies are also important indicators. In this study, the parameters, floating-point operations per second (FLOPS), training time, and testing time were selected to compare and analyze the models.

The results are compared in Table 6. The results show that the training time and test time were significantly improved by the introduction of the involution operator. Regarding the complexity of the model, compared with model II (ResNet50 + deformable ROI pooling), the number of parameters and FLOPS of model I (Mask RCNN) were slightly higher, and the corresponding training time was longer. However, the number of parameters and FLOPS of model III (Rednet50 + ROI Align), which includes a revolution operator, were greatly improved, the number of parameters was reduced by 29.8%, and the FLOPS were reduced by 10.62%. In addition, model IV (InSARNet), which has an involution operator and deformable ROI pooling module, the number of parameters and FLOPS were larger than those of model III, and the calculation time was slightly longer. However, according to the requirements for accuracy comparison and deformation anomaly area detection, it is feasible and worth sacrificing the operation speed and storage space to obtain higher accuracy.

Table 6. Comparison of parameters, floating-point operations per second (FLOPS), and training and testing times of the different modules.

Model	Param.	FLOPS	Training Time	Testing Time	Testing Time	Testing Time
	(M)	(M)	(ms/Step)	(s/Image A)	(s/Image B)	(s/Image C)
I	37.41	334.02	253	2.56	3.14	3.26
II	46.15	351.22	262	2.65	3.37	3.58
III	26.34	290.59	144	1.85	1.64	1.52
IV	32.18	298.53	157	1.81	1.96	2.27

4.4.2. Comparison of Models

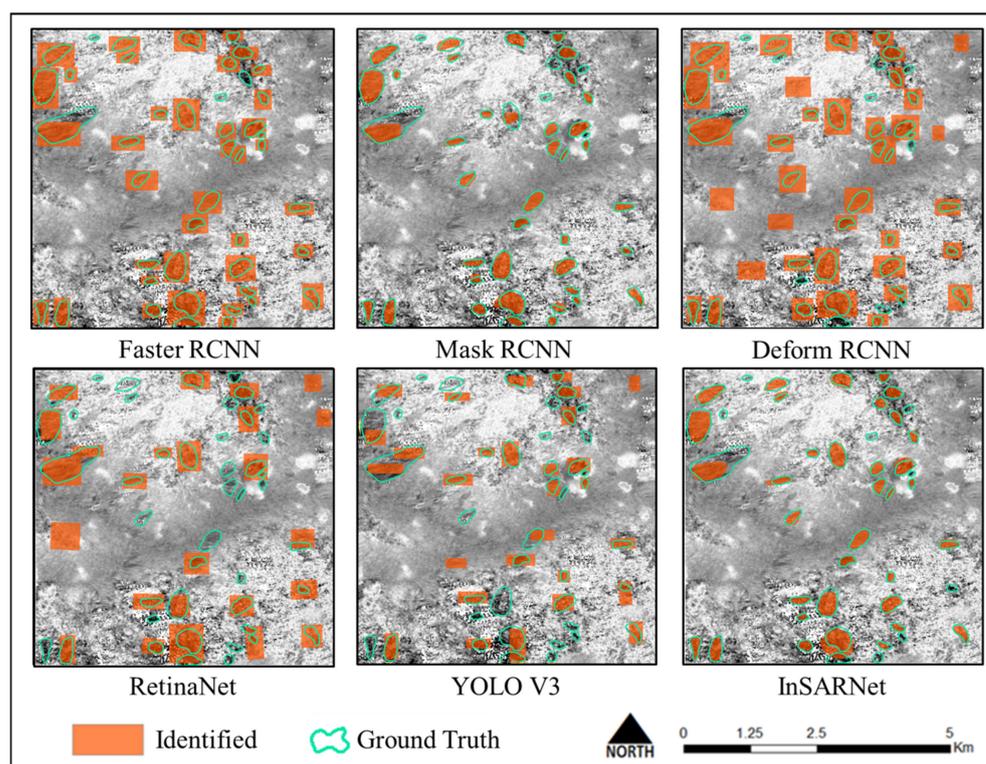
To verify the detection effect of InSARNet in regions with various shapes, five target detection models with strong universality and good effects, which have been widely verified and widely recognized, were selected for comparative analysis. The frameworks of the five models and InSARNet are described in Table 7. To ensure the fairness of the test, all of the models were trained using the same dataset, and no pre-training parameters were used in the training process. To display the results, the three sample areas described in Section 4.4.1 were also used for the comparative analysis.

Table 7. Detection models and their structures.

Model	Backbone	Neck	Head
Faster- ¹ RCNN	ResNet-50	FPN	RPN Head + ROI Pooling
Mask- ¹ RCNN	ResNet-50	FPN	RPN Head + ROI Align
Deform- ¹ RCNN	ConvNet-50	FPN	RPN Head + ROI Align
RetinaNet	ResNet-50	FPN	Retina Head
YOLO V3	Darknet-53	YOLO V3	YOLO V3
InSARNet	RedNet-50	FPN	RPN Head + Deformable ROI pooling

¹ RCNN: Region Convolutional Neural Networks.

The detection results of each detection network in area A are shown in Figure 23. Unlike the Mask RCNN and InSARNet, the other four detection models are not equipped with a mask module, so their detection results are rectangular target frames. In terms of the detection effect, the false detection rate of the one-stage model was higher, and the accuracy of the two-stage detection model was higher than that of the one-stage detection model. In area A, which had low noise, the accuracy of the two-stage detection model was almost the same, and the detection accuracy of InSARNet was slightly higher than those of the other two-stage models.

**Figure 23.** Comparison of different models in test area A.

The evaluation of the quantitative detection accuracies of the different networks in area A is shown in Table 8. The various accuracy evaluation indexes of InSARNet were better than those of the other models, which demonstrates that it had the best effect in the identification of the InSAR anomalous deformation areas. Among them, the precisions of all of the models were >80%, indicating that in area A, all of the models could identify most of the anomalous deformation areas. Among them, the accuracies of the two one-stage monitoring models were slightly lower than those of the other two-stage monitoring models, and InSARNet was the only model with an accuracy of >90%. Similarly, for the recall index, InSARNet had the highest recall rate (93.67%). Those of the Mask RCNN and

Deform RCNN were also $>90\%$, while the recall rates of the other detection models were relatively low and were concentrated at about 85% . The F1 coefficient of InSARNet was the best (0.93). For the kappa coefficient, InSARNet also had unique advantages (0.86), while those of the RetinaNet, Yolo V3, and Faster RCNN were less than 0.7. The OA values of the Mask RCNN and InSARNet exceeded 0.9, and those of the other models were also good (>0.8).

Table 8. Comparison of the accuracies of the different detection models in test area A.

Models	Precision	Recall	F1	Kappa	OA
Faster RCNN	0.8328	0.8467	0.8396	0.6729	0.8383
Mask RCNN	0.8984	0.9133	0.9058	0.8074	0.9050
Deform RCNN	0.8658	0.9033	0.8841	0.7552	0.8816
RetinaNet	0.8267	0.8547	0.8405	0.6763	0.8389
YOLO V3	0.8415	0.8500	0.8457	0.6878	0.8450
InSARNet	0.9274	0.9367	0.9320	0.8621	0.9317

The detection results of the different detection networks in area B are shown in Figure 24. As the noise increased, the false detection rates and missed detection rates of all of the models increased significantly. InSARNet still had a high detection accuracy. Among them, due to the poor recognition effect of small targets of the one-stage network, the missed detection rate increased significantly.

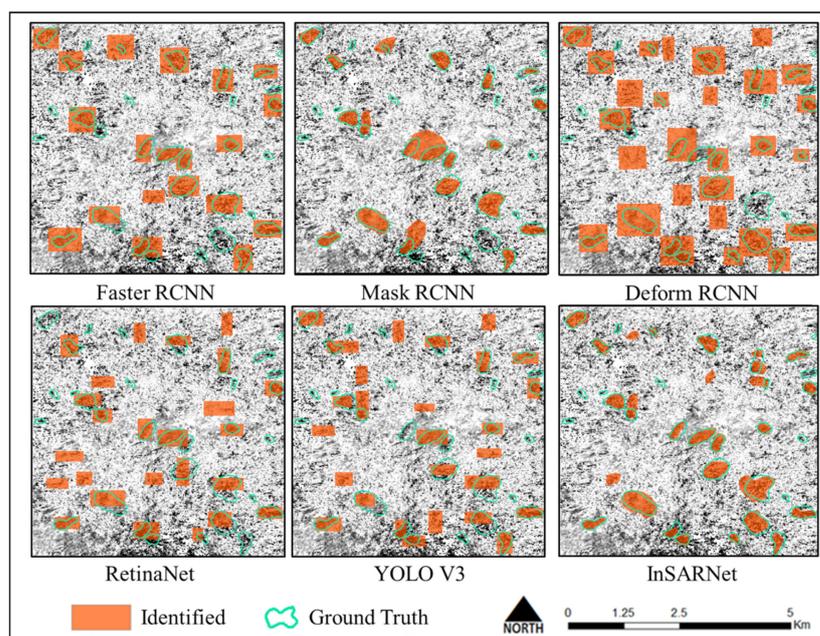


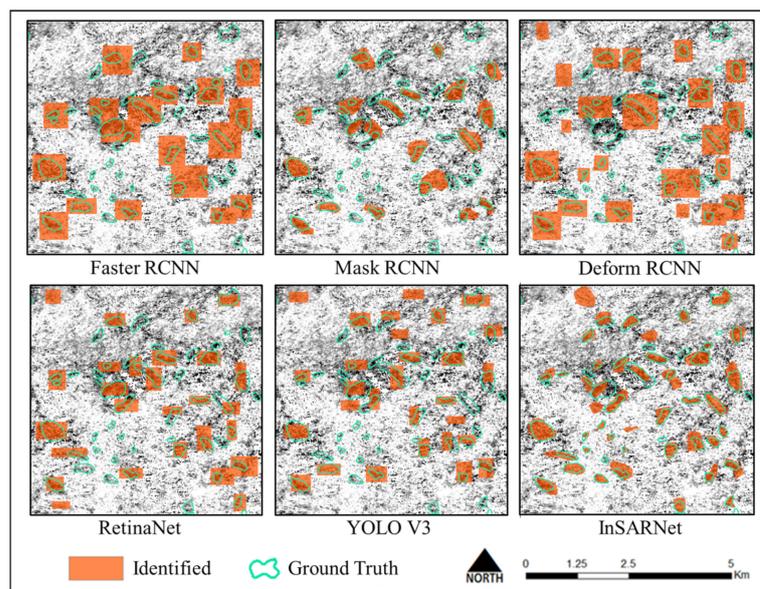
Figure 24. Comparison of different models in test area B.

The quantitative evaluation of the accuracies of the different detection models in area B is presented in Table 9. The results of the quantitative evaluation also confirm the detection ability of InSARNet in complex images. In region B, which had more noise, the indexes of all of the models were lower than in region A. The precision of the one-stage model was less than 0.8 due to its limited ability to identify small anomalous deformation areas. Among the other models, InSARNet had the highest accuracy (0.89). For the recall index and F1 coefficient, InSARNet still had the best performance (>0.9). For the kappa coefficient, InSARNet (0.81) had an absolute advantage over the other models. Of the other models, only the Deform RCNN had a value of >0.7 . For the overall accuracy index, except for that of the RetinaNet (<0.8), the other models had good detection abilities.

Table 9. Comparison of the accuracies of the different detection models in test area B.

Models	Precision	Recall	F1	Kappa	OA
Faster RCNN	0.8026	0.8000	0.8013	0.6041	0.8016
Mask RCNN	0.8289	0.8400	0.8344	0.6637	0.8333
Deform RCNN	0.8614	0.8500	0.8557	0.7160	0.8566
RetinaNet	0.7631	0.7733	0.7682	0.5299	0.7666
YOLO V3	0.7901	0.8533	0.8205	0.6069	0.8133
InSARNet	0.8990	0.9200	0.9094	0.8131	0.9083

The detection results of the detection networks in area C are shown in Figure 25. For area C, which had more noise and deformation anomaly areas with different sizes, the missed detection rate and false detection rate of each model were further improved. Except for the Mask RCNN and InSARNet, the other models were generally sensitive to the large anomalous deformation areas, while for smaller areas, they suffered from missed detection and/or target frame mismatch. The reason for this phenomenon is that the models identify these areas as anomalous deformation areas at different scales, thus retaining the largest edge frame. Its essence is the false detection of anomalous deformation areas.

**Figure 25.** Comparison of different models in test area C.

The quantitative evaluation indexes of the accuracies of the different detection models in area C are presented in Table 10. In area C, which had a lower image quality, as the accuracy index of each model continued to decrease, the advantage of InSARNet was further demonstrated. While these indexes remained optimal, the accuracy decreased slightly. Specifically, in terms of accuracy and recall, the average decrease of the other models was 6–7%, but that of InSARNet only decreased by 4%. For the F1 coefficient, only the values of the Deform RCNN and InSARNet remained above 0.8, and those of the other models were only about 0.75. The kappa coefficients of the models significantly decreased, and only that of InSARNet remained above 0.7. In particular, the kappa coefficient of the RetinaNet was less than 0.4. The gap between the overall accuracies of the various models and that of InSARNet gradually widened. Those of InSARNet and Deform RCNN were still >0.8 , so they had a good detection effect. The overall accuracies of the Faster RCNN, Mask RCNN, and Yolo V3 were about 0.75, which were lower than in the other two test areas. The overall accuracy of the RetinaNet was less than 0.7, indicating that it could not effectively detect the anomalous deformation areas under the condition of more noise interference.

Table 10. Comparison of the different detection models in test area C.

Models	Precision	Recall	F1	Kappa	OA
Faster RCNN	0.7684	0.7300	0.7487	0.5221	0.7550
Mask RCNN	0.7616	0.7667	0.7641	0.5250	0.7633
Deform RCNN	0.8148	0.8066	0.8107	0.6256	0.8116
RetinaNet	0.6953	0.7000	0.6976	0.3917	0.6966
YOLO V3	0.7361	0.7533	0.7446	0.4774	0.7416
InSARNet	0.8822	0.8733	0.8777	0.7585	0.8783

5. Discussion

Through analysis of the performances of the modules and evaluation of the accuracies of the different models, it was confirmed that InSARNet is feasible, and it can effectively delineate large-scale InSAR anomalous deformation areas. Compared with the other object detection models, InSARNet had the best accuracy and stability.

There are two reasons why we selected the object detection model to identify the anomalous deformation areas of landslides. First, the anomalous deformation areas extracted using InSAR do not correspond to potential landslides. The causes of surface deformation are diverse, and vegetation growth, human activities, and other factors may also cause surface deformation. Therefore, it is necessary to screen non-potential landslides using basic geographic information, geological information, and other factors. However, the anomalous deformation areas after screening still corresponded to the boundary of the abnormal sliding area, so it is inaccurate to regard it as the boundary of the entire landslide. Using the segmentation model for recognition is a waste of efficiency and computational power. Second, for the general landslide investigation stage, the primary task is to determine the location of the potential landslide. Compared with the shape of the landslide, it is more important to determine the location of the landslide. In the investigation and verification stages, using unmanned aerial vehicle (UAV) surveys and airborne light detection and ranging (Lidar) or ground-based 3-D laser scanning technology, combined with optical images and field investigations, the boundary and shape of the slope can be obtained more accurately, and more targeted and accurate treatment and protection can be achieved.

The InSARNet model is still in the preliminary stage. At present, it is only used as the detection of anomalous deformation areas, and there is still great potential for improvement in the future. We improved and perfected the model along two sides. (1) With the addition of different vegetation and land lithology, the recognition ability under different conditions can be improved through the learning of different vegetation indexes. Different surface environment and vegetation coverage are important factors affecting the accuracy of InSAR results. LULC and NDVI are two typical data indicating surface type and vegetation cover [36]. More landslides and corresponding data will be collected in the future, which makes the model more applicable. (2) Adding optical image and geological information, combined with a variety of information, realizes the automatic extraction of potential landslide. The causes of landslides are diverse. Surface deformation is only a response to incentives. Therefore, starting from the inducement of landslide, its geological structure, lithology, rainfall and temperature change are all potential factors to improve the accuracy of the model. In the future, we will try to input such data into the model as a factor to improve the accuracy of model recognition.

In addition, the current definition of a potential landslide is a slope that may cause harm to human production and life or roads and rivers. Therefore, the distribution of roads and rivers were taken as the center, a 1 km buffer zone was established, and the anomalous deformation areas in the buffer zone were identified as potential landslides (Figure 26). A total of 98 potential landslide points were identified. Comparing the identification results with the potential landslides provided by the Ministry of Natural Resources of China, a total of 92 identification results belong to the known potential landslides, and the identification accuracy rate is 93.88%. The results prove the accuracy of the InSARNet model.

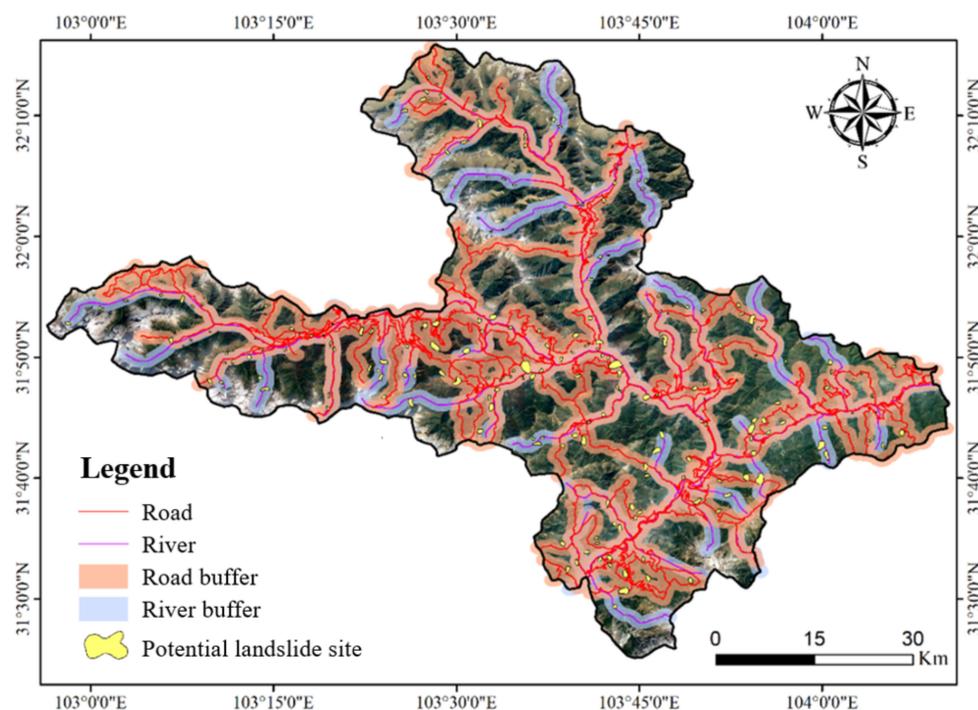


Figure 26. Extraction of potential landslide site in the study area.

The buffer method selected in this study is based on the consideration of actual needs. The treatment of potential landslide is to ensure the safety of personnel, roads and rivers. To identify potential landslides more accurately, the landslides map with the slope map defined according to thresholds recognized significant for triggering landslides is a more accurate method [57]. Deformation analysis based on landslide unit can identify landslide boundary more accurately, but there may be some misjudgment [58]. According to different needs, combined with the results of InSARNet model, appropriate methods can be selected to classify potential landslides.

After identifying the potential landslide, it is also meaningful to model and numerically simulate the landslide, according to the deformation rate and geological data, in order to assess the maximum run out and volume deposition [59]. This would provide a reliable and useful landslide map [60]. It can also provide more accurate and effective information to the local government and on-site investigators, so as to realize the effective treatment and protection of potential landslides.

6. Conclusions

A two-stage target detection model (InSARNet) was developed in this study to overcome the problem of the time-consuming and laborious nature of manual delineation and the low accuracy of large-scale InSAR anomalous deformation areas, in addition to achieving automatic extraction of InSAR anomalous deformation areas. Based on the Mask RCNN, InSARNet constructed a model suitable for InSAR detection of anomalous deformation areas by introducing an involution operator and a deformable ROI pooling module. Compared with other existing models, InSARNet had a significantly better detection accuracy and anti-noise ability. Based on the identified anomalous deformation areas, supplemented by geographical information such as rivers and roads, the potential landslides were divided to provide scientific theoretical support for local landslide treatment.

The experimental results showed that (1) InSARNet could effectively extract anomalous deformation areas from complex InSAR results, and its overall accuracy is about 90%; (2) after the introduction of the RedNet module, including the involution operator, the number of parameters and calculations of InSARNet were reduced by about 30%, and its detection accuracy was also slightly improved (by 1%); (3) after introducing the deformable

convolution module, the ability of the model to recognize small-scale deformation anomaly areas was improved, and the overall accuracy was improved by 3–4%; (4) by comparing and analyzing InSARNet with the commonly used one-stage detector and two-stage detector models, it was found that its detection accuracy was better than all of the models evaluated. Although its operation efficiency was slightly lower than that of the one-stage detector, it was worth sacrificing a little speed to obtain higher accuracy in combination with the accuracy factors and the recognition requirements of the anomalous deformation area.

Landslides are still one type of natural disaster that causes a large number of casualties every year. InSAR has been the basis of numerous achievements in the field of landslide monitoring. In practical applications, large-scale, periodic, and systematic InSAR monitoring is a common method at present. However, after calculating a large number of InSAR results, InSAR takes considerable time to manually delineate the potential landslides, and the standard is not unified. The application of deep learning in the identification of landslides enables the quick delineation of potential landslides, improves the identification efficiency, reduces labor costs, and realizes the systematization and automatic identification of potential landslides. As such, it is the inevitable trend for the development of landslide identification in the future.

Author Contributions: T.Z. designed this study. T.Z. performed the data collection, processing and analysis. D.C. optimized the figure for this work. Y.Y. and X.W. participated in the design of model code. The corresponding author W.Z. is supervisor of this work and contributed with continuous guidance during this work. T.Z. jointly wrote this manuscript, and the manuscript was edited by W.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This study was jointly financed by the Key R & D and Transformation Program of Qinghai Province [Grant No. 2020-SF-C37] and the Program of the Department of Geological Exploration Management, Ministry of Natural Resources, China [Grant No. 0733-201808076/2].

Acknowledgments: The authors are grateful to the anonymous reviewers for their constructive comments and suggestions to improve this manuscript.

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

Appendix A

Table A1. Sentinel-1 Data List.

No.	Data	Orbit	No.	Data	Orbit	No.	Data	Orbit
1	2015-01-18	Ascending	29	2016-11-20	Ascending	57	2015-10-16	Descending
2	2015-02-11	Ascending	30	2016-12-14	Ascending	58	2015-11-09	Descending
3	2015-03-07	Ascending	31	2017-01-07	Ascending	59	2015-12-03	Descending
4	2015-03-31	Ascending	32	2017-01-31	Ascending	60	2015-12-27	Descending
5	2015-04-24	Ascending	33	2017-02-12	Ascending	61	2016-01-20	Descending
6	2015-05-18	Ascending	34	2017-02-24	Ascending	62	2016-02-13	Descending
7	2015-06-11	Ascending	35	2017-03-20	Ascending	63	2016-03-08	Descending
8	2015-07-05	Ascending	36	2017-04-01	Ascending	64	2016-04-01	Descending
9	2015-07-29	Ascending	37	2017-04-13	Ascending	65	2016-04-25	Descending
10	2015-08-22	Ascending	38	2017-04-25	Ascending	66	2016-05-19	Descending
11	2015-09-15	Ascending	39	2017-05-07	Ascending	67	2016-06-12	Descending
12	2015-10-09	Ascending	40	2017-05-19	Ascending	68	2016-07-30	Descending
13	2015-11-02	Ascending	41	2017-05-31	Ascending	69	2016-08-23	Descending
14	2015-11-26	Ascending	42	2017-06-12	Ascending	70	2016-09-16	Descending
15	2015-12-20	Ascending	43	2017-06-24	Ascending	71	2017-02-19	Descending
16	2016-01-13	Ascending	44	2017-07-06	Ascending	72	2017-03-03	Descending
17	2016-02-06	Ascending	45	2017-07-18	Ascending	73	2017-03-15	Descending
18	2016-03-01	Ascending	46	2017-07-30	Ascending	74	2017-03-27	Descending
19	2016-03-25	Ascending	47	2015-01-13	Descending	75	2017-04-08	Descending
20	2016-04-18	Ascending	48	2015-02-06	Descending	76	2017-04-20	Descending
21	2016-05-12	Ascending	49	2015-03-14	Descending	77	2017-05-02	Descending
22	2016-06-05	Ascending	50	2015-04-07	Descending	78	2017-05-14	Descending
23	2016-06-29	Ascending	51	2015-05-01	Descending	79	2017-05-26	Descending
24	2016-7-23	Ascending	52	2015-5-25	Descending	80	2017-6-7	Descending
25	2016-8-16	Ascending	53	2015-6-18	Descending	81	2017-6-19	Descending
26	2016-9-9	Ascending	54	2015-7-12	Descending	82	2017-7-13	Descending
27	2016-10-3	Ascending	55	2015-8-5	Descending	83	2017-7-25	Descending
28	2016-10-27	Ascending	56	2015-8-29	Descending			

References

1. Tric, E.; Lebourg, T.; Jomard, H.; Le Cossec, J. Study of large-scale deformation induced by gravity on the La Clapière landslide (Saint-Etienne de Tinée, France) using numerical and geophysical approaches. *J. Appl. Geophys.* **2010**, *70*, 206–215. [CrossRef]
2. Mudda, L.P.; Borigarla, B.; Pydi, R.; Maddala, P. Identification of landslide/Man-made structures along transboundary rivers. *Mater. Today Proc.* **2020**, *138*, 609–616. [CrossRef]
3. Ibeh, C.U. Effect of changing groundwater level on shallow landslide at the basin scale: A case study in the Odo basin of south eastern Nigeria. *J. African Earth Sci.* **2020**, *165*, 103773. [CrossRef]
4. Tsaparas, I.; Rahardjo, H.; Toll, D.G.; Leong, E.C. Controlling parameters for rainfall-induced landslides. *Comput. Geotech.* **2002**, *29*, 1–27. [CrossRef]
5. Marin, R.J.; García, E.F.; Aristizábal, E. Effect of basin morphometric parameters on physically-based rainfall thresholds for shallow landslides. *Eng. Geol.* **2020**, *278*, 105855. [CrossRef]
6. Keefer, D.K. Statistical analysis of an earthquake-induced landslide distribution—The 1989 Loma Prieta, California event. *Eng. Geol.* **2000**, *58*, 231–249. [CrossRef]
7. Sridharan, A.; Gopalan, S. Correlations among properties of lithological units that contribute to earthquake induced landslides. *Mater. Today Proc.* **2020**, *33*, 2402–2406. [CrossRef]
8. Chen, M.; Tang, C.; Xiong, J.; Shi, Q.Y.; Li, N.; Gong, L.F.; Wang, X.D.; Tie, Y. The long-term evolution of landslide activity near the epicentral area of the 2008 Wenchuan earthquake in China. *Geomorphology* **2020**, *367*, 107317. [CrossRef]
9. Margielewski, W.; Krapiec, M.; Valde-Nowak, P.; Zernitskaya, V. A Neolithic yew bow in the Polish Carpathians. Evidence of the impact of human activity on mountainous palaeoenvironment from the Kamiennik landslide peat bog. *Catena* **2010**, *80*, 141–153. [CrossRef]
10. Persichillo, M.G.; Bordoni, M.; Cavalli, M.; Crema, S.; Meisina, C. The role of human activities on sediment connectivity of shallow landslides. *Catena* **2018**, *160*, 261–274. [CrossRef]
11. Froude, M.J.; Petley, D.N. Global fatal landslide occurrence from 2004 to 2016. *Nat. Hazards Earth Syst. Sci.* **2018**, *18*, 2161–2181. [CrossRef]
12. Ministry of Natural Resources PRC. China Geological Disasters Bulletin. 2020. Available online: http://www.gov.cn/xinwen/2020-01/17/content_5470130.htm (accessed on 17 January 2020).
13. Xu, Q.; Dong, X.; Li, W. Integrated space-air-ground early detection, monitoring and warning system for potential catastrophic geohazards. *Geomat. Inf. Sci. Wuhan Univ.* **2019**, *44*, 957–966. [CrossRef]
14. Pan, Y.; Shen, W.B.; Shum, C.K.; Chen, R. Spatially varying surface seasonal oscillations and 3-D crustal deformation of the Tibetan Plateau derived from GPS and GRACE data. *Earth Planet. Sci. Lett.* **2018**, *502*, 12–22. [CrossRef]
15. Pei, H.; Zhang, F.; Zhang, S. Development of a novel Hall element inclinometer for slope displacement monitoring. *Meas. J. Int. Meas. Confed.* **2021**, *181*, 109636. [CrossRef]
16. Zhang, Y.; Tang, H.; Li, C.; Lu, G.; Cai, Y.; Zhang, J.; Tan, F. Design and testing of a flexible inclinometer probe for model tests of landslide deep displacement measurement. *Sensors* **2018**, *18*, 224. [CrossRef]
17. Su, M.-B.; Chen, I.-H.; Liao, C.-H. Using TDR cables and GPS for landslide monitoring in high mountain area. *J. Geotech. Geoenviron. Eng.* **2009**, *135*, 1113–1121. [CrossRef]
18. Liu, Y.; Li, W.; He, J.; Liu, S.; Cai, L.; Cheng, G. Application of Brillouin optical time domain reflectometry to dynamic monitoring of overburden deformation and failure caused by underground mining. *Int. J. Rock Mech. Min. Sci.* **2018**, *106*, 133–143. [CrossRef]
19. Caviedes-Voullième, D.; Juez, C.; Murillo, J.; García-Navarro, P. 2D dry granular free-surface flow over complex topography with obstacles. Part I: Experimental study using a consumer-grade RGB-D sensor. *Comput. Geosci.* **2014**, *73*, 177–197. [CrossRef]
20. Huan, L.; Zheng, X.; Gong, J. GeoRec: Geometry-enhanced semantic 3D reconstruction of RGB-D indoor scenes. *ISPRS J. Photogramm. Remote Sens.* **2022**, *186*, 301–314. [CrossRef]
21. Bénédicte, F.; Christophe, D.; José, A. Observation and modelling of the Saint-Etienne-de-Tinée landslide using SAR interferometry. *Eur. Sp. Agency (Special Publ. ESA SP)* **1997**, *265*, 21–27.
22. Ferretti, A.; Prati, C.; Rocca, F. Permanent scatterers in SAR interferometry. *IEEE Trans. Geosci. Remote Sens.* **2001**, *39*, 8–20. [CrossRef]
23. Berardino, P.; Fornaro, G.; Lanari, R.; Sansosti, E. A new algorithm for surface deformation monitoring based on small baseline differential SAR interferograms. *IEEE Trans. Geosci. Remote Sens.* **2002**, *40*, 2375–2383. [CrossRef]
24. Zhang, T.; Xie, S.; Fan, J.; Huang, B.; Wang, Q.; Yuan, W.; Zhao, H.; Chen, J.; Li, H.; Liu, G.; et al. Detection of active landslides in Southwest China systems and ALOS-2 data detection of active landslides in Southwest China using ALOS-2 data. *Int. Conf. Heal. Soc. Care Inf. Syst. Technol.* **2021**, *181*, 1138–1145. [CrossRef]
25. Liu, P.; Li, Z.; Hoey, T.; Kincal, C.; Zhang, J.; Zeng, Q.; Muller, J.P. Using advanced InSAR time series techniques to monitor landslide movements in Badong of the Three Gorges region, China. *Int. J. Appl. Earth Obs. Geoinf.* **2012**, *21*, 253–264. [CrossRef]
26. Shi, X.; Yang, C.; Zhang, L.; Jiang, H.; Liao, M.; Zhang, L.; Liu, X. Mapping and characterizing displacements of active loess slopes along the upstream Yellow River with multi-temporal InSAR datasets. *Sci. Total Environ.* **2019**, *674*, 200–210. [CrossRef]
27. Wang, Y.; Dong, J.; Zhang, L.; Zhang, L.; Deng, S.; Zhang, G.; Liao, M.; Gong, J. Refined InSAR tropospheric delay correction for wide-area landslide identification and monitoring. *Remote Sens. Environ.* **2022**, *275*, 113013. [CrossRef]
28. Zhang, T.; Zhang, W.; Yang, R.; Liu, Y.; Jafari, M. CO₂ capture and storage monitoring based on remote sensing techniques: A review. *J. Clean. Prod.* **2020**, *281*, 124409. [CrossRef]

29. Guzzetti, F.; Reichenbach, P.; Cardinali, M.; Galli, M.; Ardizzone, F. Probabilistic landslide hazard assessment at the basin scale. *Geomorphology* **2005**, *72*, 272–299. [[CrossRef](#)]
30. Legorreta Paulin, G.; Bursik, M.; Lugo-Hubp, J.; Zamorano Orozco, J.J. Effect of pixel size on cartographic representation of shallow and deep-seated landslide, and its collateral effects on the forecasting of landslides by SINMAP and Multiple Logistic Regression landslide models. *Phys. Chem. Earth* **2010**, *35*, 137–148. [[CrossRef](#)]
31. De Luiz Rosito Listo, F.; Villaça Gomes, M.C.; Ferreira, F.S. Evaluation of shallow landslide susceptibility and Factor of Safety variation using the TRIGRS model, Serra do Mar Mountain Range, Brazil. *J. South Am. Earth Sci.* **2021**, *107*, 103011. [[CrossRef](#)]
32. Jin, J.; Chen, G.; Meng, X.; Zhang, Y.; Shi, W.; Li, Y.; Yang, Y.; Jiang, W. Prediction of river damming susceptibility by landslides based on a logistic regression model and InSAR techniques: A case study of the Bailong River Basin, China. *Eng. Geol.* **2022**, *299*, 106562. [[CrossRef](#)]
33. Marjanović, M.; Kovačević, M.; Bajat, B.; Voženilek, V. Landslide susceptibility assessment using SVM machine learning algorithm. *Eng. Geol.* **2011**, *123*, 225–234. [[CrossRef](#)]
34. Gameiro, S.; Riffel, E.S.; de Oliveira, G.G.; Guasselli, L.A. Artificial neural networks applied to landslide susceptibility: The effect of sampling areas on model capacity for generalization and extrapolation. *Appl. Geogr.* **2021**, *137*, 102598. [[CrossRef](#)]
35. 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](#)]
36. Erin, L.; Regula, F.; Denise, R.; Lorenzo, N.; Lena, R.; James, S.; Steinar, N. Multi-Temporal Satellite Image Composites in Google Earth Engine for Improved Landslide Visibility: A Case Study of a Glacial Landscape. *Remote Sens.* **2022**, *14*, 2301.
37. Solari, L.; Del Soldato, M.; Raspini, F.; Barra, A.; Bianchini, S.; Confuorto, P.; Casagli, N.; Crosetto, M. Review of satellite interferometry for landslide detection in Italy. *Remote Sens.* **2020**, *12*, 1351. [[CrossRef](#)]
38. Ma, Z.; Mei, G. Deep learning for geological hazards analysis: Data, models, applications, and opportunities. *Earth-Sci. Rev.* **2021**, *223*, 103858. [[CrossRef](#)]
39. Hinton, G.E.; Salakhutdinov, R.R. Reducing the dimensionality of data with neural networks. *Science* **2006**, *313*, 504–507. [[CrossRef](#)]
40. Alex, K.; Ilya, S.; Geoffrey, E.H. ImageNet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems*; Curran Associates, Inc.: Red Hook, NY, USA, 2012; pp. 1–9.
41. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. In Proceedings of the International Conference on Learning Representations, ICLR 2015, Conference Track Proceedings. San Diego, CA, USA, 7–9 May 2015.
42. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778. [[CrossRef](#)]
43. Girshick, R. Fast R-CNN. *Computerence* **2015**, 1440–1448.
44. Chen, L.C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *IEEE Trans. Pattern Anal. Mach. Intell.* **2018**, *40*, 834–848. [[CrossRef](#)]
45. Liu, X.; Chen, Q.; Zhao, J.; Xu, Q.; Luo, R.; Zhang, Y.; Yang, Y.; Liu, G. The spatial response pattern of coseismic landslides induced by the 2008 Wenchuan earthquake to the surface deformation and Coulomb stress change revealed from InSAR observations. *Int. J. Appl. Earth Obs. Geoinf.* **2020**, *87*, 102030. [[CrossRef](#)]
46. Yuan, X.; Shi, J.; Gu, L. A review of deep learning methods for semantic segmentation of remote sensing imagery. *Expert Syst. Appl.* **2021**, *169*, 114417. [[CrossRef](#)]
47. Sharma, V.; Mir, R.N. A comprehensive and systematic look up into deep learning based object detection techniques: A review. *Comput. Sci. Rev.* **2020**, *38*, 100301. [[CrossRef](#)]
48. Tiwari, A.; Narayan, A.B.; Dikshit, O. Deep learning networks for selection of measurement pixels in multi-temporal SAR interferometric processing. *ISPRS J. Photogramm. Remote Sens.* **2020**, *166*, 169–182. [[CrossRef](#)]
49. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; pp. 1–9. [[CrossRef](#)]
50. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely connected convolutional networks. In Proceedings of the 30th 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 2261–2269. [[CrossRef](#)]
51. Zhao, S.; Chigira, M.; Wu, X. Gigantic rockslides induced by fluvial incision in the Diexi area along the eastern margin of the Tibetan Plateau. *Geomorphology* **2019**, *338*, 27–42. [[CrossRef](#)]
52. Xu, H.; Jiang, H.; Yu, S.; Yang, H.; Chen, J. OSL and pollen concentrate ¹⁴C dating of dammed lake sediments at Maoxian, east Tibet, and implications for two historical earthquakes in AD 638 and 952. *Quat. Int.* **2015**, *371*, 290–299. [[CrossRef](#)]
53. Huang, D.; Li, Y.Q.; Song, Y.X.; Xu, Q.; Pei, X.J. Insights into the catastrophic Xinmo rock avalanche in Maoxian county, China: Combined effects of historical earthquakes and landslide amplification. *Eng. Geol.* **2019**, *258*, 105158. [[CrossRef](#)]
54. Zhao, S.; Chigira, M.; Wu, X. Buckling deformations at the 2017 Xinmo landslide site and nearby slopes, Maoxian, Sichuan, China. *Eng. Geol.* **2018**, *246*, 187–197. [[CrossRef](#)]

55. Mingsheng, W.; Hongqi, C.; Mingzhi, Z.; Hongliang, C.; Wenpei, W.; Nan, Z.; Zhe, H. Characteristics and formation mechanism analysis of the “6·24” catastrophic landslide of the June 24 of 2017, at Maoxian, Sichuan. *Chin. J. Geol. Hazard Control* **2017**, *28*, 1–7. [[CrossRef](#)]
56. Li, D.; Hu, J.; Wang, C.; Li, X.; She, Q.; Zhu, L.; Zhang, T.; Chen, Q. Involution: Inverting the inherence of convolution for visual recognition. In Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 20–25 June 2021. [[CrossRef](#)]
57. Wang, F.; Xu, P.; Wang, C.; Wang, N.; Jiang, N. Application of a gis-based slope unit method for landslide susceptibility mapping along the longzi river, southeastern tibetan plateau, China. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 172. [[CrossRef](#)]
58. Spinetti, C.; Bisson, M.; Tolomei, C.; Colini, L.; Galvani, A.; Moro, M.; Saroli, M.; Sepe, V. Landslide susceptibility mapping by remote sensing and geomorphological data: Case studies on the Sorrentina Peninsula (Southern Italy). *GISci. Remote Sens.* **2019**, *56*, 940–965. [[CrossRef](#)]
59. Zhang, W.; Xiao, D. Numerical analysis of the effect of strength parameters on the large-deformation flow process of earthquake-induced landslides. *Eng. Geol.* **2019**, *260*, 105239. [[CrossRef](#)]
60. Juez, C.; Caviedes-Voullième, D.; Murillo, J.; García-Navarro, P. 2D dry granular free-surface transient flow over complex topography with obstacles. Part II: Numerical predictions of fluid structures and benchmarking. *Comput. Geosci.* **2014**, *73*, 142–163. [[CrossRef](#)]