This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully edited content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2023.3339294

A multi-input channel U-Net landslide detection method fusing SAR multi-source remote sensing data

Hesheng Chen, Yi He, Lifeng Zhang, Wang Yang, Yaoxiang Liu, Binghai Gao, Qing Zhang, and Jiangang Lu

***Abstract*—Accurate and efficient landslide identification is an important basis for landslide disaster prevention and control. Due to the diversity of landslide features, vegetation occlusion, and the complexity of the surrounding surface environment in remote sensing images, deep learning models (such as U-Net) for landslide detection based only on optical remote sensing images will lead to false and missed detection. The detection accuracy is not high, and it is difficult to satisfy the demand. SAR has penetrability, and SAR images are highly sensitive to changes in surface morphology and structure. In this study, a multi-input channel U-Net landslide detection method fusing SAR, optical, and topographic multi-source remote sensing data is proposed. Firstly, a multi-input channel U-Net model fusing SAR multi-source remote sensing data is constructed, then an attention mechanism is introduced into the multi-input channel U-Net to adjust the spatial weights of the feature maps of the multi-source data to emphasize the landslide-related features, and finally, the proposed model is applied to the experimental scene for validation. The experimental results demonstrate that the proposed model combined with SAR multi-source remote sensing data improves the perception ability of landslide features, focuses on learning landslide-related features, improves the accuracy of landslide detection, and reduces the rate of false detections and missed detections. Compared with the traditional U-Net landslide detection method based on SAR multi-source remote sensing data and the traditional U-Net method that disregards SAR multi-source remote sensing data, the proposed method has the best quantitative evaluation indicators. Among them, the proposed method obtained the highest F1 value (66.18%), indicating that fused SAR remote sensing data can provide rich and complementary landslide feature information, simultaneously setting up a multi-channel U-Net model to input multi-source remote sensing data can effectively process landslide feature information. The proposed method can provide theoretical and technical support for landslide disaster prevention and control.**

***Index Terms*—SAR remote sensing data, Feature learning, Landslide identification, U-Net, Attention mechanism.**

I. Introduction

LANDSLIDES are considered one of the most serious natural disasters in the world, in which soil, rocks, and objects located above or below unstable slopes are affected by various factors such as rainfall, earthquakes, and human activities to move downwards and form landslides [1-4]. Landslide characteristics are strong suddenness, high harm, causing many casualties and economic losses, and causing damage to the surface environment [5-8]. After a landslide occurs, accurately obtaining and recording information on the location and size of the landslide is critical for providing disaster emergency response, disaster damage estimation, and post-disaster reconstruction [9, 10]. In addition, landslide detection is extremely important for updating landslide inventory datasets with accurate location and extent information and is also necessary for further landslide sensitivity modeling and mapping for disaster warning and risk assessment [11-13].

With the development of remote sensing technology, visual interpretation, pixel-based and object-based methods based on optical remote sensing images such as Landsat, RapidEye, Sentinel-2A, etc. have been widely used in landslide interpretation [14, 15]. The visual interpretation method generally produces more accurate results while avoiding hazardous field investigations, but the method relies on the knowledge of the interpreter [16]. Pixel-based and object-based methods are often combined with support vector machine, random forest, and other machine learning methods to achieve landslide detection [17]. The pixel-based method mainly utilizes the pixel information of remote sensing images for landslide detection, but this method only considers the characteristics of a single pixel, neglects the correlation between adjacent pixels, and is sensitive to noise [18, 19]. The object-based approach can widely utilize the shape, statistics, texture features, and contextual information of landslides for further analysis. However, the method is difficult to determine a reasonable segmentation scale for different regions, which affects the effectiveness of landslide detection [20, 21]. In recent years, the development of deep learning has provided new ideas for landslide detection and is gradually being applied to landslide detection with better results than traditional machine learning methods. Among them, the methods based on convolutional neural networks (CNN) and U-Net combined with optical remote sensing images are more frequently used in

This work was supported in part by the National Natural Scientific Foundation of China under Grant 42201459, in part by Gansu Science and Technology Program under Grant 23JRRA881, in part by Young Doctoral Fund Project of Higher Education Institutions in Gansu Province under Grant 2022QB-058. *(Corresponding author: Yi He)*

Hesheng Chen, Yi He, Lifeng Zhang, Wang Yang, Yaoxiang Liu, Binghai Gao, Qing Zhang, and Jiangang Lu are with the Faculty of Geomatics, Lanzhou Jiaotong University, Lanzhou 730700, China, also with the National-Local Joint Engineering Research Center of Technologies and Applications for National Geographic State Monitoring, Lanzhou 730700, China, and also with Gansu Provincial Engineering Laboratory for National Geographic State Monitoring, Lanzhou 730700, China (e-mail: 1625196619@qq.com; heyi@mail.lzjtu.cn).

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully e
content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2023.3339294

landslide detection tasks [22-27].

However, the occurrence of landslides is often accompanied by adverse weather conditions and complex geographical environments, such as heavy rainfall and obstruction caused by cloud or vegetation cover. Therefore, landslide detection based solely on optical remote sensing images combined with deep learning methods may have certain limitations. Synthetic aperture radar (SAR) can image normally under bad weather conditions, and at the same time has a certain penetration ability to penetrate clouds, vegetation, and other information, so it is able to observe the ground all day long and all-weather [28, 29]. In addition, the formation of landslides is closely related to topographic features [30], and SAR images are highly sensitive to changes in surface morphology and structure, so SAR images can be a better data source for landslide detection [31]. Lorenzo Nava et al. used CNN to conduct landslide detection in Iburi Prefecture, Hokkaido, Japan, using SAR data [32]. Sansar Raj Meena et al. conducted a study on automated landslide detection in the Himalayan region using RapidEye optical data and ALOS-PALSAR, derived topographic data (DEM, slope) and analyzed the potential of U-Net and machine learning methods for automated landslide detection in this region [33].

Although the landslide detection methods based on U-Net and its variant structures and various types of remote sensing data have achieved good results, there is still little attention paid to the deep learning model for landslide detection using SAR data fusion optical and terrain factors and other multi-source remote sensing data input. In addition, multi-source remote sensing data contains abundant information while also introducing redundant data. When extracting landslides, U-Net directly fuses the shallow and deep feature information extracted through the skip connection structure. The semantic differences between the two are significant, resulting in feature redundancy and a semantic gap that can interfere with the landslide features learned by the model [34]. Therefore, how to improve the deep learning model (U-Net), effectively input and process multi-source remote sensing data, strengthen the learning of landslide features, effectively distinguish landslides from similar ground objects, and improve the accuracy of landslide detection in complex scenes is a challenging problem faced by using multi-source remote sensing data for landslide detection.

This study constructs a multi-input channel U-Net landslide detection method based on multi-source remote sensing data such as SAR, optics, and terrain factors. The goal of this article is to: (1) to construct a multi-input channel U-Net landslide detection framework that integrates SAR multi-source remote sensing data; (2) to verify the proposed model in Bailong River basin experimental scene; (3) to compare and analyze the results and performance of landslide detection with traditional U-Net models and different types of remote sensing data.

II. Data Sources and Methods

A. Data Sources

The data used in this study include landslide vector data, SAR (Sentinel-1A) data, optical remote sensing image (Sentinel-2A) data, and terrain (slope, aspect) data. The landslide vector data is processed as labeled data, which provides a definition of the true class of each pixel in the image for subsequent training and testing of the deep learning model. SAR, optical remote sensing image, and terrain data are used as the image data for landslide detection, which provides the visual information required by the model to understand the semantic content in the image and classify the image at the pixel level.

1) Landslide vector data: Based on the historical landslide data and combined with fieldwork verification, landslides in the study area were vectorized and labeled by visual interpretation on Google Earth to obtain landslide boundary information, and a total of 732 landslide vectors were obtained. Then the vectorized and labeled landslide data were imported into ArcMap 10.7, converted into raster data, and exported as binary image, which was used as label data for deep learning model training. The labeled image contained two types of pixels, white for landslide areas and black for non-landslide areas.

2) SAR remote sensing data: SAR data are adopted from Sentinel-1A data obtained from the Alaska Satellite Facility (ASF, https://asf.alaska.edu/) website. The Sentinel-1A satellite is equipped with SAR sensors, which can observe the Earth's surface in an all-weather and all-day manner. Compared to optical sensors, SAR has the advantage of providing high-quality observation data even under night, cloud cover, and low visibility conditions. When downloading SAR data, the data obtained was in L1 Single Look Complex (SLC) mode on May 29, 2021, using “Vertical Transmit and Horizontal Receive” (VH) and “Vertical Transmit and Vertical Receive” (VV) polarization methods and Interferometric Wide (IW) acquisition mode. SAR intensity images are heavily affected by noise, so this study conducts "filtering" processing on SAR data to minimize the impact of noise as much as possible. SAR images are affected by geometric distortions, and this study used terrain correction for compression and other distortions, while shadows are difficult to correct. However, in this study, we mainly used multi-source data (optical, SAR, and terrain) for landslide identification, with the aim of utilizing the rich and complementary information of these data to comprehensively present diverse landslide features and improve the performance of landslide detection. Firstly, the Sentinel-1A data and digital elevation model (DEM) data were imported into SARscape software, the VH and VV polarization modes were selected, and then the operation of multilooking, filtering, geocoding, and radiometric calibration was carried out sequentially (specific parameters are shown in Table 3, Table 4, Table 5). Finally, the pre-processed data were imported into ArcMap10.7 software, and the results were exported in PNG format after resampling to obtain the Sentinel-1A image data of VH and VV polarization modes in the experimental scene (Fig. 1(a), (b)).

3) Optical remote sensing data: Sentinel-2A data from European Space Agency (ESA, https://scihub.copernicus.eu/) were used as optical remote sensing data. The cloud-free Sentinel-2A data on April 29, 2021, were selected according to the setup. Sentinel-2A is a high-resolution multi-spectral

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully e
content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2023.3339294

imaging satellite, which includes visible light band, near-infrared band, and mid-infrared band, and provides image data with three resolutions of 10 m, 20 m, and 60 m. The Sentinel-2A data obtained from ESA at the L1C level were not atmospherically corrected, and to eliminate atmospheric effects, they were imported into the Sen2Cor plug-in in the SNAP software developed by ESA for atmospheric correction to obtain L2A data. Subsequently, the SNAP software was used to resample the L2A-level data at 10 m resolution to obtain Sentinel-2A data in 12 bands. Finally, the resampled images of each band were imported into ArcMap 10.7 software, and the experimental scene was embedded and cropped to obtain the Sentinel-2A images of each band, which were exported in PNG format, and some of the images of the bands are shown in Fig. 1(c) and (d).

4) Terrain data: Terrain data include slope and aspect, which provides important terrain auxiliary data for subsequent landslide detection. Slope and aspect were calculated based on DEM data with ALOS 12.5m DEM Resolution. ALOS 12.5m DEM data was elevation data produced by using Advanced Land Observing Satellite (ALOS) phased array L-band synthetic aperture radar (PALSAR). After generating slope and aspect data, imported them into ArcMap10.7 software for resampling, and exported the results in PNG format. Through the above steps, slope and aspect image data in the experimental scene were obtained (Fig. 1(e), (f)).

Table 1 Multilooking parameters

	Main Parameters
Multilooking	Time Domain
Range Looks	1
Azimuth Looks	1
Grid Size for Suggested Looks	15

Table 2 Filtering parameters

	Main Parameters
Filter Method	Lee
Equivalent Number of Looks	-1
Azimuth Window Size	5
Range Window Size	5
Iteration Number	2
Directionality Number	12

Table 3 Geocoding and radiometric calibration parameters

	Main Parameters
X Grid Size	15
Y Grid Size	15
Radiometric Calibration	True
Scattering Area Method	Local Incidence Angle
Radiometric Normalization	True
Local Incidence Angle	False
Output type	Linear

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2023.3339294

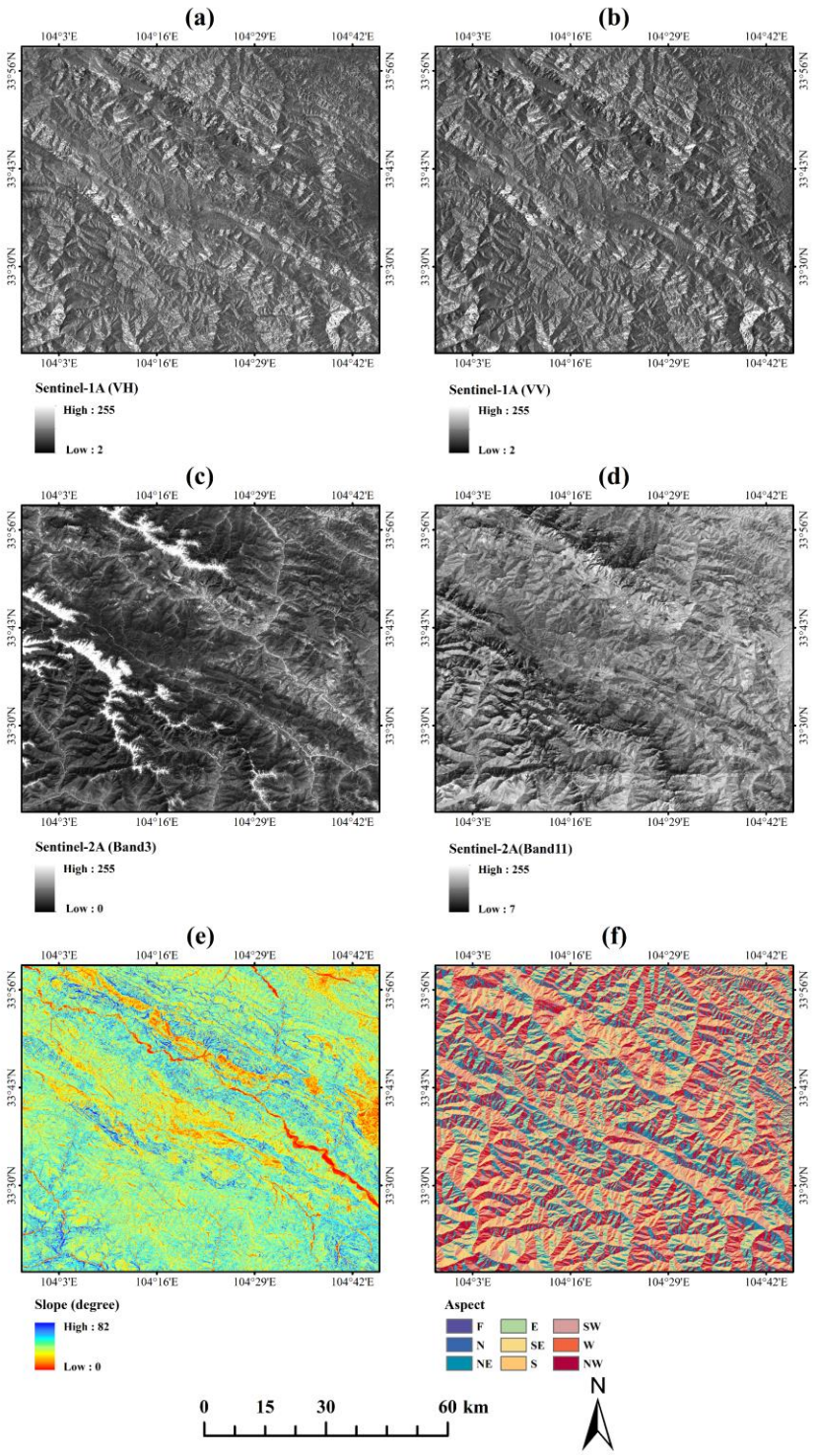


Fig. 1. SAR, optical, and terrain image data. (a) Sentinel-1A (VH), (b) Sentinel-1A (VV), (c) Sentinel-2A(Band3), (d) Sentinel-2A(Band11), (e) Slope, (f) Aspect.

B. Methods

Firstly, the acquired Sentinel-1A, Sentinel-2A, slope, and aspect data were preprocessed and converted into image data suitable for deep learning model training, and the corresponding datasets were created. Then, the multi-input channel U-Net network model was constructed by combining Python language with TensorFlow as the base framework. The experimental scene was selected to train the model using the dataset in the training area and test the model using the testing area. Finally, the landslide detection results of different models combined with different data types were evaluated according to the visual effect and relevant quantitative evaluation indexes, and the model performance was analyzed. The specific flow is shown in Fig. 2, and the following subsections provide detailed information about various aspects of the methodology of this study

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully edited content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2023.3339294

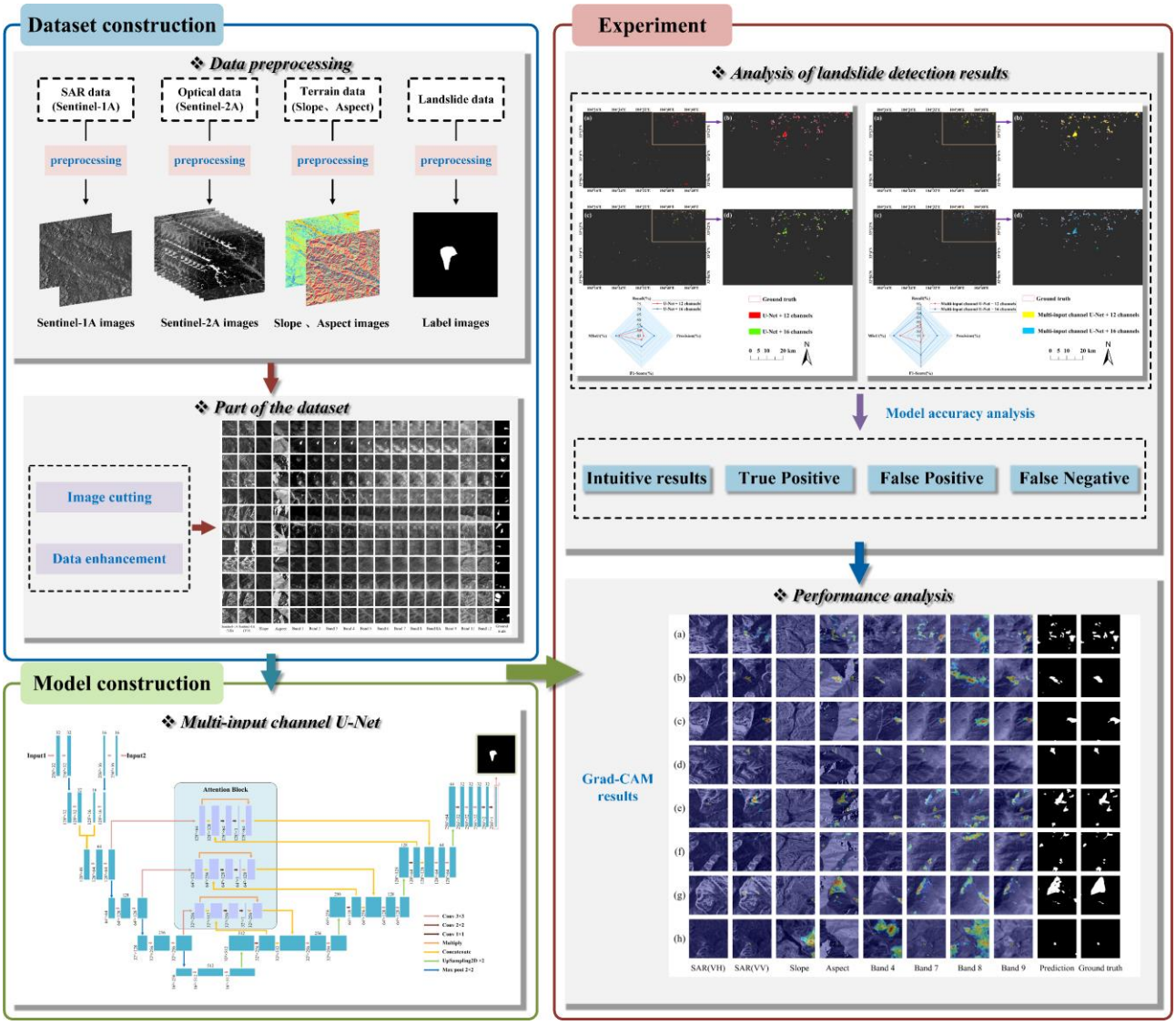


Fig. 2. Flowchart of a multi-input channel U-Net landslide detection method fusing SAR multi-source remote sensing data.

1) Dataset construction: Because the size of original remote sensing image is too large to be directly input into the model for training, images were cut into several sub-images of the same size. At the same time, the size selection of sub-images had a certain impact on the performance of the model. Too small size led to failure to cover large-scale landslide areas, while too large size led to GPU memory overflow [35]. Through the experimental test, considering the landslide scale and hardware constraints, the sub-image size was set to 256*256 pixels in this study. In the process of image cutting, the overlapping cutting strategy with a step size of 128 was adopted. The Sentinel-1A images, Sentinel-2A images, slope images, aspect images, and ground truth images were cut into several sub-images with a size of 256*256 according to the same cutting strategy, and each sub-image contains at least one landslide area, to ensure the effective learning of landslide characteristics by the model.

To increase the number and diversity of samples, improve the generalization ability and robustness of the model, and make the deep learning model better adapt to different data distributions and real scenes, this study used some data enhancement strategies to expand the dataset [36]. Specific data enhancement methods included mirroring and rotating the cut image and ground truth (90 ° counterclockwise rotation). Through these operations, 3153 samples were finally obtained. Then, the obtained samples were randomly divided according to the ratio of the training set to the verification set of 8:2. After the division, the training set contained 2850 samples, and the verification set contained 573 samples. The training set and the verification set were mutually exclusive (Fig. 3). The training set is used to train the parameters and weights of the model, and the validation set is used to evaluate the performance of the model and select the optimal model parameters [37]. Through the data enhancement strategy and data set partition, we can make full use of the sample data and improve the performance and robustness of the model.

In this study, the experimental scene is divided into training

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully e
content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2023.3339294

area and testing area, which are completely independent and mutually exclusive. The data from the training area is used for model training, and the data from the test area is used to test the landslide detection effect of the model and evaluate the performance of the model in landslide detection. It is worth noting that all training samples are generated from the training area, and no image data of the test area is used during model training. This strict distinction can ensure that the model is faced with unprecedented data during the test phase. By dividing the training and testing areas, the generalization ability and adaptability of the model in the real scene can be evaluated. This method verifies the reliability of the model and can provide an objective evaluation of the landslide detection effect.

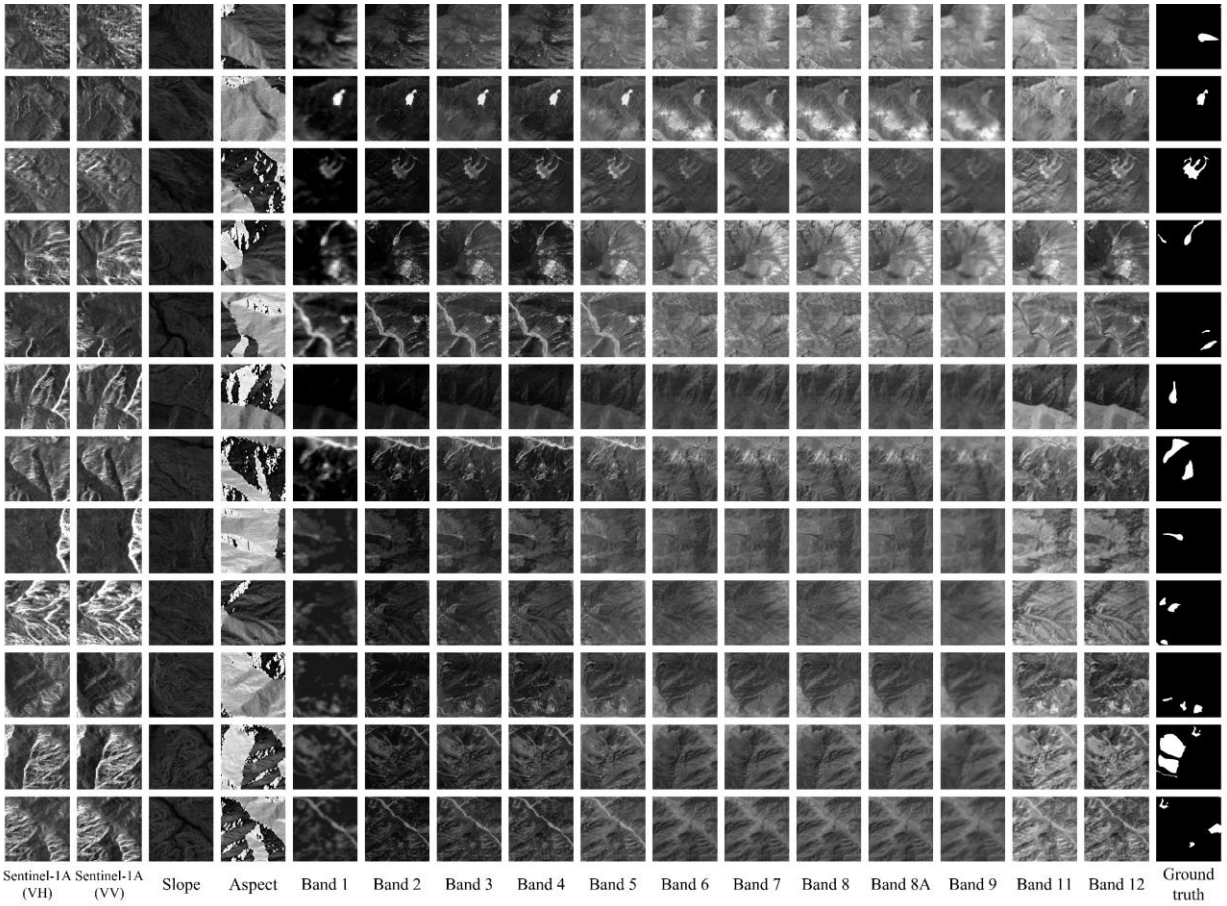


Fig. 3. Partial dataset.

2) U-Net mode: U-Net proposed by Ronneberger et al. in 2015 is a two-dimensional image semantic segmentation network based on fully convolutional networks (FCN) [38, 39]. Due to its robustness, it makes this network widely used in remote sensing image segmentation as well. U-Net is a classic "U" network structure, which is mainly composed of encoder, decoder, and skip connection. The encoder part includes convolution layer and pooling layer. The convolution layer is used to extract the spatial features of the image, and the pooling layer is used to reduce the dimension of the features and increase the receptive field while reducing the size of the feature map; The decoder maps deep abstract features, including upsampling layer and convolution layer. According to the upsampling, it restores a certain spatial size, and uses convolution layer to gradually restore the spatial details; Skip connection is one of the key ideas of U-Net. By copying the feature map of the corresponding depth in the encoder and decoder and concatenating the dimension with the upsampling results in the decoder in the channel dimension, it realizes the fusion of shallow features and deep features, and achieves the purpose of feature fusion in the encoding and decoding process. In this study, U-Net is used as the base network for the landslide detection model, and its structure is shown in Fig. 4.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2023.3339294

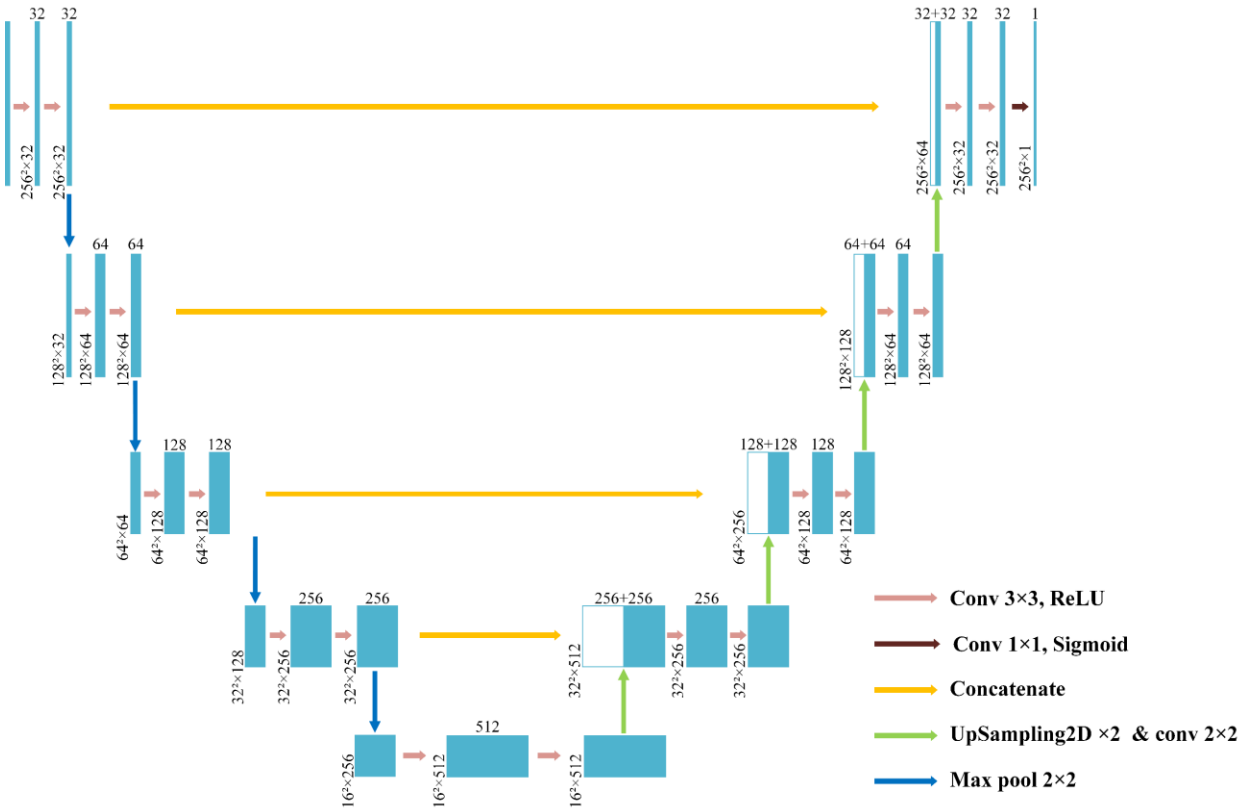


Fig. 4. The structure of U-Net.

3) Multi-input channel U-Net model: In this study, fusing SAR multi-source remote sensing data for landslide detection, and how to input the data effectively into the model is the current problem to be considered. In addition, the non-landslide background occupies most of the area in the image, and the background information other than landslides will also affect the accuracy of landslide detection. Therefore, this study firstly constructs a multiple input channel to abstract the multi-source remote sensing data into multiple input channels so that they can be inputted better. In the field of computer vision, the attention mechanism enables the model to selectively focus on the subject feature information and perform weighted aggregation calculations in the spatial dimension of the feature map, and its basic function is to highlight the subject-related feature part of the feature map, so that the related information will have higher weights in the subsequent calculations [40, 41]. Therefore, for the non-landslide background, during the skip connection between the encoder and the decoder, this study abandons the traditional U-Net direct splicing method and introduces an attention mechanism module to process it in order to adjust the spatial weights of the feature map, focusing on the landslide features [42]. It contains the concatenate operation, two 1×1 convolution operations, and the final Hadamard product operation. The structure of the attention mechanism module is shown in Fig. 5 and the structure of the multi-input channel U-Net model is shown in Fig. 6.

The specific coding form of the multi-input channel U-Net model is to first construct a multi-input channel, divide the multi-source data into two categories (SAR and optical remote sensing images slope and aspect images) to enter the network in the form of multiple inputs, and then after a scale of convolution-pooling feature extraction, the feature maps of the two types of data will undergo the operation of concatenate, so as to make them merge into a single feature map in the channel dimension to facilitate the subsequent step-by-step encoding of landslide features. The specific operation of the attention mechanism module is to present the feature maps after each feature extraction of the U-Net encoder, which are further feature processed by convolution as the input (x) of the attention mechanism module. At the same time, the feature maps of the same size in the U-Net decoder of the corresponding depth (g) and the input feature maps after feature processing are concatenated and merged in the channel dimension to realize the feature fusion between the encoder and the decoder in U-Net; and then for the fused feature maps after two 1×1 convolution operations, Relu and Sigmoid are selected sequentially as activation functions. The Relu function is used to suppress overfitting in the attention mechanism module, and then the Sigmoid function is operated to form a single-channel spatial attention matrix that suppresses redundant features, and then the weight matrix and the feature map inputted in the first step undergo the Hadamard product operation to form a feature map that incorporates the up-sampling results of the decoder and then co-calculates the spatial weights of the attention map to assign higher weights to the landslide area and realize the integration of the landslide area into the decoder and the

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully edited content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2023.3339294

decoder. Higher weights can be assigned to landslide areas to realize the focus on landslide feature areas. Finally, the output map of the attention module is fused with the up-sampling results of the decoder at the corresponding level, and then the detail reconstruction and decoding process of the feature map is continued. This mechanism can not only complete the feature fusion of the corresponding depth of the encoder and decoder in the semantic segmentation process of the U-Net network, but also solve the shortcomings of the fixed spatial weight and the semantic gap between the encoder and decoder feature maps of the traditional U-Net network structure in the corresponding depth feature fusion, so as to improve the performance of the model for landslide feature learning, and then improve the accuracy of landslide detection results.

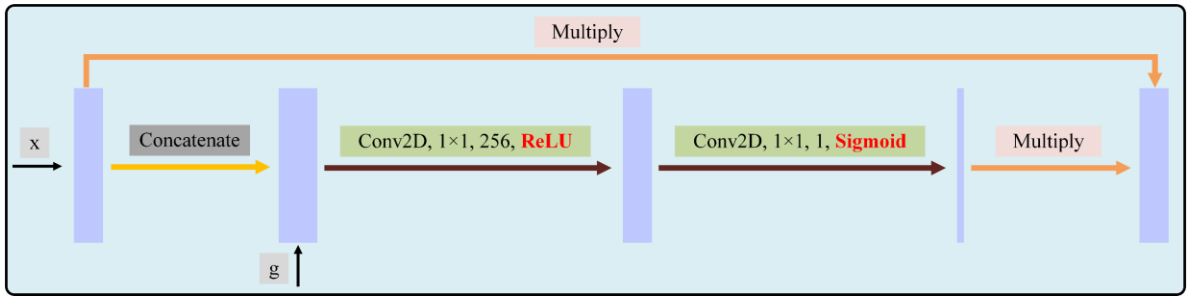


Fig. 5. Attention Block.

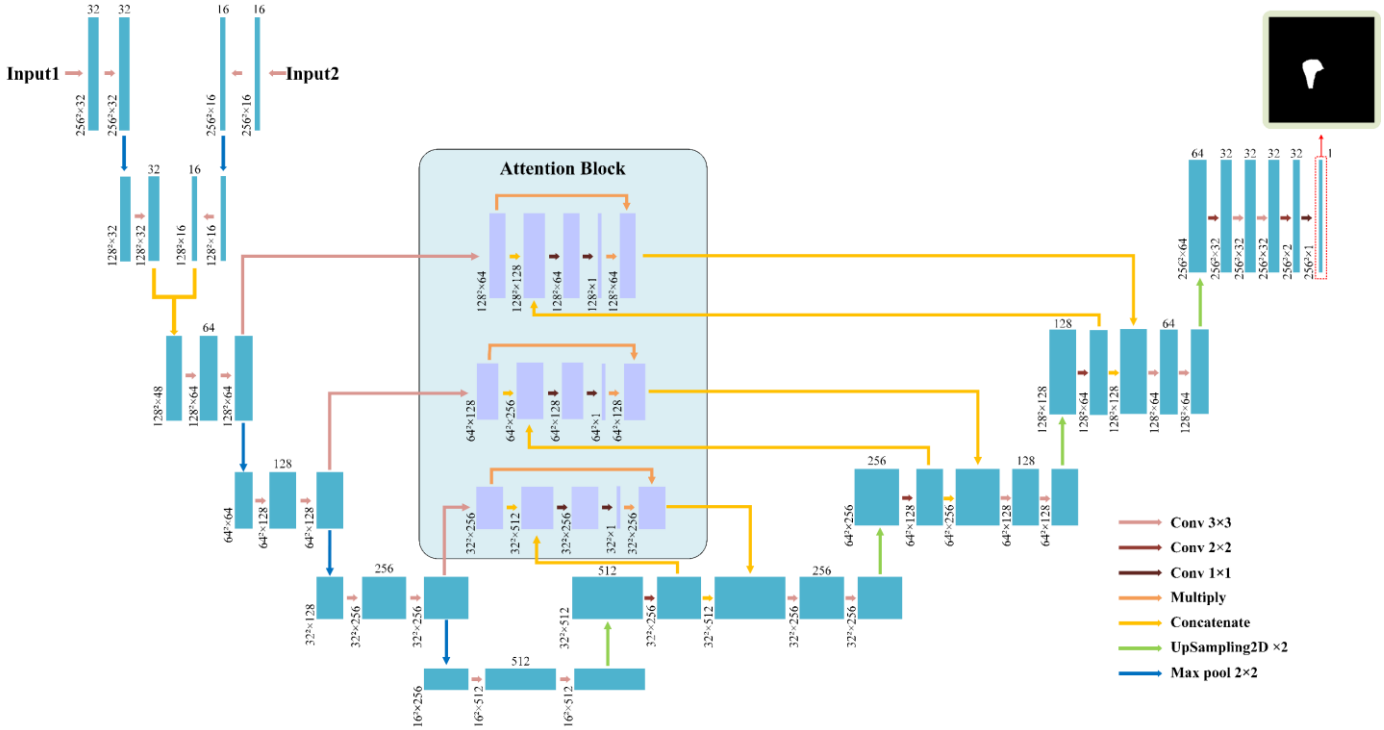


Fig. 6. Multi-input channel U-Net.

4) Model training: Table 4 shows the hardware and operating system configuration used for the experiment and Table 5 lists the details of the software configuration. After the pre-experimental test, considering the hardware limitations and the computational efficiency of the model, the number of iterations (epoch) in the experimental process is set to 128, the batch size is set to 16, the learning rate is set to 0.0001, and the optimizer is selected Adaptive moment estimation (Adam) [43]. The Adam algorithm is an optimization of the gradient descent Adagrad algorithm and the Rmsprop algorithm. At the same time, the Adam algorithm is not affected by the gradient scale transformation when the parameters are updated. It can automatically adjust the learning rate, which is easy to implement and has high computational efficiency [44]. This paper studies the binary classification of image pixels, so binary cross-entropy (BCE) is selected as the loss function. It is defined as follows:

$$Loss = -\frac{1}{N}\sum_{i=1}^{N} y_i \cdot log\left(p\left(y_i\right)\right) + (1-y) \cdot log\left(1-p\left(y_i\right)\right) \quad (1)$$

Where N is the number of categories (N=2 in this paper), y is

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2023.3339294

the binary label 0 or 1, and *p(y)* is the probability that the output belongs to the *y* label.

During the training process, the area under curve (AUC) is selected to monitor the training process of the model. All the deep learning models in this paper are implemented in Python programming language and are developed and implemented based on Tensorflow API and Keras library.

Table 4 The experimental platform

Project	Operating system	CPU	Memory	Hard disk	GPU
Version	Windows 10 Professional	Intel Core i5-13400F	DDR4 32G 3200MHz	KINGSTON SNV2S500G	NVIDIA GeForce RTX 3060

Table 5 The core software

Project	CUDA	CUDNN	Python	Keras	Tensorflow
Version	12.0.133	8.1.1	3.10.10	2.10.0	2.10.1

5) Evaluation index: To quantitatively evaluate the model landslide detection results, this study first calculated the confusion matrix of the landslide detection results (Fig. 7). Subsequently, based on the confusion matrix, this study calculated four quantitative evaluation indicators, namely precision, recall, F1 score, and mean intersection over union (MIoU), to conduct a more comprehensive quantitative evaluation of the model.

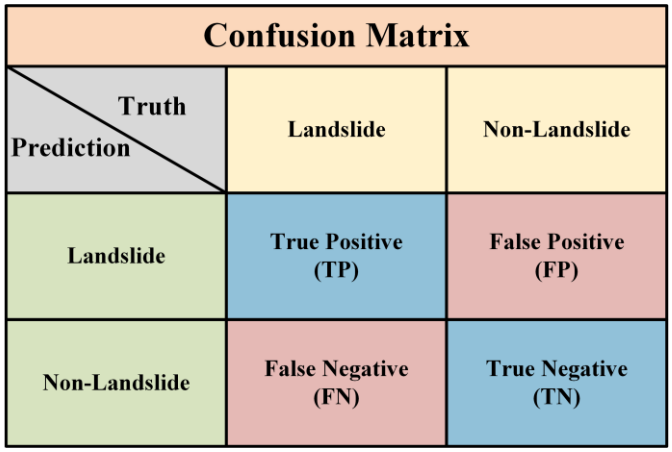


Fig. 7. Confusion matrix of landslide detection results.

Precision is used to evaluate the proportion of samples predicted by the model as positive cases (landslides), and it is an indicator to measure the accuracy of the model in predicting positive cases. The calculation formula is as follows:

$$Precision = \frac{TP}{TP+FP} \quad \ldots\ldots\ldots\ldots\ldots\ldots (2)$$

Recall is used to evaluate the proportion of samples correctly predicted as positive examples (landslides) by the model to all actual positive examples. It is an indicator that measures the model's ability to cover positive examples. The calculation formula is as follows:

$$Recall = \frac{TP}{TP+FN} \quad \ldots\ldots\ldots\ldots\ldots\ldots (3)$$

F1 score (F1) is the harmonic mean of precision and recall, considering both precision and recall indicators. For imbalanced datasets, F1 can provide a more comprehensive performance evaluation, a more balanced indicator, and better reflect the comprehensive performance of the model. The larger the F1, the better the performance of the model. The calculation formula is as follows:

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (4)$$

MIoU is a commonly used evaluation metric for semantic segmentation models, used to measure the performance of the model in semantic segmentation tasks. It provides an evaluation of the overall segmentation accuracy by considering the overlap of pixel levels in various categories. The calculation formula is as follows:

$$MIoU = \frac{\frac{TP}{FP+TP+FN} + \frac{TN}{FP+TN+FN}}{2} \quad (5)$$

III. Results and Analysis

In this section, an experimental scene is selected to apply and comprehensively evaluate the proposed method. Firstly, the proposed model is trained according to the datasets of different data types in the training area, and then the trained model is used to detect landslides in the test area, to compare and analyze the effect of landslide detection of different types of data from the perspective of visual and quantitative evaluation indexes, and to prove the effectiveness of fusing SAR multi-source remote sensing data for landslide detection. In addition, to demonstrate the effectiveness of the multi-input channel U-Net model fusing SAR multi-source remote sensing data for landslide detection, as well as to comparatively analyze its landslide detection effect from different perspectives, we compute the TP, FP, and FN diagrams of the landslide detection results of the different network models fusing SAR multi-source remote sensing data. The comparative experiments and visualization analysis can comprehensively

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully edited content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2023.3339294

evaluate the landslide detection performance of the proposed method in the experimental scene.

A. Experimental scene

Bailong River Basin is located in the southeast of Gansu Province, China. It is a tributary of the Jialing River, a tributary of the Yangtze River. It is in the transition zone from the second-level ladder to the third-level ladder in mainland China. It belongs to the transition area from the Qinghai-Tibet Plateau to the Qinba Mountains and the Loess Plateau (Fig. 8). The climate conditions in the basin are well, the annual average temperature is 2-15 °C, and the annual average precipitation is about 700 mm. The topography and geomorphology in the basin are complex, the alpine gorge landform and the loess landform are distributed alternately, the neotectonic movement is strong, the active faults are developed, and the weak rock and soil bodies (phyllite, etc.) are widely distributed, forming a fragile and sensitive geological environment. At the same time, frequent seismic activities and deep valleys provide favorable conditions for the occurrence of landslide disasters, making landslide disasters frequent [45, 46]. Affected by complex factors such as complex terrain, vegetation cover, and weather conditions in the basin, it is difficult to detect landslides only using optical remote sensing data. Therefore, it is necessary to use other types of remote sensing data in combination to make up for the limitations of optical remote sensing data.

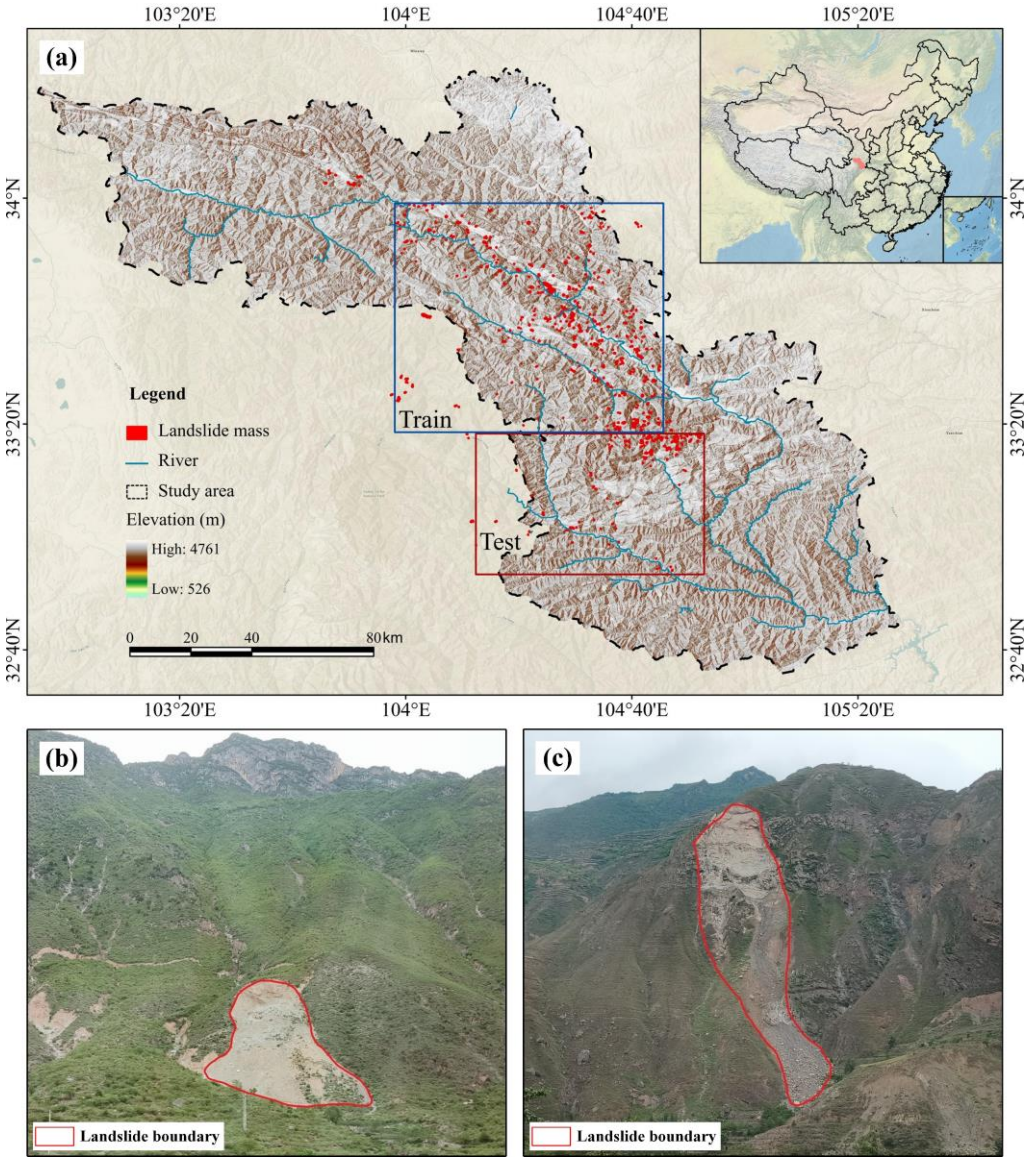


Fig. 8. Experimental scene. (a) Location, (b) and (c) Field investigation images of experimental scene.

B. Analysis of landslide detection results fusing SAR multi-source remote sensing data

To further validate the usability and effectiveness of the proposed method fusing SAR multi-source data for landslide detection, a total of four sets of experiments were designed for two-by-two comparisons in this study. (1) traditional U-Net

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully e
content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2023.3339294

combined with 12 bands of Sentinel-2A images (U-Net + 12 channels); (2) traditional U-Net combined with SAR multi-source remote sensing data (two polarization modes of Sentinel-1A images, 12 bands of Sentinel-2A images, slope, and aspect images, U-Net + 16 channels); (3) multi-input channel U-Net combined with 12 bands of Sentinel-2A images (multi-input channel U-Net + 12 channels); (4) multi-input channel U-Net combined with SAR multi-source remote sensing data (two polarization modes of Sentinel-1A images, 12 bands of Sentinel-2A images, slope, and aspect images, U-Net + 16 channels). In this study, the model was trained based on the training dataset produced in the same training area, and the model was tested in the same testing area. To obtain an overview of the landslide detection results of the different methods, all the sub-images of the detection results were combined to form a complete landslide image of the testing area based on their geographic coordinates. Meanwhile, we calculated the four evaluation indexes, recall, precision, F1, and MIoU (Table 6), respectively, based on the confusion matrix to further quantitatively evaluate the landslide detection results (Fig. 9 and 10).

Comparing Fig. 9(a), (b) and (c), (d), and Fig. 10(a), (b) and (c), (d), it can be found that whether based on the traditional U-Net model or the multi-input-channel U-Net model, the landslide detection results with the fusion of the SAR multi-source remote sensing data are better than the landslide detection results using only the Sentinel-2A optical remote sensing images in terms of overall visual effect, and the overall results of the detection are more refined, with fewer false detections and missed detections. Meanwhile, the landslide detection results with the fusion of SAR multi-source remote sensing data are also superior to the landslide detection results using only Sentinel-2A optical remote sensing images for all four quantitative evaluation indexes. Among them, high recall means that the coverage rate of landslide detection is high, that is, more landslides are detected, regardless of how many objects are mistakenly classified as landslides (In Figure 11, the multi-input channel U-Net model combined with SAR multi-source remote sensing data obtained the maximum landslide area, indicating that this method can comprehensively and completely detect landslides). High precision indicates that the object detected as a landslide has a high probability of being a landslide. F1 comprehensively considers the model's detection coverage of landslides and the accuracy of being detected as landslides and provides an overall performance based on recall and precision. MIoU can be used to measure the similarity between the model's prediction results of landslides and non-landslides at the pixel level and the real labels. In this scene, the number of background pixels representing non-landslides is much higher than the number of images representing landslides, and the distribution between sample categories is unbalanced. Therefore, MIoU is higher than the three metrics of recall, precision, and F1 in the four sets of experiments. On both traditional U-Net and multi-input channel U-Net models, F1 values of landslide detection results incorporating SAR multi-source remote sensing data are higher by about 6.6% and 14.2%, respectively, which comprehensively proves that the incorporation of SAR multi-source remote sensing data can deal with the phenomena of false detections and missed detections better in the face of the complex scene.

The comparison between different models shows that, due to the complex geographic environment of Bailong River (complex topography, vegetation cover, and cloud cover, etc.), it is difficult to provide enough landslide feature information for the model to learn based on only optical remote sensing images. The landslide detection method fusing SAR multi-source remote sensing data can effectively introduce and utilize the rich and comprehensive landslide feature information between multi-source data, which facilitates the model to understand and learn the landslide features from different perspectives, effectively differentiating the landslide from its surrounding similar and complex features, and then improve the landslide detection accuracy. The visualization and quantitative evaluation results demonstrate the usability and effectiveness of fusing SAR multi-source remote sensing data for landslide detection.

Table 6 Results of evaluation index

Models	Recall (%)	Precision (%)	F1-score (%)	MIoU (%)
U-Net + 12 channels	50.56	46.95	48.69	66.00
U-Net + 16 channels	53.77	56.88	55.28	69.03
Multi-input channel U-Net + 12 channels	57.45	47.49	52.00	67.48
Multi-input channel U-Net + 16 channels	69.74	62.96	66.18	74.67

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully edited content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2023.3339294

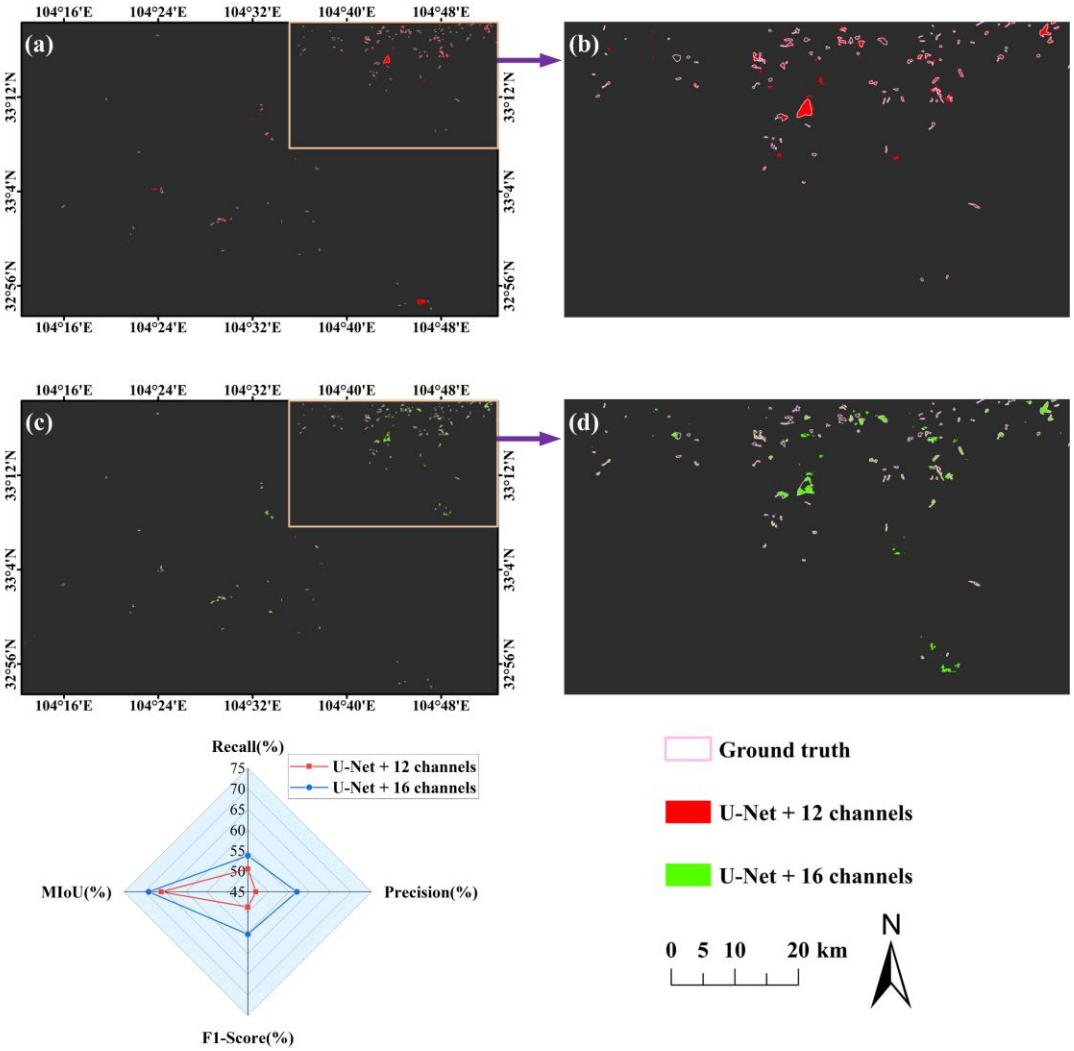


Fig. 9. Landslide detection results of traditional U-Net fusion of different data in the test area. (a) landslide detection results from traditional U-Net fused optical remote sensing imagery, (b) specific area in a, (c) landslide detection results from traditional U-Net fused SAR multi-source remote sensing data, (d) specific area in c.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully edited content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2023.3339294

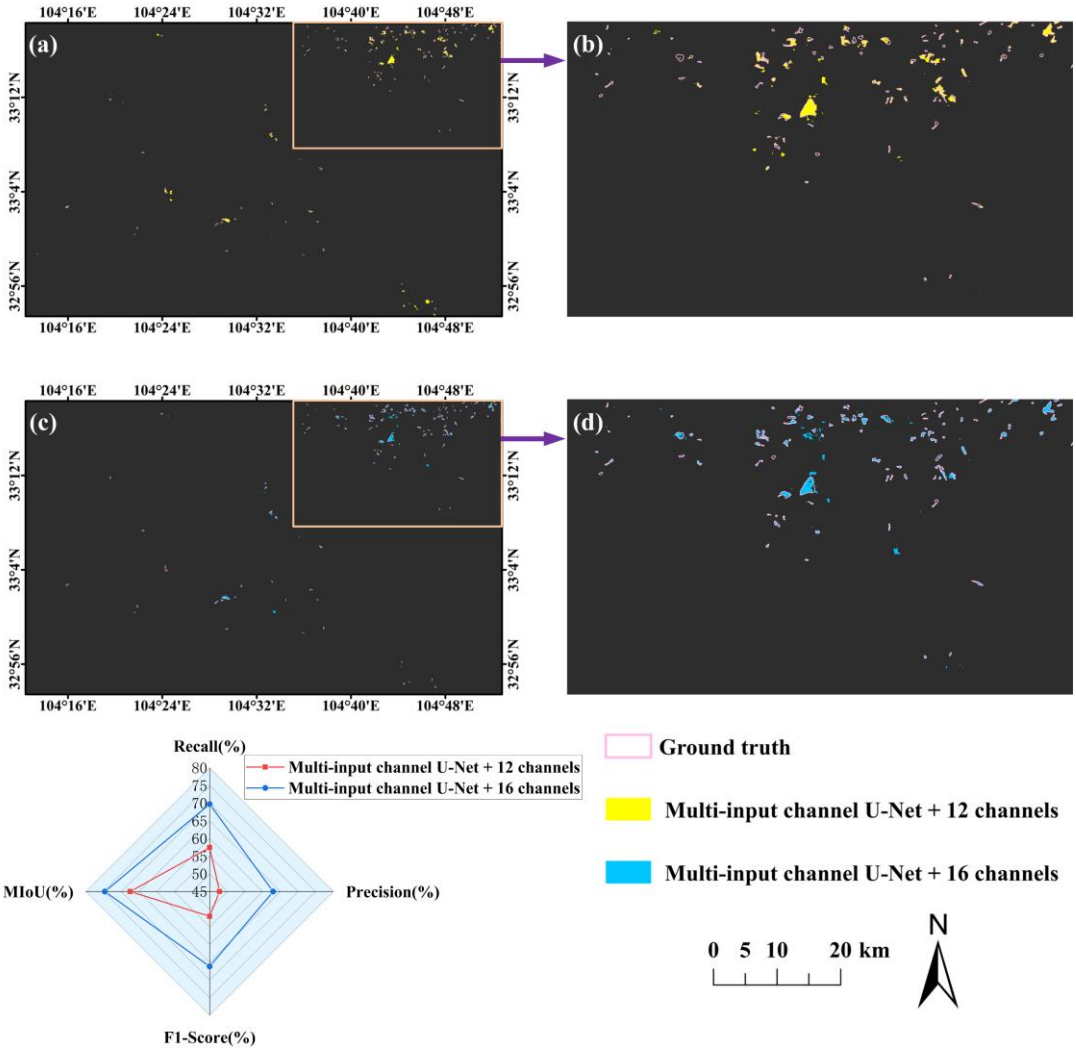


Fig. 10. Landslide detection results of multi-input U-Net fusion of different data in the test area. (a) landslide detection results from multi-input channel U-Net fused optical remote sensing imagery, (b) specific area in a, (c) landslide detection results from multi-input channel U-Net fused SAR multi-source remote sensing data, (d) specific area in c.

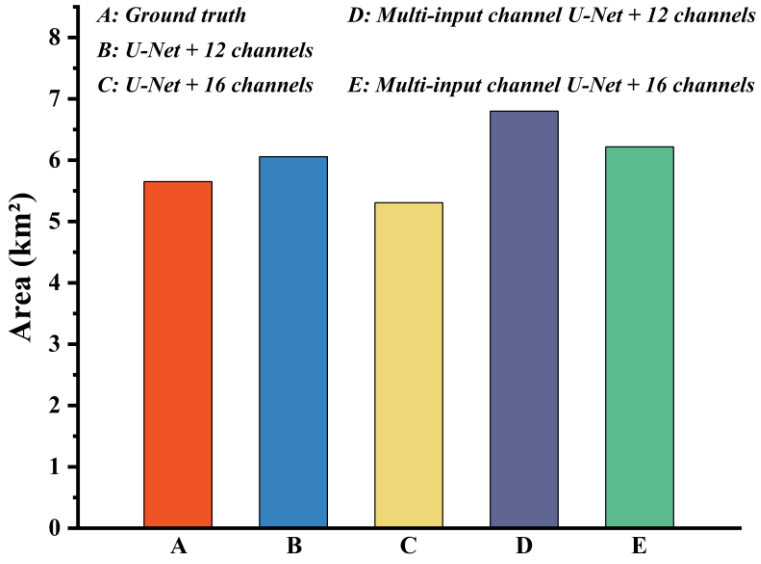


Fig. 11. Total area of detected landslides

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully edited content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2023.3339294

C. Accuracy analysis of the proposed model

In this study, multi-input channel U-Net is constructed as a deep learning model for landslide detection and combined with multi-source remote sensing data such as SAR (Sentinel-1A), optical (Sentinel-2A), and terrain (slope and aspect) for model training and landslide detection. To validate the effectiveness of the landslide detection model in this study and to demonstrate the superiority of its performance, we compare the multi-input channel U-Net landslide detection results with the traditional U-Net landslide detection results for the test area. The results show that both models can detect most of the landslides in the test area. Some areas are randomly selected and enlarged (Fig. 12).

Comparing the landslide detection results of ground truth with that of multi-input channel U-Net and traditional U-Net, it can be found that the traditional U-Net model can detect relatively complete landslide boundary information, but there are many false detections (yellow circles in Fig. 12(c)) and missed detections (red circles in Fig. 12(c)). The false detections mostly occur near known landslide areas, and the scale of the objects that were mistakenly detected as landslides is relatively small, some landslides that are close to each other have misconnected at the boundary. At the same time, the phenomenon of missed detection is more significant on small-scale landslides, and the detection results are not precise enough, indicating that traditional U-Net models have certain shortcomings in landslide feature extraction and learning. Compared with traditional U-Net models, the overall detection performance of the multi-input channel U-Net model is better (Fig. 12(d)), which is closer to the ground truth. The landslide detection results are more complete, and there are fewer false detections and missed detections. This effectively improves the recognition ability of landslide and non-landslide feature differences, reflecting the comprehensive ability of the multi-input channel U-Net model in the landslide detection task of this study. Multi-input channel U-Net model can better learn landslide features and achieve more accurate landslide detection.

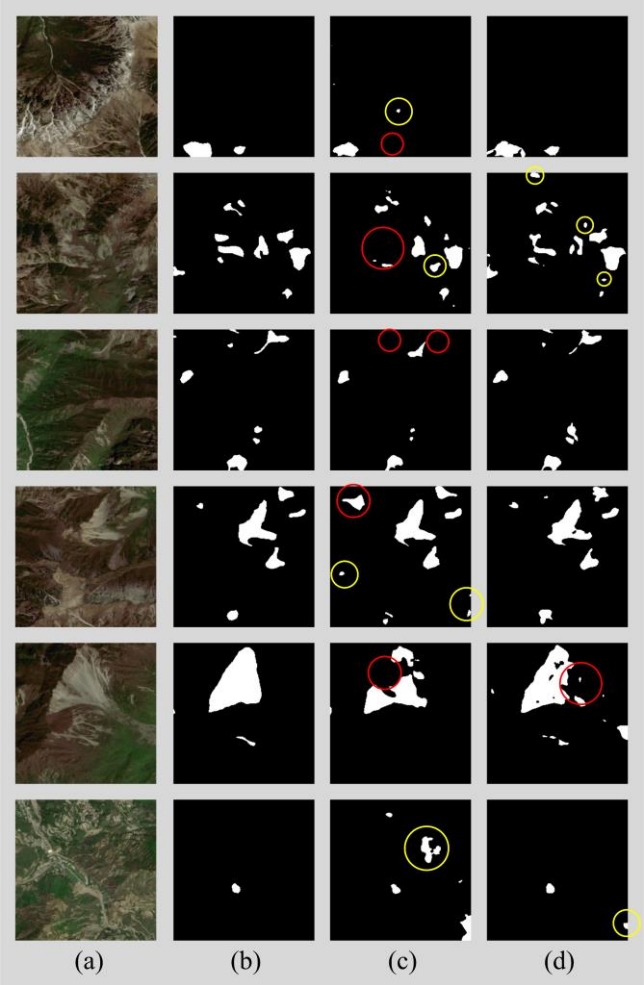


Fig. 12. Partial landslide detection results of different models fusing SAR multi-source remote sensing data in the test area. (a) optical (Sentinel-2A) image, (b) ground truth, (c) U-Net results, and (d) multi-input channel U-Net results.

To further analyze the accuracy of the landslide detection results of the proposed model, we calculate the TP, FP, and FN of the landslide detection results of the traditional U-Net and multi-input channel U-Net models based on ground truth and landslide detection results, and visualize them (Fig. 13). The accuracy of the detection results is compared from the overall large-scale area. As shown in Fig. 13, compared to traditional U-Net (Fig. 13(a), (b)), multi-input channel U-Net has more TP pixels (Fig. 13(c), (d) blue area, the model predicted landslide, actual landslide), fewer FP pixels (Fig. 13(c), (d) red area, the model predicted landslide, actual landslide), and fewer FN pixels (Fig. 13(c), (d) yellow area, the model predicted non-landslide, actual landslide), this phenomenon indicates that the multi-input channel U-Net model can better balance precision and recall, indicating that the multi-input channel U-Net model has achieved more accurate detection of landslides, with fewer false detections and missed detections. This proves the effectiveness of the multi-input channel U-Net model in the research method of this article. In addition, we also randomly select some sub-regions to plot TP, FP, and FN results and

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully edited content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2023.3339294

enlarge them (Fig. 14) for visual comparison of accuracy at different scales. The results also demonstrate the superiority of the multi-input channel U-Net model for landslide detection (compared to traditional U-Net, the multi-input channel U-Net landslide detection results have more TP, less FP, and FN).

Overall, this study constructs a multi-input channel U-Net model can effectively introduce SAR multi-source remote sensing data by adding multiple input channels. The attention mechanism is introduced between the U-Net jump connections to adjust the spatial weights of the feature maps, so that the model can better deal with the complex background information and focus on and learn the landslide features. While completing the U-Net network feature fusion, the deficiency of U-Net in corresponding to the depth feature fusion with fixed spatial weights and the semantic gap between the encoder and decoder feature maps is solved, and the landslide and non-landslide features can be better distinguished. The multi-input channel U-Net model combined with SAR, optical, and terrain factor data achieves better landslide detection.

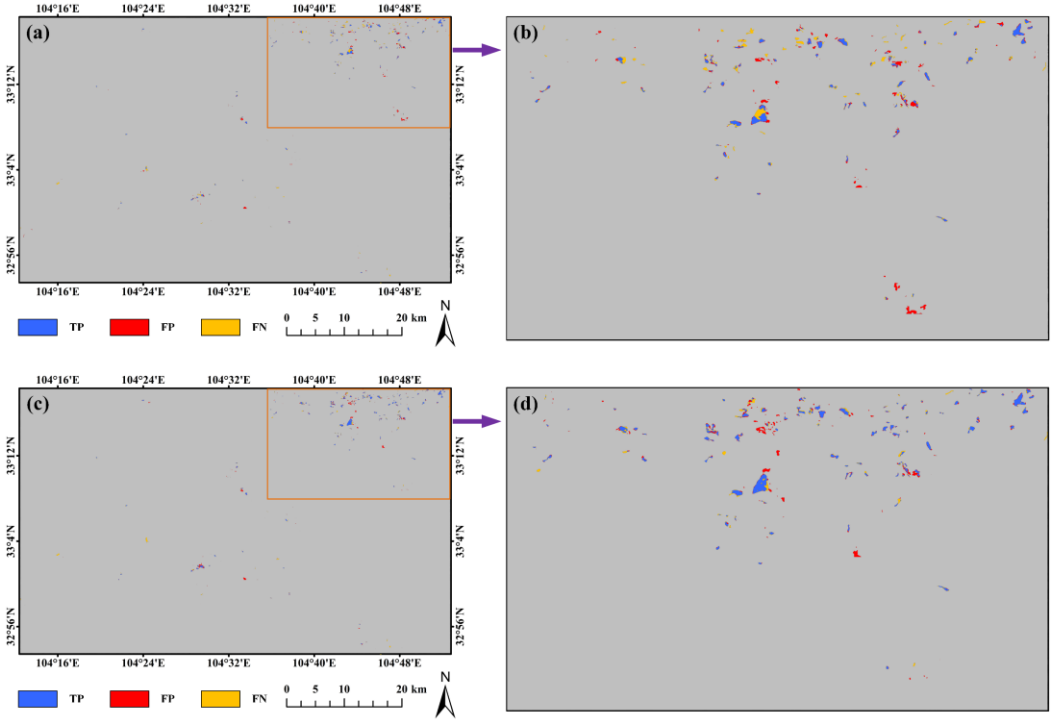


Fig. 13. Visualization of TP, FP, and FN for landslide detection results. (a) U-Net results, (b) specific area in a, (c) multi-input channel U-Net results, (d) specific area in c.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2023.3339294

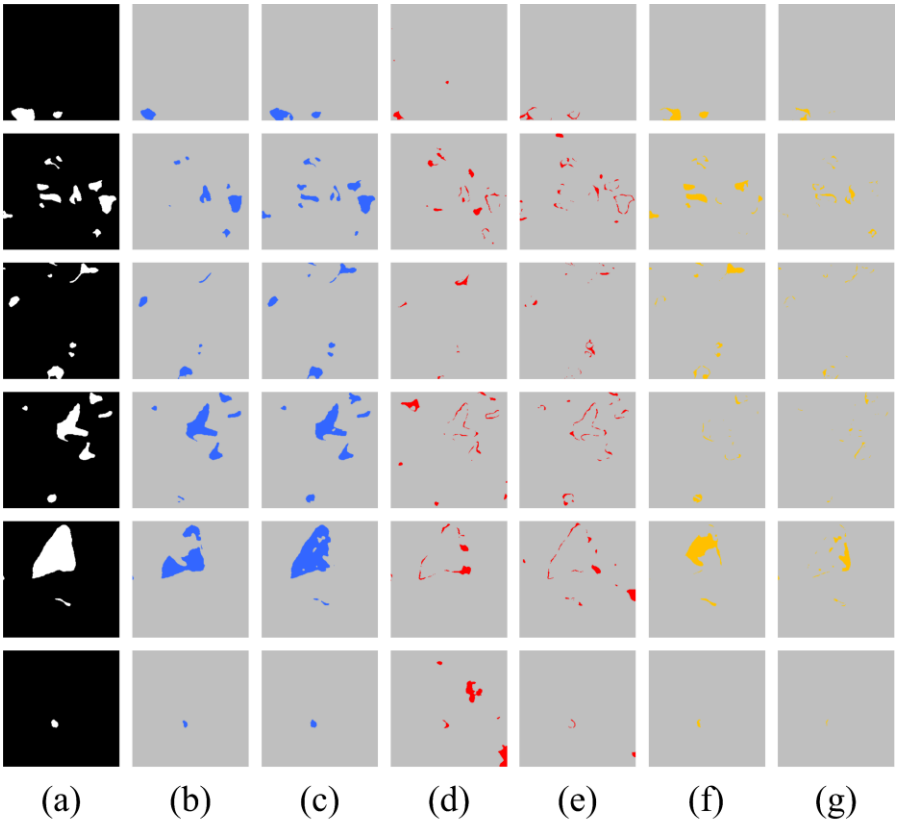


Fig. 14. TP, FP, and FN diagrams of landslide detection results for randomly selected regions. (a) ground truth, (b) U-Net TP, (c) multi-input channel U-Net TP, (d) U-Net FP, (e) multi-input channel U-Net FP, (f) U-Net FN, and (g) multi-input channel U-Net FN.

IV. Discussion

A. Model performance analysis

Gradient-weighted Class Activation Mapping (Grad-CAM) is a deep network visualization method based on gradient localization, which helps explain the decision basis of the model and identify important features by generating heat maps to visualize the regions of interest of the model. Grad-CAM can help users better understand and validate the behavior and performance of the models [47, 48]. Before training the multi-input channel U-Net model, a total of 16 inputs were constructed by abstracting all types of data into 16 bands of images, and all types of data were jointly input into the model for training. To enhance the interpretability of the proposed method, the heat maps of each input band image were generated separately using the Grad-CAM algorithm. Then superimposed it on the corresponding band image to display, visualized the region of interest of the multi-input channel U-Net model on each band image, and explained the decision basis of the model on each band image. Fig. 15 shows the corresponding results.

In Fig. 15, the red region is the part that the model pays more attention to. The darker color indicates that the model pays more attention to this region, which is considered to contribute more to the landslide pixel classification results. This indicates that the features in this region play an important role in the landslide detection task. On the contrary, the blue region is the part that the model pays less attention to. The lighter color indicates that the model pays less attention to this region and considers that the region contributes less to the landslide pixel classification results. This indicates that the features in this region play a minor role in the landslide detection task [49]. Based on the observation of the heat map, it can be found that each band of the image makes a certain contribution in landslide detection.

This study fuses SAR multi-source remote sensing data for landslide detection. Multi-source remote sensing data contain different types of rich and diverse information due to the differences in data sources, physical characteristics, resolution, and sensor parameters. In particular, SAR can penetrate vegetation and clouds to obtain subsurface information, which is valuable for landslide identification because some landslides will be covered under vegetation and clouds [50-52]. Meanwhile, SAR has a strong ability to perceive changes in the surface and surface structure, and landslides usually lead to changes in surface morphology and destruction of surface structure [53]. These phenomena can be captured in SAR images, thus contributing to the identification and monitoring of landslides. This study used SAR (Sentinel-1A) data with two polarization methods, VH and VV. VH polarization emits electromagnetic waves in the vertical direction and receives electromagnetic waves in the horizontal direction. VV polarization is the transmission and reception of electromagnetic waves in a vertical direction [54]. The echo

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully e
content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2023.3339294

intensity of the same object varies under different polarization modes, thus increasing the information of landslide targets. The common use of SAR data with different polarization methods can provide rich polarization information from multiple perspectives. This polarization information can be used to extract more details of surface features, enhance landslide recognition ability, reduce false and missed detection rates, and obtain better visual effects and evaluation indicators (Fig. 9, 10). From Fig. 15, the model's attention to landslide features is relatively rough in some bands of optical images, while it is more detailed in SAR images (Fig. 15(a), (c), (h)). In Fig. 15(h), the model generates more erroneous attention (Band4, Band8) in some bands of optical images, while the attention on SAR images is more accurate and clearer. Slope and aspect data provide terrain information, which can supplement missing or difficult-to-observe terrain features in optical images. Landslides are usually related to the steepness and aspect of the terrain. Increasing slope and aspect data can provide a more comprehensive and clear understanding of terrain changes, thereby better helping deep learning models capture the terrain features of landslide areas (Fig. 15(e), (g), (h)). Therefore, fusing SAR multi-source remote sensing data can provide rich and complementary information to help the model perceive landslides from different perspectives, analyze the surface situation more comprehensively, and capture more comprehensive and diversified landslide features, thus improving the performance of landslide detection.

In this study, we construct a multi-input channel U-Net model for landslide detection by introducing the multi-input channel and attention mechanism module, which realizes the effective input and processing for multi-source data. The model has strong information processing and feature learning ability, and can learn the information related to landslide features in different types of data by means of adaptive weighting. The introduced attention mechanism module (Fig. 5) can solve the problem of spatial weight fixation of traditional U-Net when corresponding to deep feature fusion. It can effectively capture the important landslide-related features in the input multi-source data, automatically learn and focus on the key areas (darker color area, red area in Fig. 15), and enhance the learning ability of landslide features. Ultimately, it improves the ability to accurately identify and localize landslide areas [42, 55]. By fusing information from different data sources, the multi-input channel U-Net can obtain landslide feature information from multiple viewpoints and data sources. This information is then processed effectively to improve the accuracy and robustness of landslide detection, thus enhancing the performance of landslide detection.

In addition, from the above subsection 3.2, the landslide detection accuracy can be improved to a certain extent by optimizing either the network structure or the dataset. It is worth noting that the landslide detection results of traditional U-Net combined with SAR multi-source remote sensing data are better than those of multi-input channel U-Net combined with optical images in all four evaluation indexes. The multi-input channel U-Net realizes the effective input of multi-source remote sensing data and focuses on landslide features by introducing multiple input channels and adding the attention mechanism. The SAR multi-source remote sensing data can provide rich information to help the model better understand and learn landslide features. Therefore, it can be inferred that the effect of fusing multi-source data is better than the effect of changing the network structure of the model in this research experiment. This suggests that the proposed method of enriching data information by introducing multi-source data to improve model performance is feasible in landslide detection.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2023.3339294

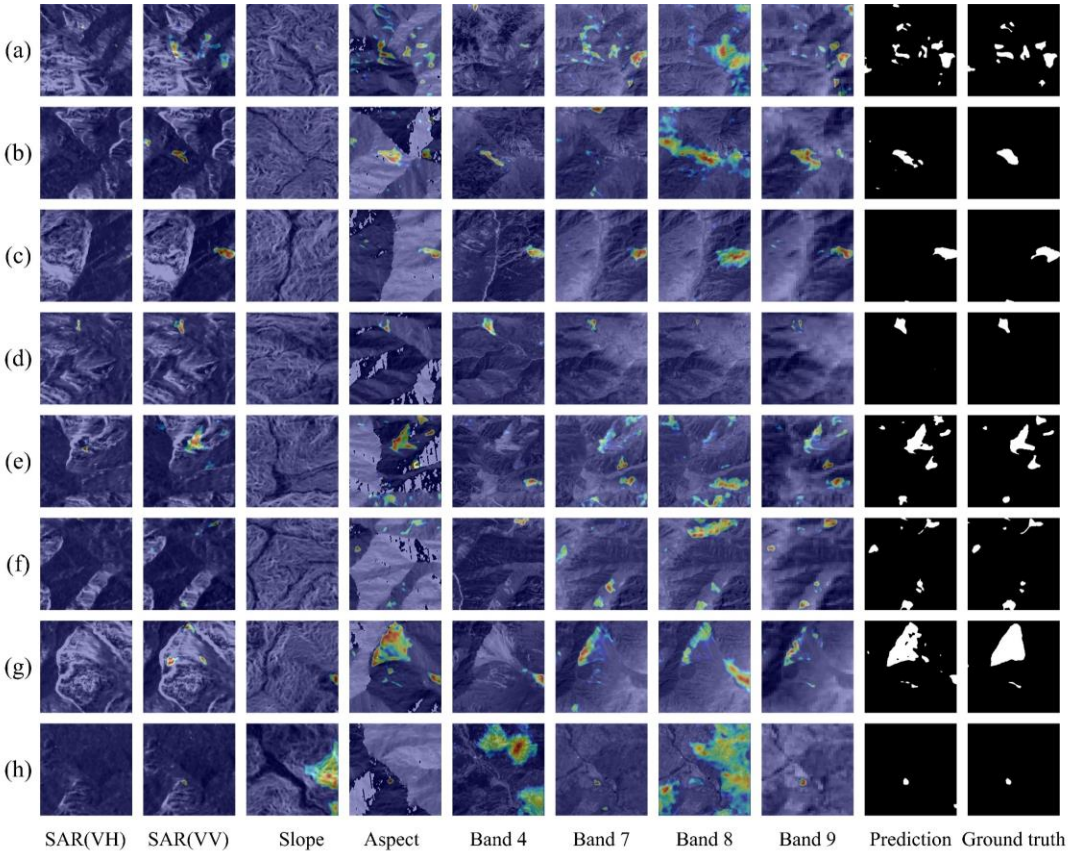


Fig. 15. Heat maps of landslide detection by multi-input channel U-Net combined with SAR multi-source remote sensing data.

B. Prospect

This study utilizes SAR multi-source remote sensing data combined with the multi-input channel U-Net model for landslide detection and applies it to the Bailong River Basin. The multi-input channel U-Net model can effectively input the SAR multi-source remote sensing data and better learn the rich landslide feature information provided by these data, and the detection results proved its validity, but there still existed some errors and omissions (Fig. 9 and 10). This is because landslides are geological phenomena with different spatial scales, and their scale varies [56, 57]. The spatial resolution of the data used in this study is limited, so some smaller landslides may be overlooked, resulting in missed detection. Meanwhile, deep learning methods are data-driven, which requires many training data [58-60], and the number of landslide samples used for training in this study is relatively small. Therefore, we will consider using higher spatial resolution remote sensing data for landslide detection in future work to improve the accuracy of landslide detection. In addition, in future application scenarios, facing the problem of insufficient data samples, how to effectively expand data samples or accurately detect landslides based on current samples is also a problem that needs to be considered in future work.

V. Conclusion

A multi-input channel U-Net landslide detection method fusing SAR multi-source remote sensing data is proposed and utilized for landslide detection. SAR multi-source remote sensing data can provide rich information. The multi-input channel U-Net model can input the data effectively and process the most valuable and noteworthy parts of the multi-source data. Then the landslide-related parts are associated with higher weights to improve the model's ability to perceive and recognize the key features of landslides. This improves the model's ability to learn landslide features. Mutually independent training and testing regions are selected to apply the multi-input channel U-Net method fusing SAR multi-source remote sensing data for landslide detection experiments, and compared with the traditional U-Net with SAR multi-source remote sensing data and the traditional U-Net without regard to SAR multi-source remote sensing data for landslide detection methods. The proposed multi-input channel U-Net landslide detection method fusing SAR multi-source remote sensing data can comprehensively utilize the rich information provided by different types of multi-source remote sensing data, with a focus on learning landslide features. As a result, better performance was achieved, with higher landslide detection accuracy and lower false and missed detection rates. The quantitative results show that each evaluation index of the multi-input channel landslide detection method fusing SAR multi-source remote sensing data proposed in this study is optimal, and the highest F1 value (66.18%) is obtained, which verifies the effectiveness of the proposed method. The proposed method can provide technical support for landslide disaster assessment.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2023.3339294

ACKNOWLEDGMENT

The authors would like to thank the editor and anonymous reviewers for their valuable time on this manuscript.

REFERENCES

[1] K. Dai et al., "Entering the Era of Earth Observation-Based Landslide Warning Systems: A Novel and Exciting Framework," IEEE Geoscience and Remote Sensing Magazine, vol. 8, no. 1, pp. 136-153, 2020.

[2] Y. He et al., "A unified network of information considering superimposed landslide factors sequence and pixel spatial neighbourhood for landslide susceptibility mapping," International Journal of Applied Earth Observation and Geoinformation, vol. 104, p. 102508, 2021/12/15/ 2021.

[3] Q. Xu et al., "Remote sensing for landslide investigations: A progress report from China," Engineering Geology, vol. 321, p. 107156, 2023/08/01/ 2023.

[4] S. Ji, D. Yu, C. Shen, W. Li, and Q. Xu, "Landslide detection from an open satellite imagery and digital elevation model dataset using attention boosted convolutional neural networks," Landslides, vol. 17, no. 6, pp. 1337-1352, 2020/06/01 2020.

[5] B. Gao et al., "Landslide Risk Evaluation in Shenzhen Based on Stacking Ensemble Learning and InSAR," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 16, pp. 1-18, 2023.

[6] P. Lu, S. Bai, V. Tofani, and N. Casagli, "Landslides detection through optimized hot spot analysis on persistent scatterers and distributed scatterers," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 156, pp. 147-159, 2019/10/01/ 2019.

[7] W. Jiang et al., "Deep Learning for Landslide Detection and Segmentation in High-Resolution Optical Images along the Sichuan-Tibet Transportation Corridor," Remote Sensing, vol. 14, no. 21. doi: 10.3390/rs14215490

[8] Y. He et al., "An identification method of potential landslide zones using InSAR data and landslide susceptibility," Geomatics, Natural Hazards and Risk, vol. 14, no. 1, p. 2185120, 2023/12/31 2023.

[9] Y. Li, P. Wang, Q. Feng, X. Ji, D. Jin, and J. Gong, "Landslide detection based on shipborne images and deep learning models: a case study in the Three Gorges Reservoir Area in China," Landslides, vol. 20, no. 3, pp. 547-558, 2023/03/01 2023.

[10] B. Yu, N. Wang, C. Xu, F. Chen, and L. Wang, "A Network for Landslide Detection Using Large-Area Remote Sensing Images with Multiple Spatial Resolutions," Remote Sensing, vol. 14, no. 22. doi: 10.3390/rs14225759

[11] S. O. Y. Amankwah et al., "Landslide detection from bitemporal satellite imagery using attention-based deep neural networks," Landslides, vol. 19, no. 10, pp. 2459-2471, 2022/10/01 2022.

[12] X. Zhang, W. Yu, M.-O. Pun, and W. Shi, "Cross-domain landslide mapping from large-scale remote sensing images using prototype-guided domain-aware progressive representation learning," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 197, pp. 1-17, 2023/03/01/ 2023.

[13] Z. a. Zhao et al., "A comparative study of different neural network models for landslide susceptibility mapping," Advances in Space Research, vol. 70, no. 2, pp. 383-401, 2022/07/15/ 2022.

[14] F. Chen, B. Yu, and B. Li, "A practical trial of landslide detection from single-temporal Landsat8 images using contour-based proposals and random forest: a case study of national Nepal," Landslides, vol. 15, no. 3, pp. 453-464, 2018/03/01 2018.

[15] O. Ghorbanzadeh, K. Gholamnia, and P. Ghamisi, "The application of ResU-net and OBIA for landslide detection from multi-temporal sentinel-2 images," Big Earth Data, pp. 1-26, 2022.

[16] H. Chen et al., "A landslide extraction method of channel attention mechanism U-Net network based on Sentinel-2A remote sensing images," International Journal of Digital Earth, vol. 16, no. 1, pp. 552-577, 2023/12/31 2023.

[17] O. Ghorbanzadeh, Y. Xu, P. Ghamisi, M. Kopp, and D. Kreil, "Landslide4Sense: Reference Benchmark Data and Deep Learning Models for Landslide Detection," IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1-17, 2022.

[18] M. I. Sameen and B. Pradhan, "Landslide Detection Using Residual Networks and the Fusion of Spectral and Topographic Information," IEEE Access, vol. 7, pp. 114363-114373, 2019.

[19] X. Yang and L. Chen, "Using multi-temporal remote sensor imagery to detect earthquake-triggered landslides," International Journal of Applied Earth Observation and Geoinformation, vol. 12, no. 6, pp. 487-495, 2010/12/01/ 2010.

[20] S. W. Myint, P. Gober, A. Brazel, S. Grossman-Clarke, and Q. Weng, "Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery," Remote Sensing of Environment, vol. 115, no. 5, pp. 1145-1161, 2011/05/15/ 2011.

[21] T. R. Martha, N. Kerle, V. Jetten, C. J. van Westen, and K. V. Kumar, "Characterising spectral, spatial and morphometric properties of landslides for semi-automatic detection using object-oriented methods," Geomorphology, vol. 116, no. 1, pp. 24-36, 2010/03/15/ 2010.

[22] N. Prakash, A. Manconi, and S. Loew, "A new strategy to map landslides with a generalized convolutional neural network," Scientific Reports, vol. 11, no. 1, p. 9722, 2021/05/06 2021.

[23] C. Fang, X. Fan, H. Zhong, L. Lombardo, H. Tanyas, and X. Wang, "A Novel Historical Landslide Detection Approach Based on LiDAR and Lightweight Attention U-Net," Remote Sensing, vol. 14, no. 17. doi: 10.3390/rs14174357

[24] K. Bhuyan, S. R. Meena, L. Nava, C. van Westen, M. Floris, and F. Catani, "Mapping landslides through a temporal lens: an insight toward multi-temporal landslide mapping using the u-net deep learning model," GIScience & Remote Sensing, vol. 60, no. 1, p. 2182057, 2023/12/31 2023.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully e
content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2023.3339294

[25] O. Ghorbanzadeh, S. R. Meena, T. Blaschke, and J. Aryal, "UAV-Based Slope Failure Detection Using Deep-Learning Convolutional Neural Networks," Remote Sensing, vol. 11, no. 17. doi: 10.3390/rs11172046

[26] P. Dou, H. Shen, Z. Li, X. Guan, and W. Huang, "Remote Sensing Image Classification Using Deep–Shallow Learning," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 14, pp. 3070-3083, 2021.

[27] P. Dou, H. Shen, Z. Li, and X. Guan, "Time series remote sensing image classification framework using combination of deep learning and multiple classifiers system," International Journal of Applied Earth Observation and Geoinformation, vol. 103, p. 102477, 2021/12/01/ 2021.

[28] S. Plank, A. Twele, and S. Martinis, "Landslide Mapping in Vegetated Areas Using Change Detection Based on Optical and Polarimetric SAR Data," Remote Sensing, vol. 8, no. 4. doi: 10.3390/rs8040307

[29] Z. Qin, Z. Chaoying, and C. Xuerong, "Technical progress and development trend of geological hazards early identification with multi-source remote sensing," Acta Geodaetica et Cartographica Sinica, vol. 51, no. 6, pp. 885-896, 2022.

[30] F. C. Dai, C. F. Lee, and Y. Y. Ngai, "Landslide risk assessment and management: an overview," Engineering Geology, vol. 64, no. 1, pp. 65-87, 2002/04/01/ 2002.

[31] L. Nava, K. Bhuyan, S. R. Meena, O. Monserrat, and F. Catani, "Rapid Mapping of Landslides on SAR Data by Attention U-Net," Remote Sensing, vol. 14, no. 6. doi: 10.3390/rs14061449

[32] L. Nava, O. Monserrat, and F. Catani, "Improving Landslide Detection on SAR Data Through Deep Learning," IEEE Geoscience and Remote Sensing Letters, vol. 19, pp. 1-5, 2022.

[33] S. R. Meena et al., "Landslide detection in the Himalayas using machine learning algorithms and U-Net," Landslides, vol. 19, no. 5, pp. 1209-1229, 2022/05/01 2022.

[34] Y. He et al., "An Extraction Method for Glacial Lakes Based on Landsat-8 Imagery Using an Improved U-Net Network," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 14, pp. 6544-6558, 2021.

[35] Z. Zhang, W. Lu, J. Cao, and G. Xie, "MKANet: An Efficient Network with Sobel Boundary Loss for Land-Cover Classification of Satellite Remote Sensing Imagery," Remote Sensing, vol. 14, no. 18. doi: 10.3390/rs14184514

[36] X. Yu, X. Wu, C. Luo, and P. Ren, "Deep learning in remote sensing scene classification: a data augmentation enhanced convolutional neural network framework," GIScience & Remote Sensing, vol. 54, no. 5, pp. 741-758, 2017/09/03 2017.

[37] M. B. Lyons, D. A. Keith, S. R. Phinn, T. J. Mason, and J. Elith, "A comparison of resampling methods for remote sensing classification and accuracy assessment," Remote Sensing of Environment, vol. 208, pp. 145-153, 2018/04/01/ 2018.

[38] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 3431-3440.

[39] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, Cham, 2015: Springer International Publishing, pp. 234-241.

[40] Z. Niu, G. Zhong, and H. Yu, "A review on the attention mechanism of deep learning," Neurocomputing, vol. 452, pp. 48-62, 2021/09/10/ 2021.

[41] G. Brauwers and F. Frasincar, "A General Survey on Attention Mechanisms in Deep Learning," IEEE Transactions on Knowledge and Data Engineering, vol. 35, no. 4, pp. 3279-3298, 2023.

[42] O. Oktay et al., "Attention u-net: Learning where to look for the pancreas," arXiv preprint arXiv:1804.03999, 2018.

[43] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv preprint arXiv:1412.6980, 2014.

[44] S. Ruder, "An overview of gradient descent optimization algorithms," arXiv preprint arXiv:1609.04747, 2016.

[45] Y. Zhang, X. Meng, G. Chen, L. Qiao, R. Zeng, and J. Chang, "Detection of geohazards in the Bailong River Basin using synthetic aperture radar interferometry," Landslides, vol. 13, no. 5, pp. 1273-1284, 2016/10/01 2016.

[46] Y. Li et al., "Early identification and characteristics of potential landslides in the Bailong River Basin using InSAR technique," National Remote Sensing Bulletin, vol. 25, no. 02, pp. 677-690, 2021.

[47] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba, "Learning deep features for discriminative localization," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 2921-2929.

[48] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-cam: Visual explanations from deep networks via gradient-based localization," in Proceedings of the IEEE international conference on computer vision, 2017, pp. 618-626.

[49] H. Panwar, P. K. Gupta, M. K. Siddiqui, R. Morales-Menendez, P. Bhardwaj, and V. Singh, "A deep learning and grad-CAM based color visualization approach for fast detection of COVID-19 cases using chest X-ray and CT-Scan images," Chaos, Solitons & Fractals, vol. 140, p. 110190, 2020/11/01/ 2020.

[50] K. Burrows, R. J. Walters, D. Milledge, K. Spaans, and A. L. Densmore, "A New Method for Large-Scale Landslide Classification from Satellite Radar," Remote Sensing, vol. 11, no. 3. doi: 10.3390/rs11030237

[51] C. Zhao et al., "Integration of Sentinel-1A, ALOS-2 and GF-1 Datasets for Identifying Landslides in the Three Parallel Rivers Region, China," Remote Sensing, vol. 14, no. 19. doi: 10.3390/rs14195031

[52] G. R. Faqe Ibrahim, A. Rasul, and H. Abdullah, "Improving Crop Classification Accuracy with Integrated Sentinel-1 and Sentinel-2 Data: a Case Study of Barley and

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2023.3339294

Wheat," Journal of Geovisualization and Spatial Analysis, vol. 7, no. 2, p. 22, 2023/07/19 2023.
[53] H. Yoshimatsu and S. Abe, "A review of landslide hazards in Japan and assessment of their susceptibility using an analytical hierarchic process (AHP) method," Landslides, vol. 3, no. 2, pp. 149-158, 2006/05/01 2006.
[54] R. Sonobe, "Combining ASNARO-2 XSAR HH and Sentinel-1 C-SAR VH/VV Polarization Data for Improved Crop Mapping," Remote Sensing, vol. 11, no. 16. doi: 10.3390/rs11161920
[55] Z. Dong, S. An, J. Zhang, J. Yu, J. Li, and D. Xu, "L-Unet: A Landslide Extraction Model Using Multi-Scale Feature Fusion and Attention Mechanism," Remote Sensing, vol. 14, no. 11. doi: 10.3390/rs14112552
[56] S. Tavakkoli Piralilou et al., "Landslide Detection Using Multi-Scale Image Segmentation and Different Machine Learning Models in the Higher Himalayas," Remote Sensing, vol. 11, no. 21. doi: 10.3390/rs11212575
[57] S. F. Tebbens, "Landslide Scaling: A Review," Earth and Space Science, https://doi.org/10.1029/2019EA000662 vol. 7, no. 1, p. e2019EA000662, 2020/01/01 2020.
[58] Y. Yi, Z. Zhang, W. Zhang, C. Zhang, W. Li, and T. Zhao, "Semantic Segmentation of Urban Buildings from VHR Remote Sensing Imagery Using a Deep Convolutional Neural Network," Remote Sensing, vol. 11, no. 15. doi: 10.3390/rs11151774
[59] Y. Wei and S. Ji, "Scribble-Based Weakly Supervised Deep Learning for Road Surface Extraction From Remote Sensing Images," IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1-12, 2022.
[60] M. Wang, X. Zhang, X. Niu, F. Wang, and X. Zhang, "Scene Classification of High-Resolution Remotely Sensed Image Based on ResNet," Journal of Geovisualization and Spatial Analysis, vol. 3, no. 2, p. 16, 2019/10/10 2019.



Hesheng Chen received the B.E. degree in Network Engineering from Henan University, Kaifeng, China, in 2021. He is currently working toward the M.S. degree at faculty of geomatics, Lanzhou Jiaotong University, Lanzhou, China.

His research interests include landslide detection and extraction based on deep learning combined with remote sensing data.



Yi He received the B.S. in geographic information system from Lanzhou Jiaotong University, Lanzhou, China, in 2011, and the Ph.D. degree in earth system science from Lanzhou University, China, in 2016. He has been a Postdoctoral Researcher with the school of environment and municipal engineering, Lanzhou Jiaotong University, Lanzhou, China. He is currently a professor with the faculty of geomatics, Lanzhou Jiaotong University, Lanzhou, China.

His research interests include disaster remote sensing, ecological remote sensing, image processing and time series InSAR prediction based on deep learning.



Lifeng Zhang received the M.S. degree in cartography and geographic information system from Lanzhou Jiaotong University, Lanzhou, China, in 2010, and a Ph.D. degree in environmental science and engineering from Lanzhou Jiaotong University, Lanzhou, China in 2017. He is a professor, Lanzhou Jiaotong University, Lanzhou, China.

His research interests include ecological environment remote sensing monitoring, visualized analysis, and analysis of land use change.



Wang Yang received the M.S. degree in Surveying and Mapping from Lanzhou Jiaotong University, Lanzhou, China, in 2023. He is currently working toward the Ph.D. degree at faculty of geomatics, Lanzhou Jiaotong University, Lanzhou, China.

His research interests include InSAR data processing technology, machine learning, and remote sensing image information extraction.



Yaoxiang Liu received the B.E. degree in Remote sensing science and technology from the Lanzhou Jiaotong University, Lanzhou, China, in 2019. He is currently working toward the M.S. degree at the faculty of geomatics, Lanzhou Jiaotong University, Lanzhou, China.

His research interests include InSAR data processing technology, machine learning and remote sensing image information extraction.



Binghai Gao received the B.E. degree in remote sensing science and technology from the Lanzhou Jiaotong University, Lanzhou, China, in 2021. He is currently working toward the M.S. degree in mapping engineering with the Faculty of Geomatics, Lanzhou Jiaotong University, Lanzhou, China.

His research interests include InSAR data processing technology, deep learning, and remote sensing image information extraction.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully e
content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2023.3339294



Qing Zhang received the B.E. degree in surveying and mapping engineering from the Lanzhou Jiaotong University, Lanzhou, China, in 2022. He is currently working toward the M.S. degree in mapping science and technology with the Faculty of Geomatics, Lanzhou Jiaotong University, Lanzhou, China.

His research interests include InSAR data processing technology, machine learning, and remote sensing image processing.



Jiangang Lu received the B.E. degree in remote sensing science and technology from the Shandong University of Science and Technology, Qingdao, China, in 2020. He is currently working toward the M.S. degree in mapping science and technology with the Faculty of Geomatics, Lanzhou Jiaotong University, Lanzhou, China.

His research interests include landslide hazard risk assessment, machine learning, and remote sensing image information extraction.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/