Feature-Based Constraint Deep CNN Method for Mapping Rainfall-Induced Landslides in Remote Regions With Mountainous Terrain: An Application to Brazil

Guosen Xu , Yi Wang , *Member, IEEE*, Lizhe Wang , *Fellow, IEEE*, Lucas Pedrosa Soares , and Carlos H. Grohmann

Abstract—Landslides have caused tremendous damage to human lives and property safety. However, the complex environment of mountain landslides and the vegetation coverage around landslides make it difficult to identify landslides quickly and efficiently using high-resolution images. To address this challenge, this article presents a feature-based constraint deep U-Net (FCDU-Net) method to detect rainfall-induced mountainous landslides. Usually, the vegetation in the landslide area is severely damaged, and the vegetation coverage can indirectly reflect the spatial extent of the landslide. Meanwhile, the texture features of high-resolution images can characterize the surface environment of landslide hazards to a certain extent. We first introduce auxiliary features of normalized difference vegetation index and gray-level co-occurrence matrix into the proposed method to further improve the detection performance. Then, to minimize the information redundancy of these features and the image, we combine Relief-F and Deep U-Net to screen the optimal features to effectively identify accurate and detailed landslide boundaries. Compared with traditional semantic segmentation methods, the FCDU-Net method can capture fine-grained details in high-resolution images and produce more accurate segmentation results. We conducted experiments by applying the proposed method and other most popular semantic segmentation methods to a high-resolution RapidEye image in Rio de Janeiro, Brazil. The results demonstrate that the FCDU-Net method can achieve better landslide detection results than the other semantic segmentation methods, and the evaluation measures of Precision, F1 score, and mean Intersection-over-Union are as high as 88.87%, 81.17%, and 83.19%, respectively. Furthermore, we quantitatively analyze the effect of the convolution input window size on the performance of FCDU-Net in detecting landslides. We believe that FCDU-Net can serve as a reliable tool for fast and accurate regional landslide hazard surveys.

Index Terms—Auxiliary features, deep convolutional neural networks (DCNNs), feature selection, high-resolution image, landslide detection.

Guosen Xu is with the Institute of Geophysics and Geomatics, China University of Geosciences, Wuhan 430074, China (e-mail: guosennn@163.com).

Yi Wang is with the Institute of Geophysics and Geomatics, and State Key Laboratory of Biogeology and Environmental Geology, China University of Geosciences, Wuhan 430074, China (e-mail: cug.yi.wang@gmail.com).

Lizhe Wang is with the Institute of Computer Science and State Key Laboratory of Biogeology and Environmental Geology, China University of Geosciences, Wuhan 430074, China (e-mail: lizhe.wang@gmail.com).

Lucas Pedrosa Soares is with the Institute of Geosciences, University of São Paulo, São Paulo 05508-080, Brazil (e-mail: lpsoares@usp.br).

Carlos H. Grohmann is with the Institute of Energy and Environment, University of São Paulo, São Paulo 05508-010, Brazil (e-mail: guano@usp.br).

I. INTRODUCTION

AS A common geological disaster, landslides cause immeasurable damage to infrastructure and seriously threaten the safety of human life and property [1]–[3]. Therefore, efficiently and accurately locating of landslide areas is crucial for emergency rescue and disaster warning. Landslide inventory data is the main resource for landslide research at different scales [4]. The training and testing of landslide prediction begins with the creation of a landslide inventory map [5]. Traditional landslide inventory mapping methods mainly rely on field survey and visual interpretation of aerial images [6], but it requires personnel with professional knowledge to complete it, wastes a lot of manpower and material resources, and cannot meet the needs of efficient and accurate positioning of landslides.

Remote sensing data can cover hundreds of square kilometers, providing rich image data for disaster observation. Optical and radar remote sensing images have been widely used to observe the landscape changes of landslides [7]–[12]. Traditional image processing techniques mainly detect landslides on remote sensing images through statistical methods [10] and semi-automatic methods [13]. However, these methods need to obtain the image features of the landslide area in advance by artificial means, and also require a lot of parameter modulations, and the process is complicated.

Satellite images are significant for identifying landslides following large-scale events, especially when the disaster occurs in remote or hard-to-reach areas [14]. Generally, landslides can be automatically identified in aerial photographs, high-resolution

and very-high-resolution remote sensing images by using pixelwise identification methods and object-based information extraction (OBIE) methods [15]. Among them, pixelwise methods can overcome the shortcomings of human interpretation, and can be mainly separated into two main categories: supervised learning and unsupervised learning. The former includes discriminant analysis, feature analysis, maximum likelihood analysis, support vector machine, etc., and the latter includes iterative self-organizing data analysis technique algorithm and K-means. These pixelwise methods were used to identify landslides in Taiwan and compared with visual interpretation results [16]. After discussing six different image fusion methods, Marcelino *et al.* [17] proved that the intensity-hue-saturation technique can better preserve the information of the original image and used it to identify landslide scars. In addition, the change detection technique allows the extraction of landslides from images at different time phases [18]–[25]. The pixelwise methods only exploit the spectral features of different materials in the image without considering the geometric and contextual information in the image. Therefore, these methods are prone to "pepper and salt" noise in the resultant map. However, OBIE-based methods can consider multiple features, such as spectral (pixel value, hue, color, etc.), spatial (landslide area size, shape, contour, etc.), and background features [26], which are widely used for landslide detection using airborne and satellite images [27], [28]. For example, Martin *et al.* [29] took the image segmentation result as the basic unit and established landslide recognition constraints based on shape, texture, and adjacent features, and the classification accuracy reached 65%. Based on FORMOSAT-2 images and DEM data, Rau *et al.* [30] identified landslides in the mountainous areas of central and southern Taiwan through multilevel segmentation and hierarchical semantic network, and achieved good experimental results. Meanwhile, conventional classification methods such as support vector machine [31] and random forest [28] have also been introduced into the OBIE-based landslide detection, and their effectiveness has been verified. However, some empirical parameters of OBIE rely too much on experts and the degree of automation is low. Thus, there are still certain limitations in practical applications [32], [33].

Convolutional neural networks (CNNs) can effectively capture abstract features from the original image [34]. In recent years, CNNs have attracted extensive attention and applications in the fields of large-scale image recognition, target detection, and semantic segmentation due to their advantages of local receptive, weight sharing, and connection sparsity [35]–[38]. However, only a few articles have used CNNs for landslide detection based on remote sensing image [2]. Ding *et al.* [39] first used texture and spectral features to remove the interference of background features such as buildings, water bodies, and vegetation, and then used CNNs to capture remote sensing image information before and after the landslide and calculated the Euclidean distance of the two features to determine whether there is a landslide on the image. Yu *et al.* [40] first screened out candidate images with landslides from remote sensing images through a shallow CNN, and then used a region growing algorithm to determine landslide boundaries and centers. Ghorbanzadeh *et al.* [1] performed different numbers of convolutions

and different levels of feature fusion on remote sensing images of southern Nepal. To further reduce the influence of background on the image, Ji *et al.* [2] developed a deep convolutional neural network (DCNN) for end-to-end landslide target detection. Yi *et al.* [41] constructed a cascaded end-to-end DCNN to intelligently detect earthquake-triggered landslides. Qin *et al.* [42] applied distant domain transfer learning and traditional CNNs for landslide detection.

In summary, traditional feature extraction methods and CNNs have been used for landslide detection in high-resolution images. However, only a few articles have used deep learning techniques to detect landslides in areas with complex terrain, especially those triggered by extreme weather such as heavy rain. We believe that the task of detecting landslides in mountainous areas caused by heavy rainfall still faces the following challenges. First, the background of remote sensing images with mountainous terrain is very complex, including quarries, terraces, slopes, and riverbeds, which may be similar in characteristics to landslides and easily interfere with detecting them. Second, landslides usually appear at different scales on remote sensing images, and the length of landslides may vary from a few meters to several kilometers. Finally, there are obvious seasonal changes in remote sensing images of mountainous areas, and the specific characteristics of coverage in different seasons may make it difficult to identify landslides.

To fill a gap in the previous articles, we present a feature-based constraint deep U-Net (FCDU-Net) method to map rainfall-induced landslides in remote areas with mountainous terrain. The main contributions are summarized as follows:

1) We only need to acquire postdisaster optical remote sensing imagery and process landslide data through data augmentation strategies, without inputting other data such as topographic factors and multitemporal images.

2) Deep U-Net (DU-Net) is used as a prediction model for landslide detection. Since it combines the advantages of U-Net and DenseNet, it can capture fine-grained details in high-resolution images and produce more accurate semantic segmentation results than most popular networks.

3) The FCDU-Net method combines auxiliary features with the original image to improve the identification accuracy of landslides. Usually, the vegetation in the landslide area is severely damaged, and the vegetation coverage can indirectly reflect the spatial extent of the landslide. Meanwhile, the texture features of high-resolution images can characterize the surface environment of landslide hazards to a certain extent. On this basis, normalized difference vegetation index (NDVI) and gray-level co-occurrence matrix (GLCM) are input into the network as auxiliary features. Furthermore, to minimize the information redundancy of these features and the image, we combine Relief-F and DU-Net to screen the optimal features for subsequent processes, and effectively identify accurate and detailed landslide boundaries.

The rest of this article is organized as follows. Section II introduces the study area. Section III describes the auxiliary features used in this article and the modified DU-Net method. Section IV analyzes landslide detection results of different
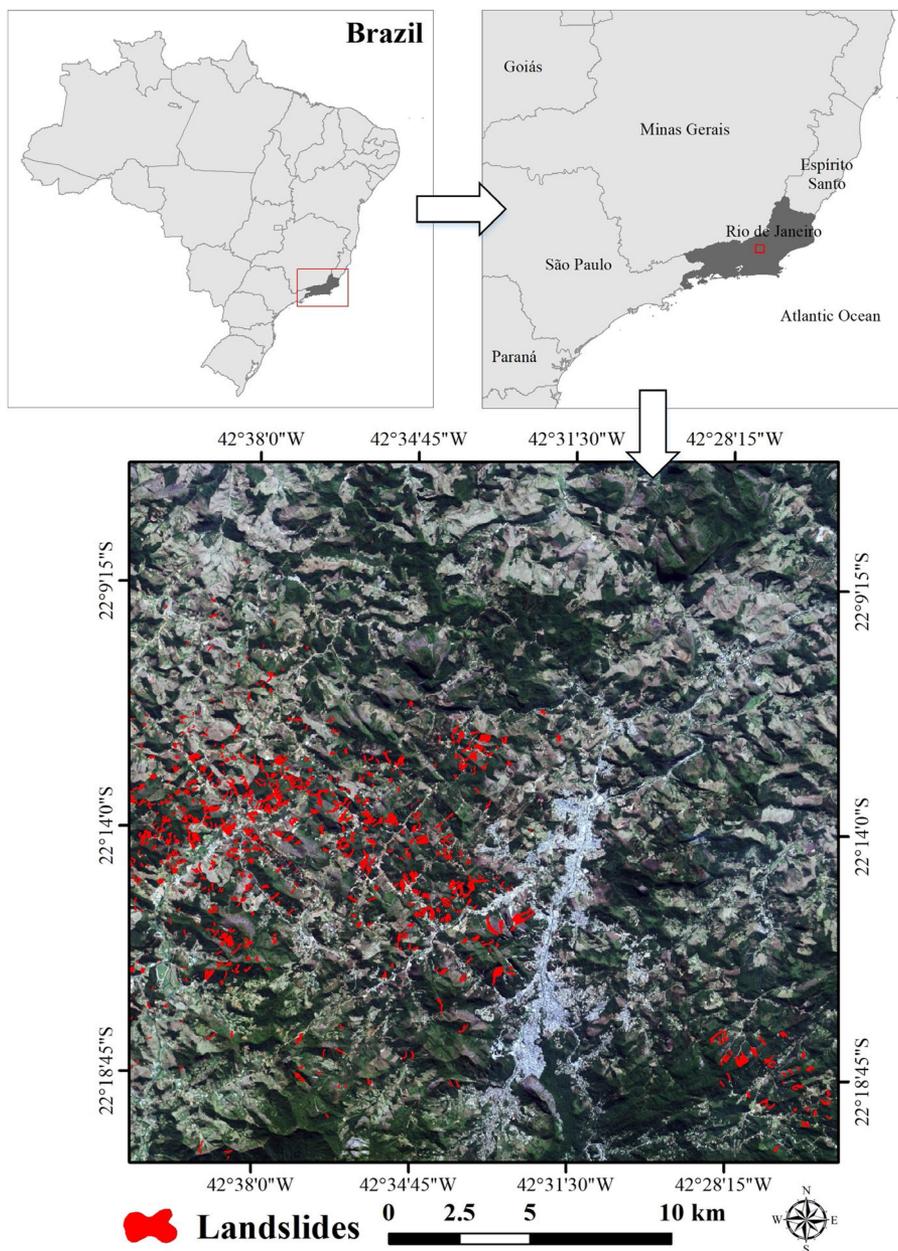
Fig. 1. Location of the study area and annotated rainfall-induced landslides (red polygons).

methods. Section V discusses the effectiveness of the proposed method. Finally, Section VI concludes the article.

II. STUDY AREA AND DATA

The study area is located in the Nova Friburgo mountains in Rio de Janeiro, Brazil, as shown in Fig. 1. In January 2011, a heavy rainfall event of 350 mm/48 h occurred in Nova Friburgo, triggering at least 3500 translational landslides with a total area of about 5.56 km^2, and more than 1500 people died from the rainfall-induced landslides [43]. The landslide inventory map of the study area was manually generated by experts from high-resolution satellite images based on shape and tone information, and validated using Google Earth Pro and compared with landslide maps produced by other scholars [44]. The high-resolution image used in this article was acquired by the RapidEye sensor on August 13, 2011. It has 5000×5000 pixels and five bands in the range of 440–850 nm, with a spatial and temporal resolution of 5 m and 5.5 days, respectively. In the image, 816 landslides caused by heavy rainfall were marked as solid polygons of different sizes, ranging from 200.32 to 78117.35 m^2, as shown in Fig. 1.

III. METHODOLOGY

A. Auxiliary Features

1) NDVI: NDVI, a remote sensing index reflecting the status of land vegetation cover, has been widely employed in land use

TABLE I
CALCULATION FORMULA OF TEXTURE FEATURE FACTOR

Texture feature factor	Formula
Mean	$Mea = \sum_{i,j=0}^{N-1} i \cdot P_{i,j}$
Variance	$Var = \sum_{i,j=0}^{N-1} i \cdot P_{i,j} (i - Mea)^2$
Homogeneity	$Hom = \sum_{i,j=0}^{N-1} \frac{i \cdot P_{i,j}}{1 + (i-j)^2}$
Contrast	$Con = \sum_{i,j=0}^{N-1} (i-j)^2 \cdot P_{i,j}$
Dissimilarity	$Dis = \sum_{i,j=0}^{N-1} i \cdot P_{i,j} \lvert i-j \rvert$
Entropy	$Ent = \sum_{i,j=0}^{N-1} i \cdot P_{i,j} \ln P_{i,j}$
Second Moment	$Sec = \sum_{i,j=0}^{N-1} i \cdot P_{i,j}^{~2}$
Correlation	$Cor = \sum_{i,j=0}^{N-1} i \cdot P_{i,j} \frac{(i-Mea)(j-Mea)}{\sqrt{Var_1 \cdot Var_2}}$

and land cover change. In the study of landslide identification, it is generally believed that the vegetation coverage inside the landslide body will change with the movement of the landslide. Therefore, this article introduces the index to strengthen the distinction between vegetation areas and nonvegetation areas

$$NDVI = \frac{NIR - R}{NIR + R} \tag{1}$$

where NIR and R represent the reflectivity in the near-infrared band and in the red band, respectively.

2) GLCM: The GLCM is a classic analysis method that extracts texture features by calculating the conditional probability density between gray levels of an image, and is widely used for information extraction of remote sensing images. The texture of the landslide is significantly different from the surrounding geological environment. In order to synthesize the information of each band and reduce the dimensionality of data, the principal component analysis is used on the original image, and eight GLCM texture features of the first principal component are calculated. Table I lists the texture features of high-resolution images used for deep learning modeling. The **Mean** indicates how regular the texture distribution is. The stronger the regularity, the greater the mean. **Variance** is the difference between each pixel and the mean. If the gray value of the pixel changes greatly, the variance value is larger. **Homogeneity** is a measure of uniformity within a local area. In this area, the more uniform the image distributed, the greater the homogeneity. **Contrast** represents how the matrix values are distributed and to what extent local variations in the image reflect the sharpness of the image and the depth of the grooves in the texture. The greater the contrast, the deeper the grooves, and the sharper the effect. **Dissimilarity** is similar to **Contrast**. In a local area, the higher the contrast, the stronger the dissimilarity. **Entropy** is a measure

of the randomness of the information contained in an image. Entropy is greatest when all means in the GLCM are equal or the pixel values show the greatest randomness. Therefore, the higher the entropy value, the more complex the grayscale distribution of the image. The **second moment** represents the uniformity of the grayscale distribution of the image and texture thickness. When the image texture is fine and the grayscale distribution is uniform, the second angular moment is larger. **Correlation** refers to the similarity of the grayscale of the image in the row or column direction. The texture reflects the local grayscale correlation of the image. The larger the value is, the greater the correlation.

B. Deep Convolutional Neural Networks

1) Fully Convolutional Network: To apply deep learning CNNs to semantic segmentation, a full convolutional network (FCN) was proposed [45], which replaces the last part of the CNN with a convolution layer (a kernel of size 1×1), allowing prediction of two-dimensional dense class label maps [46]. The semantic segmentation effect of point-to-point classification is achieved by restoring the feature map to the input size while preserving the spatial structure of the image. It modifies the fully connected layers of deep CNNs (AlexNet, VGG, GoogleNet, etc.) for image classification into convolutional layers, and proposes an end-to-end semantic segmentation model.

2) U-Net: U-Net was proposed to improve the situation where the target edges are not fine enough in the resultant segmentation map [47]. In this network, the number of channels of the feature map is doubled during each downsampling procedure to allow more feature information (e.g., boundaries, colors, shapes, etc.) to propagate among the convolutional layers, while this number is halved during each upsampling procedure. The skip connection is performed and the stacking operation of dimensional concatenation to fuse feature maps of different scales. This process preserves more high-resolution details to help restore the spatial resolution of the output map and improve segmentation accuracy [48]. This network has been widely used in target extraction from remote sensing images [49]–[51].

3) DeepLabv3+: DeepLabv3+ was proposed to use atrous convolutions to improve segmentation performance [52]. Specifically, it uses atrous convolutions [53] with different strides [54]. The receptive field is expanded through atrous convolution without increasing the network training parameters and reducing the resolution of the output feature map. In essence, DeepLabv3+ fuses the shallow features of the encoder with the deep features of the atrous spatial pyramid pooling [55] module to output more accurate semantic segmentation results. Shallow features can provide better object details, and deep features can provide more abstract semantic information and location information of objects. The fusion of the two features can output higher precision semantic segmentation results.

To better compare the aforementioned methods, Table II lists their advantages and disadvantages.

C. Feature Selection Based on Relief-F

Relief-F is a multiclass feature selection algorithm based on mathematical statistics, calculating the weights of each

TABLE II
COMPARISON OF CHARACTERISTICS OF THREE SEMANTIC SEGMENTATION METHODS

Algorithm	Characteristics	Advantages	Disadvantages
FCN	It replaces the fully connected layers in the CNN model with fully convolutional layer for pixel-level dense prediction.	FCN solves the problems of repeated storage and convolution computation caused by the use of pixel blocks in traditional segmentation networks.	The deconvolution process is coarse and insensitive to image details.
U-Net	It uses skip connections to stitch the feature maps of the encoding network onto the corresponding feature maps of the decoding network.	On the basis of FCN, a decoding network is added to effectively fuse low-level and high-level features to restore the spatial dimension and boundary information of the image.	The network structure is complex, the number of parameters is large, and the object boundary segmentation performance is not good.
DeepLabv3+	It use the ASPP module to process multi-scale contextual content information for optimization.	The receptive field is enlarged, and more feature information can be extracted.	The computational cost is high, some position information of the pixels is lost, and the local consistency of the feature map is affected.
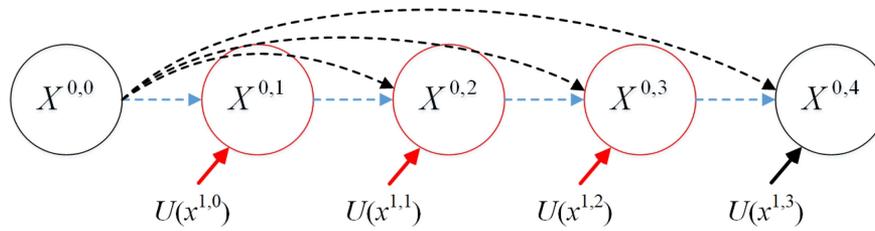
Fig. 2.   Input and output of the first layer of DU-Net.

feature variable and rank the magnitude of the weight values by randomly selecting samples [56]. The Relief-F method can determine the relative importance of the features to landslide occurrence, and it evaluates the value of features by considering the correlation between features and categories [57].

The Relief-F algorithm first selects a random sample R from the dataset D, and then selects k nearest neighbor samples from the similar and different classes of the sample R. In the feature set F, if the distance between samples of different classes is greater than that of samples of similar classes in one feature, the feature is conducive to classification, and its weight is increased. The process repeats n times to take the mean of all the results as the final weight value of each feature. Finally, the feature weight is defined as follows:

$$\omega\left(F_j\right)=\omega\left(F_j\right)-\frac{1}{n \cdot k} \sum_{h \in H}\left|R_j-h_j\right|+\frac{1}{n \cdot k} \sum_{m \in M}\left|R_j-m_j\right|$$
(2)

where $\omega(F_j)$ is the weight value of the jth feature, k is the number of nearest neighbor samples, $\sum_{h \in H}\left|R_j-h_j\right|$ indicates the sum of the distances between the selected k similar class neighbor samples and the sample R on the jth feature, and $\sum_{m \in M}\left|R_j-m_j\right|$ represents the sum of distances between k heterogeneous nearest neighbor samples and the sample R on the jth feature.

D. Deep U-Net Architecture

DU-Net [58] can effectively combine the advantages of traditional U-Net [47] and DenseNet [59], using skip connection to combine low-level and high-level information in U-Net, while using dense connection to enhance feature propagation in DenseNet. In U-Net, the feature maps of the encoder are received directly in the decoder. However, in DU-Net, the redesigned dense skip connections change the connectivity of the encoder and decoder subnetworks, and the number of convolutional layers depends on the network level. For example, as shown in Fig. 2, the layer-jumping path between nodes $x^{0,0}$ and $x^{0,3}$ is composed of dense convolutional blocks with three convolutional layers. Each convolutional layer is preceded by a connection layer that merges the output of the previous convolutional layer from the same dense block and the corresponding upsampled output of the lower dense block. The DU-Net structure can capture subtle information, obtaining better output than U-Net. Therefore, the use of DU-Net for semantic segmentation of remote sensing images has great application value.

The formula for DU-Net's dense skip connection is given by

$$x^{a,b}=\begin{cases} H\left(x^{a-1,b}\right), & b=0 \\ H\{[[x^{a,k}]_{k=0}^{b-1}, U\left(x^{a+1,b-1}\right)]\} & b>0 \end{cases}$$
(3)

where $x^{a,b}$ represents the output of the b th node of the a th layer of the network, and $H\left(\cdot\right)$, $U\left(\cdot\right)$, and $[\cdot]$ represent convolution operation, upsampling operation, and concatenation operation, respectively. When $b=0$, the node only takes as input from the downsampling layer of the upper layer. There are two inputs passed to the node with $b=1$, both of which come from the encoding subnetwork and are two consecutive layers. When $b+1$ inputs are passed to a node with $b>1$, b of those inputs
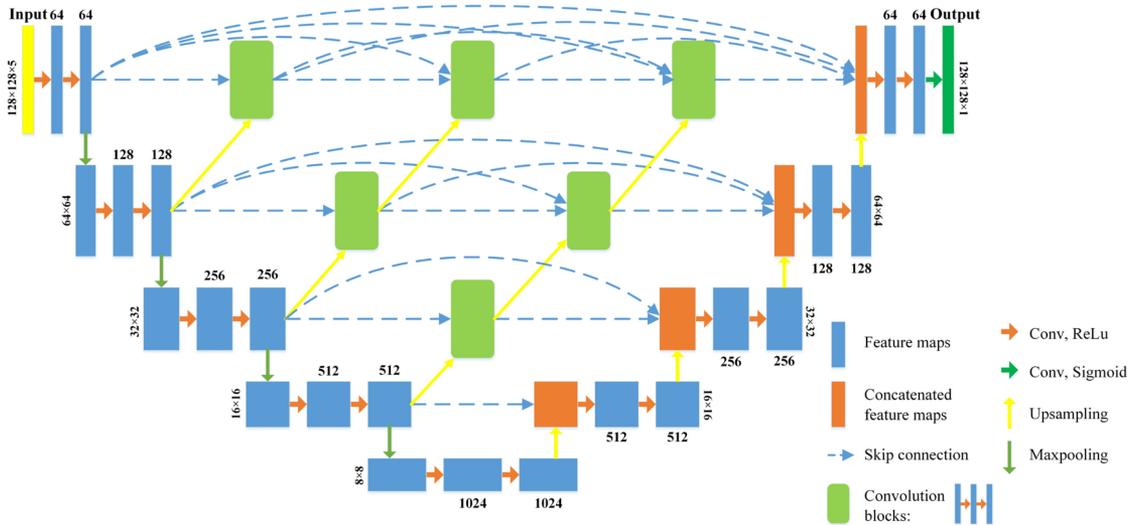
Fig. 3. Architecture of DU-Net.

come from the output of the previous node in the skip path consistent with that node, while one more input is obtained from a skip path lower than that node after an upsampling operation.

We propose an FCDU-Net method to detect landslides, which fuses spectral bands, NDVI, GLCM texture features, and uses the Relief-F algorithm to filter out positive factors from auxiliary features and uses the DU-Net model for feature extraction and training. Fig. 3 shows the DU-Net architecture, consisting of an encoding structure, a decoding structure, and dense skip connections. The feature maps of the same layer have the same size, and from top to bottom, they are the first to fifth layers. According to [58], three, two, and one convolutional modules are included between the first, second, and third layers of the encoder–decoder path, respectively. Each convolution module is composed of two sets of convolutional layers with a kernel size of 3 and a rectified linear unit. The overall feature fusion structure of DU-Net is in the shape of an inverted pyramid. Each layer of the pyramid is equivalent to a dense block, and the features are transferred backward through dense skip connections. Among different layers, the features are deeply aggregated from top to bottom, which shortens the semantic gap when the encoder features and decoder features skip and connect, and improves the feature learning ability of the model.

E. Flowchart of Landslide Inventory Mapping

Fig. 4 illustrates the overall flow of the proposed method. First, training samples are generated through data preprocessing, including image normalization, feature selection, and landslide data processing. The proposed method is supervised and trained based on the training dataset and real samples. Meanwhile, the model is continuously calculated and analyzed based on validation data to ensure satisfactory generalization to unknown datasets. After the validation loss reaches the minimum value, the training process terminates. The model inference process

adopts the edge-ignoring prediction method [60], and the pros and cons of the model are evaluated based on the test dataset.

1) Data Preprocessing: Data preprocessing mainly includes image cropping, normalization, and preparation of landslide training datasets. Different features of the target often have different dimensional units, which will affect the results of data analysis. To eliminate the influence between indicators, image normalization is usually required to make the data indicators in the same order of magnitude and facilitate the training of deep CNNs. In this article, each channel of the original image is normalized by Min–Max normalization so that the input has a similar distribution. Due to the diversity of landslide shapes and sizes, multiple input window sizes are used for landslide detection. First, an *n*-band high-resolution image and a single-band rasterized ground-truth label layer are stacked to form a (*n*+1)-channel array to ensure that feature transformations remain in the image. We then scan the image using a sliding window algorithm to generate training patches, as shown in Fig. 5.

To better train a deep learning network, it is necessary to select a sufficient number of samples to train the network. If the number of training image sets is small, the scenes contained in these image data may be relatively homogeneous, and the network model cannot learn features extensively, so the generalization ability of the trained model is not strong. Inspired by previous articles in similar fields [61], [62], we use a data augmentation strategy to generate more training samples. Next, these patches are expanded horizontally, vertically, and diagonally, which increases the invariance and robustness of the network to such deformations. All training samples are generated from the training area, and images from the test area are not used during training.

2) Model Training: Using a small-batch training strategy, only a small portion of the training dataset is used as input in each iteration to avoid local minimizing of training errors and achieve rapid convergence in the parameter optimization process [63]. In addition, the scales of landslides are different. We randomly
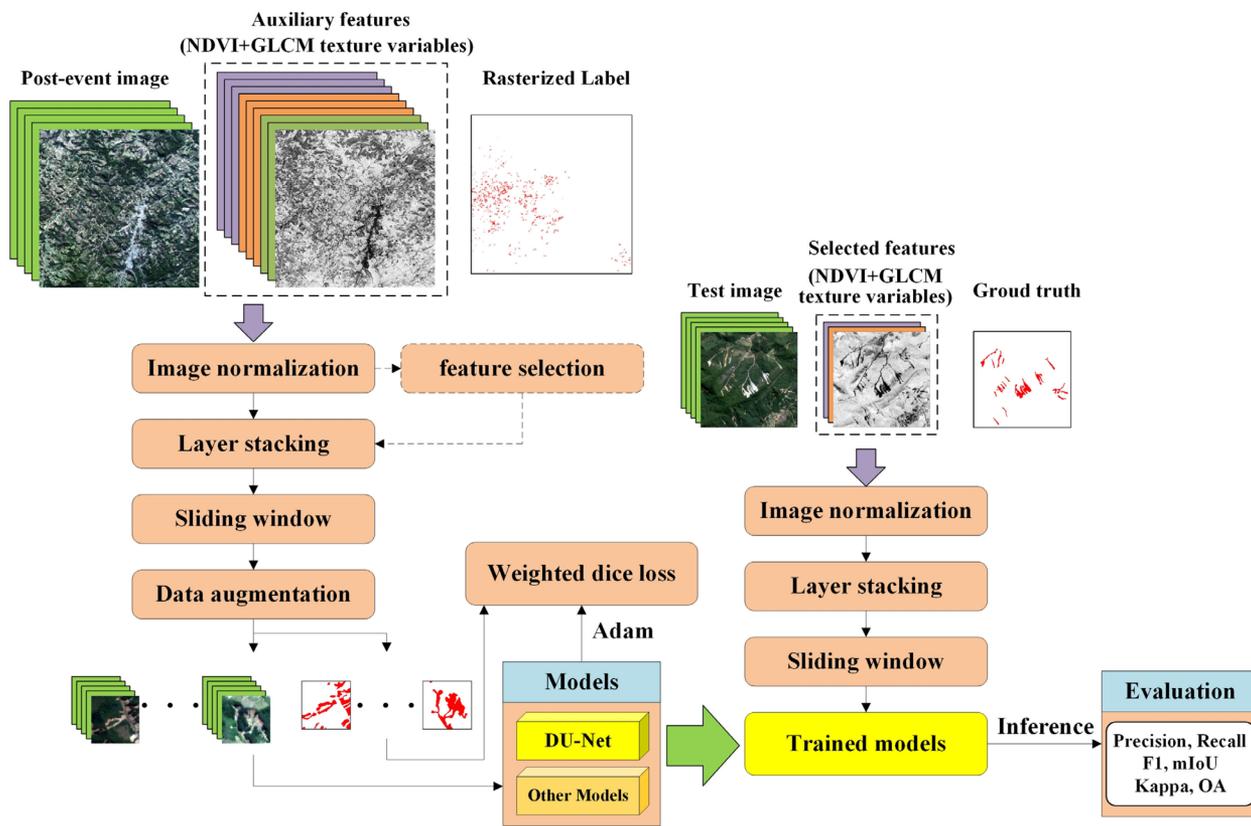
Fig. 4. Flowchart of this article.
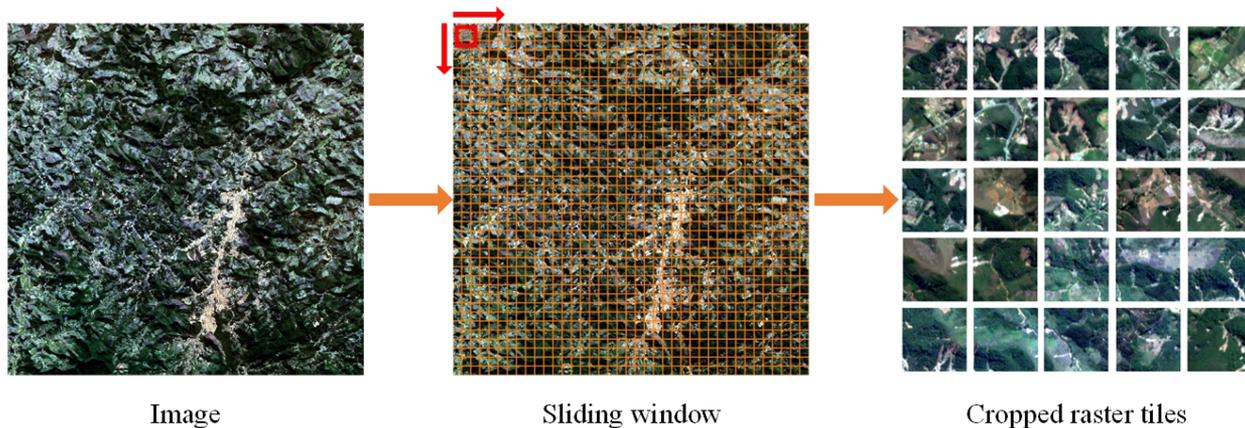
Fig. 5. Flowchart of cropping a high-resolution image to generate raster tiles through scanning the image using a sliding window.

selected three test areas. The total training area contains 749 landslides (88%) and test area contains 98 landslides (12%). For imbalanced datasets, this may result in the minority categories being easily ignored since classifiers are usually designed to optimize overall accuracy (OA) [64], [65]. Dice loss [66] is especially proposed as a loss function for medical images with too large proportion of background information and too small proportion of foreground information, resulting in imbalance of positive and negative sample categories. Taking into account the difference in landslide and background proportions, this article uses the weighted dice loss. Specifically, we weight the losses

for different categories and calculate the weight for each class. The higher the frequency, the smaller the lower of the category.

3) Model Validation and Inference: During training, 30% of each dataset is used as validation data, and they are not augmented to evaluate the generalization ability of DCNNs. In this article, a threshold of 0.5 is chosen to classify the results into binary classes. Specifically, if the predicted probability value of a certain pixel is greater than 0.5, the pixel is considered to be a landslide unit, and the pixel with a probability value less than 0.5 is a nonlandslide unit. Next, we conduct model testing to analyze the quality of the network. In particular, we select a
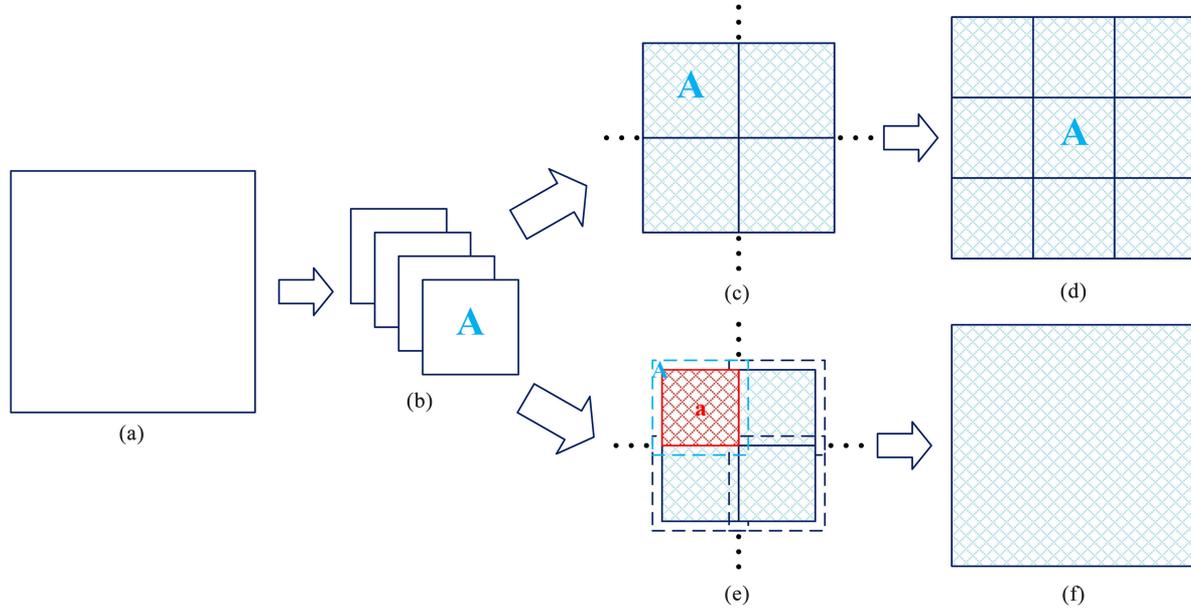
Fig. 6.   Sketch map of model inference. (a) Test area image. (b) Cropping the image. (c) Inference results are directly spliced together. (d) Direct splicing results. (e) Splicing inference results while ignoring the edges. (f) Edge-ignoring inference result.

robust model that does not overfit in the training dataset, and evaluate it qualitative and quantitative through visual inspection and different objective metrics.

What is more, the automatic landslide detection model uses datasets from the test area for evaluation. The test area image is sliced into a series of smaller images to be input to the network for inference, and the prediction results are merged into a mosaic to fully cover the original range, as shown in Fig. 6. Direct stitching prediction methods often result in splicing traces, affecting landslide inventory mapping. In this article, we perform edge-ignoring prediction, i.e., cropping overlapping images and adopt the edge-ignoring strategy when splicing. In Fig. 6(e) and (f), the actual cropped image is predicted to be A, then the stitching result is a. The percentage of a in area A is r, and the overlap ratio of adjacent cropped images is $1 - \sqrt{r}$, where the value of r is 0.5.

F. Accuracy Evaluation

To quantitatively analyze the effectiveness of the proposed method, the experiments calculate several evaluation measures of Precision, Recall, and F1 score (F1). Precision indicates that the correctly recognized landslides are divided by the aggregate number of landslides recognized by the model. Recall means that the correctly recognized landslides are divided by the aggregate number of actual landslides. F1 is used as a standard measure of model performance, which represents the harmonic average of the measures of Precision and Recall. Besides, two statistical measures of Kappa and OA are applied to evaluate the predictive ability of different landslide inventory mapping methods. The aforementioned measures are given by

$$\text{Precision} = \frac{T_P}{T_P + F_P} \tag{4}$$

$$\text{Recall} = \frac{T_P}{T_P + F_N} \tag{5}$$

$$F_1 = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \tag{6}$$

$$\text{Kappa} = \frac{P_c - P_{\exp}}{1 - P_{\exp}} \tag{7}$$

$$P_{\exp} = \frac{(T_P + F_N)(T_P + F_P) + (F_P + T_N)(F_N + T_N)}{(T_P + T_N + F_N + F_P)^2} \tag{8}$$

$$\text{Overall Accuracy} = \frac{T_P + T_N}{T_P + F_P + T_N + F_N} \tag{9}$$

where T_P and T_N represent the number of positive and negative samples correctly classified by the model, respectively, and F_P and F_N denote the number of positive and negative samples incorrectly classified by the model, respectively. P_c is the observed agreement rate meaning the percentage of type consistent fractions in the two datasets, and $P_{\exp}$ is the expected probability of change agreement. The Kappa coefficient is generally calculated to be between -1 and 1, but usually between 0 and 1. The higher the F1, the better the prediction.

Mean Intersection-over-Union (mIoU) is a commonly adopted validation metric in computer vision. The intersection and union of the predicted area and the true area are obtained, and mIoU is the ratio of these two components. In this article, these two sets are the landslide interpretation map and the landslide prediction map. The value ranges from 0 to 1, where 1 means the best result. The measure of mIoU is given by

$$\text{mIoU} = \frac{1}{m+1} \sum_{j=0}^{m} P_{ij} \frac{P_{ii}}{\sum_{j=0}^{m} P_{ij} + \sum_{j=0}^{m} P_{ji} - P_{ii}} \tag{10}$$
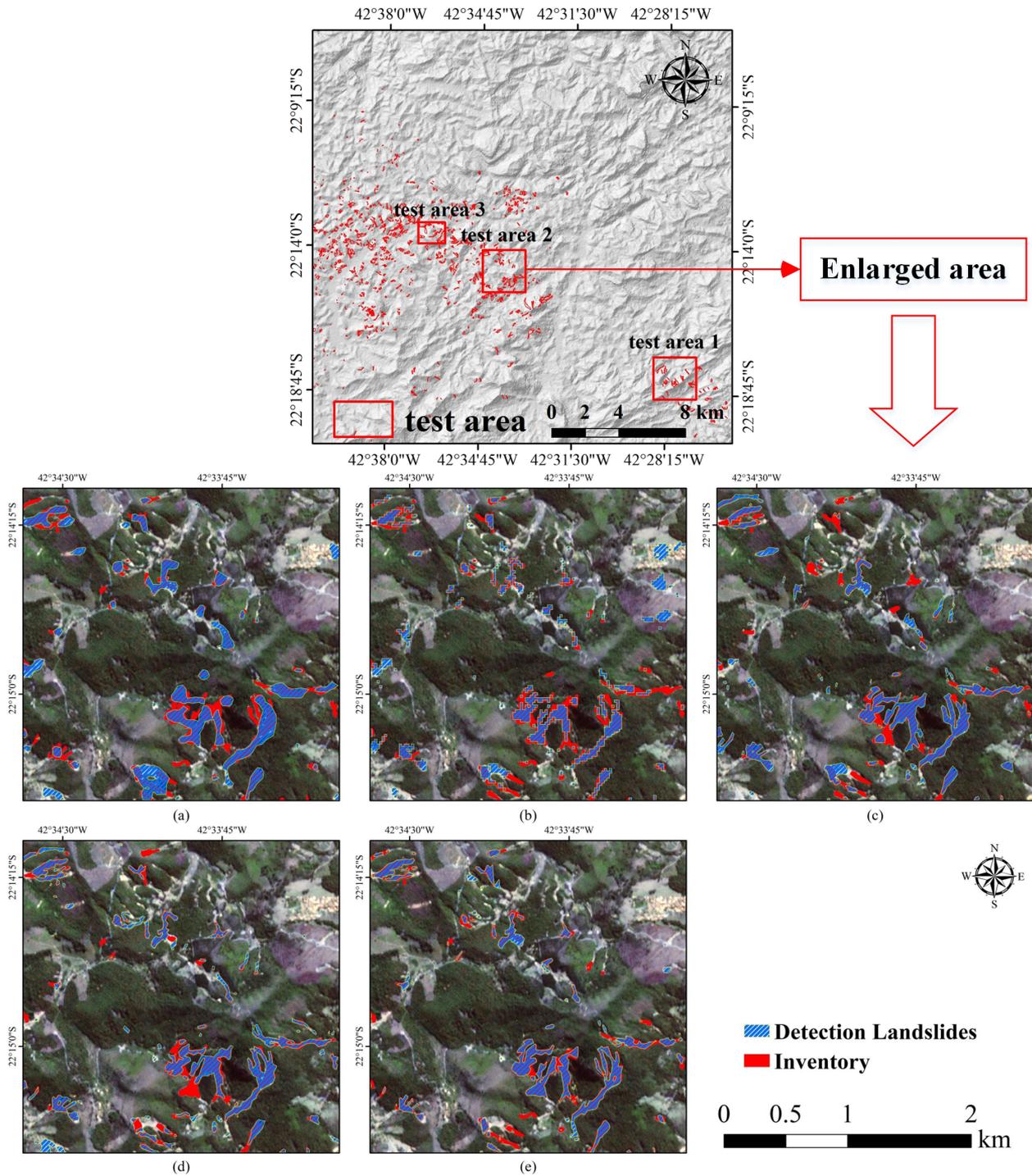
Fig. 7.　Landslide inventory mapping results by different DCNN methods. (a) FCN. (b) DeepLabv3+. (c) U-Net. (d) DU-Net. (e) FCDU-Net.

where $m = 1$ in this article and $m + 1$ is the number of classes. P_{ij} represents the number of real i, but predicted to be j.

IV. RESULTS

A. Experimental Settings

To illustrate the performance of FCDU-Net in landslide inventory mapping, four deep CNNs were compared, including FCN, DeepLabv3+, U-Net, and DU-Net. All deep CNNs were performed on Linux platform using Tensorflow-based Keras framework, and the training and testing process was carried out on a deep learning machine with 64G CPU memory. Calculations were performed using two NVIDIA GeForce RTX 2080Ti GPUs (12 GB RAM) in the NVIDIA CUDA Toolkit 10.1 environment. Experiments all restarted training the network without involving any pretrained model weights. The commonly
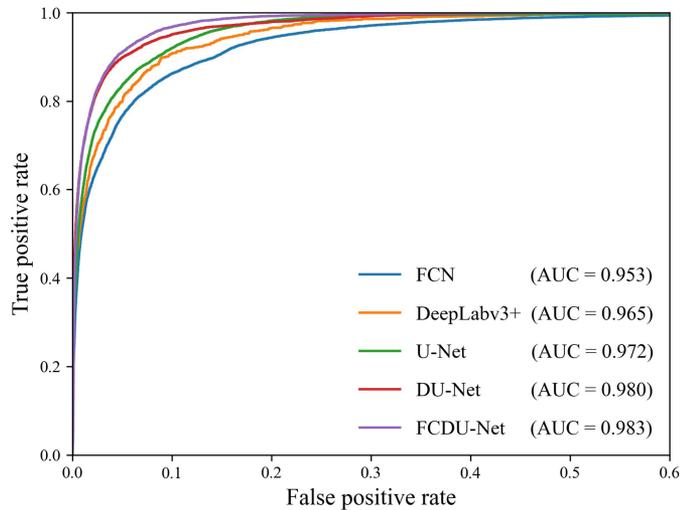
Fig. 8. ROC curves of different methods using the test set.

used Adam optimizer [67] was chosen to update the hyperparameters in the network. Through extensive experiments, 0.001 was used as the learning rate. To avoid overfitting, an early-stopping measure is used to save network parameters when the validation loss curve reaches a minimum. During the training phase, a maximum of 200 training epochs were conducted. To balance training time and model inference performance, a batch size of 64 was chosen. All experimental hyperparameters, training data, and other variables were set identically.

B. Landslide Inventory Mapping Results

All the mentioned deep learning methods were trained using the training dataset, and tested on three subareas of the study area. Referring to previous articles and landslide distribution, all methods used an input window size of 128 × 128. Fig. 7 shows landslide detection maps for the five DCNN methods. It is clear from Fig. 7 that FCN and DeepLabv3+ produced more omission errors in the upper left and lower right corners of the study area, respectively. In contrast, the DU-Net model had fewer false positives and false negatives, compared to U-Net. Fig. 7(e) shows that FCDU-Net has a good recognition for small landslides that are indistinguishable to the naked eye, degree. The landslide boundaries inferred by FCDU-Net match the true distribution of landslides better than U-Net and DU-Net. The visual inspection results show that FCDU-Net has the best prediction performance.

Fig. 8 plots receiver operating characteristic curves of the five models. The area under the curve (AUC) is a quantitative measure of the accuracy of each method in distinguishing categories. In this article, the two categories are landslide areas and nonlandslide areas. Theoretically, the minimum value of AUC is 0 and the maximum is 1. Moreover, the closer the AUC value is to 1, the better the model's ability to identify samples of different classes. As can be seen from Fig. 8, the AUC values of all the models are above 0.95, indicating that DCNN can exhibit very satisfactory predictive ability. Furthermore, the FCDU-Net

method outperforms the other methods with the highest AUC value of 0.983.

Table III lists the detection accuracy evaluations of different methods for Precision, Recall, F1, mIoU, Kappa, and OA, demonstrating the quantitative assessment accuracy of each method based on the test dataset. It can be observed that FCDU-Net has the highest mIoU at 83.19%, which is about 7% higher than U-Net (76.47%). In addition, FCDU-Net achieves the highest Precision, Recall, F1, Kappa, and OA values. Except for Recall, the FCN model performs worse than the other three models in terms of all other statistical metrics.

C. Parametric Analysis

1) Influence of Convolutional Input Window Size: To analyze the influence of convolutional input window size, we conducted landslide detection experiments using FCDU-Net with different convolutional input window sizes. Specifically, the size of 128 × 128 was used as the large input window size of the model, while the sizes of 64 × 64 and 96 × 96 were considered as two different versions of the small window. On this basis, 3060, 1880, and 1296 training samples of size 64 × 64, 96 × 96, and 128 × 128 were derived, respectively. Fig. 9 shows landslide detection maps for two landslide subareas in the study area using FCDU-Net with three different input window sizes. It can be seen that different convolutional input window sizes can reflect the geological environmental conditions of different scales of landslides. As a result, the model can extract different spatial contextual information during training.

To quantitatively analyze the usability of the FCDU-Net method with three different input window sizes, a confusion matrix is used to calculate the four evaluation measures mentioned in Section III-B, as listed in Table IV. The size from 64 × 64 to 96 × 96 improves Precision, but results in lower Recall, mainly because the larger input window has a negative impact on the classification of random points distributed in the landslide polygon. Some randomly distributed points will be close to the boundary of the landslide area. As a result, the nonlandslide area will increase with the size of the input window. However, the lowest F1 and Kappa were obtained based on the size of 96 × 96. In short, FCDU-Net with 128 × 128 can perform better on all evaluation metrics compared to other sizes.

2) Influence of the Relief-F Feature Selection: To verify the effectiveness of DU-Net integrated with Relief-F, the feature importance (FI) of each auxiliary feature is shown in Fig. 10. The higher the FI value, the more important the conditioning factor is to modeling. As can be observed from Fig. 10, the NDVI factor reached the highest FI value. The correlation with landslide occurrence raises as the value increases. Therefore, the conditioning factors that obtained the first three FI values were selected for subsequent modeling.

In this section, we compare the landslide detection performance of DU-Net and FCDU-Net. All experiments used only an input window size of 128 × 128. The test area covers a variety of land use types such as roads, shrubs, woodlands, and houses, and soil landslides mainly occur in woodlands and grasslands. Each landslide varies in shape and size, and some landslides have
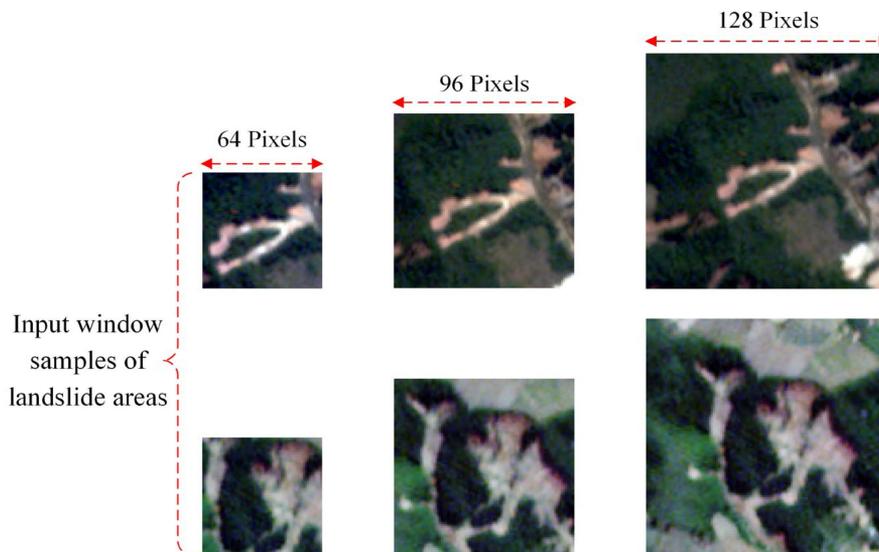
Fig. 9.   Three convolution input window sizes from two landslide subareas.
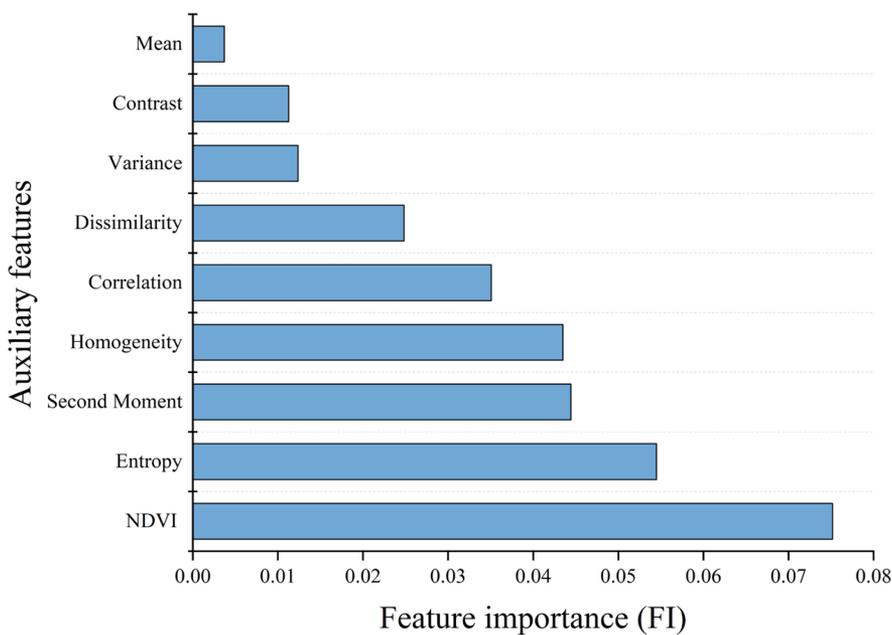
Fig. 10.   Feature importance of each auxiliary feature factor using Relief-F.

TABLE III
DETECTION ACCURACY ASSESSMENTS BY DIFFERENT METHODS IN ALL TEST AREAS

Methods	mIoU	Precision	Recall	F1	Kappa	OA
FCN	69.41%	54.92%	68.07%	60.79%	58.34%	95.33%
DeepLabv3+	72.23%	68.96%	61.39%	64.96%	63.11%	96.47%
U-Net	76.47%	72.41%	71.24%	71.82%	70.25%	97.02%
DU-Net	78.27%	76.63%	72.38%	74.44%	73.05%	97.35%
FCDU-Net	**83.19%**	**88.87%**	**74.70%**	**81.17%**	**80.21%**	**98.15%**

The highest accuracies are indicated in bold in each statistical measure.
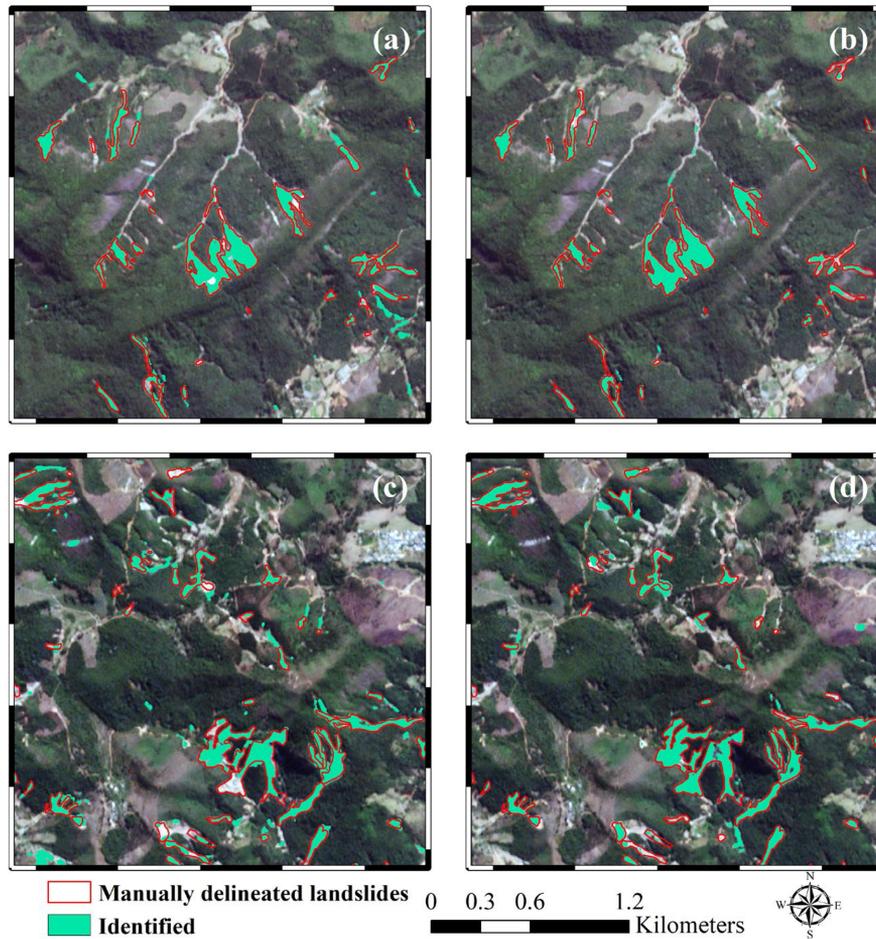
Fig. 11. Landslide detection maps of two test areas obtained by DU-Net and FCDU-Net. (a) DU-Net on test area 1. (b) FCDU-Net on test area 1. (c) DU-Net on test area 2. (d) FCDU-Net on test area 2.
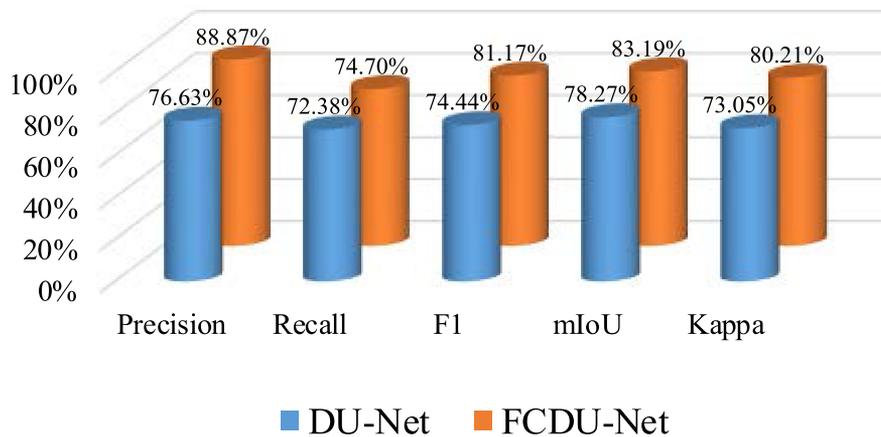
Fig. 12. Performance of DU-Net and FCDU-Net in terms of statistical measures.

spectral signatures similar to bare soil, posing a great challenge to landslide detection. Fig. 11 shows the landslide inventory mapping results for DU-Net and FCDU-Net. It can be seen from Fig. 11(a) and (c) that most of the large landslides can be distinguished from the background by DU-Net, and the boundaries of some slender landslides can also be accurately identified.

However, the results reveal some incomplete and overdetection of small-scale landslides for two reasons. One is the serious imbalance between landslide and nonlandslide samples, and the other is due to the impact of imaging techniques or vegetation coverage, some small landslides cannot be detected by remote sensing images alone. From Fig. 11(b) and (d), we can observe
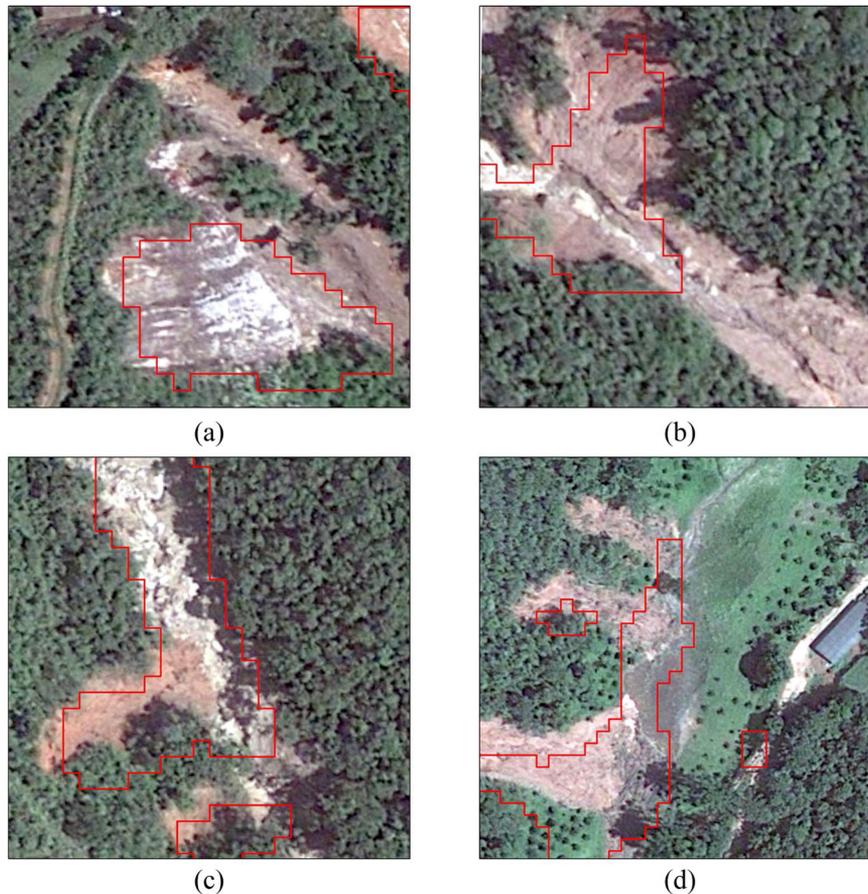
Fig. 13. Very-high-resolution images from Google Earth overlapped with the landslide detection results by FCDU-Net. Figures (a)–(d) denote landslides that were not recorded in the previous landslide inventory data.

TABLE IV
PERFORMANCE OF FCDU-NET USING DIFFERENT INPUT WINDOW SIZES

Sizes	Recall	Precision	F1	Kappa
64×64	73.32%	71.94%	72.62%	71.07%
96×96	65.79%	77.90%	71.34%	69.87%
128×128	74.70%	88.87%	81.17%	80.21%

that the entire landslide map in both figures is satisfactory, reducing the number of overidentified and incorrectly identified landslides, and most of the landslide scars are consistent with the boundaries of the true inventory data. shows the performance of DU-Net and FCDU-Net in terms of five statistical measures Fig. 12previously mentioned in Section III-B. As can be seen from this figure, FCDU-Net performs more balanced in all test domains, and FCDU-Net is better than DU-Net in terms of the statistical measures. Specifically, the Precision of FCDU-Net is 88.87%, which is more than 12% higher than that of DU-Net (76.63%), and the F1, mIoU, and Kappa of FCDU-Net are all 5% higher than that of DU-Net.

V. DISCUSSION

Landslide inventory mapping is the key to emergency rescue and landslide disaster loss assessment [68]. Meanwhile, improving the efficiency of spatial prediction is also important for technical experts to obtain detailed landslide disaster distribution. Therefore, this article compares several DCNNs such as FCN, DeepLabv3+, U-Net, and DU-Net for landslide detection. Furthermore, we evaluate the impact of using different input window sizes on prediction accuracy and ensemble performance of DU-Net and Relief-F.

Generally speaking, high-resolution images have fewer available spectral bands, the spectral distribution of objects in the image varies greatly, and the spectra of different objects overlap each other. However, the geometry and textural information of the object displayed in the image are more obvious. Landslides are usually covered with vegetation. Therefore, it is necessary to combine spectral features with other auxiliary features for landslide detection. Based on the results by Relief-F, NDVI, texture features of Entropy, and Second Moment are the eigenfactors because they represent the characteristic properties and structural information of the landslide. The experimental results show that FCDU-Net achieved higher prediction accuracy than

DU-Net alone because the filtering factor of Relief-F has a positive impact on the detection model and reduces the prediction bias of the model.

To further improve the effectiveness of the proposed method, we apply FCDU-Net to other high-resolution images to identify potential or unlabeled landslide hazards. As shown in Fig. 13, FCDU-Net can manually infer unlabeled landslides in different test areas, which further proves that it can perform well in landslide detection. Furthermore, DCNNs play an active role in updating or supplementing existing landslide databases as they can detect areas that were previously neglected.

In previous articles, some scholars began to explore the application of deep learning or transfer learning in landslide inventory mapping [69]–[74]. A comprehensive quantitative comparison is not conducted herein due to differences in image size and geographic location. They used OBIE, change detection, or long time series detection algorithms, and included several topographic factors (e.g., DEM and its derivatives). For example, Lu and Ma [70] combined transfer learning with OBIE to achieve edge extraction for large landslides. Shi *et al.* [72] proposed a new integrated method combining DCNN and change detection for efficient landslide mapping in the Hong Kong area. Ghorbanzadeh *et al.* [74] showed that topographic factors such as elevation, slope, and plane curvature can improve the prediction performance of deep learning networks. In contrast to these articles, we did not use high-resolution DEM and any postprocessing methods. The excellent performance demonstrates the effectiveness and robustness of the proposed method.

Through comparative analysis, the application of DCNN methods has greater advantages. Specifically, they require less manual fine-tuning than traditional methods and can be easily used in new domains that require only a small amount of training data. Moreover, the trained DU-Net will allow continuous processing and segmentation of high-resolution images, driving the development of software for continuous detection of new landslides.

One of the aims of this article is to explore the effect of landslides at different scales on the detection accuracy of deep learning methods, so we have adopted the strategy of different convolutional input window sizes for the proposed FCDU-Net method. The experimental results demonstrate that FCDU-Net with the convolutional input window size of 128 × 128 can perform the best. It can found that multitemporal or long time series data may be required to train the proposed method to well remove the seasonal changes in mountainous areas. However, by integrating NDVI and eight GLCM texture features with spectral bands in the original image, this article proposes a strategy for landslide feature extraction in high-resolution remote sensing images in southern latitudes in winter, and we will investigate the generalizability of this method in the future.

VI. CONCLUSION

In this article, we propose an effective FCDU-Net semantic segmentation method. The FCDU-Net training model has the advantages of fewer parameters and lower error probability, and can be put into use quickly. To validate the effectiveness of the proposed method, we apply it to landslide detection in remote areas of the Novo Fribourg Mountains. In addition, we also analyze the influence of different convolution input window sizes on the landslide detection results. The experimental results demonstrate that FCDU-Net using an input window size of 128 × 128 can produce the best performance and achieve higher prediction results than classical methods of FCN, DeepLabv3+, U-Net, and DU-Net. Therefore, FCDU-Net has greater application potential in the landslide inventory dynamic mapping system, which will expand the landslide database and promote the subsequent studies of landslide susceptibility mapping. In the future, we will increase landslide samples from other rainfall events and continue to train the proposed model to enhance its robustness to other disaster areas.

REFERENCES

[1] O. Ghorbanzadeh, T. Blaschke, K. Gholamnia, S. R. Meena, D. Tiede, and J. Aryal, "Evaluation of different machine learning methods and deep-learning convolutional neural networks for landslide detection," *Remote Sens.*, vol. 11, no. 2, 2019, Art. no. 196.

[2] 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.

[3] A. Mohan, A. K. Singh, B. Kumar, and R. Dwivedi, "Review on remote sensing methods for landslide detection using machine and deep learning," *Trans. Emerg. Telecommun. Technol.*, vol. 32, no. 7, 2021, Art. no. e3998.

[4] F. Raspini *et al.*, "Updated landslide inventory of the area between the Furiano and Rosmarino creeks (Sicily, Italy)," *J. Maps*, vol. 12, no. 5, pp. 1010–1019, 2016.

[5] G. Danneels, E. Pirard, and H.-B. Havenith, "Automatic landslide detection from remote sensing images using supervised classification methods," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2007, pp. 3014–3017.

[6] P. Lu, A. Stumpf, N. Kerle, and N. Casagli, "Object-oriented change detection for landslide rapid mapping," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 4, pp. 701–705, Jul. 2011.

[7] J. Dou *et al.*, "An integrated artificial neural network model for the landslide susceptibility assessment of Osado Island, Japan," *Natural Hazards*, vol. 78, no. 3, pp. 1749–1776, 2015.

[8] C. Lira *et al.*, "Automatic detection of Landslide features with remote sensing techniques: Application to Madeira Island," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2011, pp. 1997–2000.

[9] J. E. Nichol, A. Shaker, and M.-S. Wong, "Application of high-resolution stereo satellite images to detailed landslide hazard assessment," *Geomorphology*, vol. 76, no. 1/2, pp. 68–75, 2006.

[10] P. Tarolli, G. Sofia, and G. Dalla Fontana, "Geomorphic features extraction from high-resolution topography: Landslide crowns and bank erosion," *Natural Hazards*, vol. 61, no. 1, pp. 65–83, 2012.

[11] Y. Zhang, H. Ran, Y. Peng, and Y. Zheng, "Automatic prediction of landslides over InSAR techniques and differential detection using high-resolution remote sensing images: Application to Jinsha river," in *Proc. China High Resolution Earth Observ. Conf.*, 2019, pp. 471–486.

[12] C. Zhao, Y. Kang, Q. Zhang, Z. Lu, and B. Li, "Landslide identification and monitoring along the Jinsha river catchment (Wudongde Reservoir area), China, using the InSAR method," *Remote Sens.*, vol. 10, no. 7, 2018, Art. no. 993.

[13] E. L. Harp, D. K. Keefer, H. P. Sato, and H. Yagi, "Landslide inventories: The essential part of seismic landslide hazard analyses," *Eng. Geol.*, vol. 122, no. 1/2, pp. 9–21, 2011.

[14] D. Hölbling, B. Friedl, and C. Eisank, "An object-based approach for semi-automated landslide change detection and attribution of changes to landslide classes in northern Taiwan," *Earth Sci. Informat.*, vol. 8, no. 2, pp. 327–335, 2015.

[15] C. Zhong *et al.*, "Landslide mapping with remote sensing: Challenges and opportunities," *Int. J. Remote Sens.*, vol. 41, no. 4, pp. 1555–1581, 2020.

[16] A. Borghuis, K. Chang, and H. Lee, "Comparison between automated and manual mapping of Typhoon-triggered landslides from SPOT-5 imagery," *Int. J. Remote Sens.*, vol. 28, no. 8, pp. 1843–1856, 2007.

[17] E. V. Marcelino, A. R. Formaggio, and E. E. Maeda, "Landslide inventory using image fusion techniques in Brazil," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 11, no. 3, pp. 181–191, 2009.

[18] R. Behling, S. Roessner, H. Kaufmann, and B. Kleinschmit, "Automated spatiotemporal landslide mapping over large areas using rapideye time series data," *Remote Sens.*, vol. 6, no. 9, pp. 8026–8055, 2014.

[19] K. S. Cheng, C. Wei, and S. Chang, "Locating landslides using multi-temporal satellite images," *Adv. Space Res.*, vol. 33, no. 3, pp. 296–301, 2004.

[20] J. Hervás, J. I. Barredo, P. L. Rosin, A. Pasuto, F. Mantovani, and S. Silvano, "Monitoring landslides from optical remotely sensed imagery: The case history of Tessina landslide, Italy," *Geomorphology*, vol. 54, no. 1–2, pp. 63–75, 2003.

[21] J. Nichol and M. S. Wong, "Satellite remote sensing for detailed landslide inventories using change detection and image fusion," *Int. J. Remote Sens.*, vol. 26, no. 9, pp. 1913–1926, 2005.

[22] A. A. Othman and R. Gloaguen, "Automatic extraction and size distribution of landslides in Kurdistan Region, NE Iraq," *Remote Sens.*, vol. 5, no. 5, pp. 2389–2410, 2013.

[23] N. W. Park and K. H. Chi, "Quantitative assessment of landslide susceptibility using high-resolution remote sensing data and a generalized additive model," *Int. J. Remote Sens.*, vol. 29, no. 1, pp. 247–264, 2008.

[24] W. Yang, M. Wang, and P. Shi, "Using MODIS NDVI time series to identify geographic patterns of landslides in vegetated regions," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 4, pp. 707–710, Jul. 2012.

[25] W. Zhang, J. Lin, J. Peng, and Q. Lu, "Estimating Wenchuan earthquake induced landslides based on remote sensing," *Int. J. Remote Sens.*, vol. 31, no. 13, pp. 3495–3508, 2010.

[26] 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. 1/2, pp. 24–36, 2010.

[27] T. R. Martha, N. Kerle, C. J. Van Westen, V. Jetten, and K. V. Kumar, "Object-oriented analysis of multi-temporal panchromatic images for creation of historical landslide inventories," *ISPRS J. Photogramm. Remote Sens.*, vol. 67, pp. 105–119, 2012.

[28] A. Stumpf and N. Kerle, "Object-oriented mapping of landslides using random forests," *Remote Sens. Environ.*, vol. 115, no. 10, pp. 2564–2577, 2011.

[29] Y. Martin and S. Franklin, "Classification of soil-and bedrock-dominated landslides in British Columbia using segmentation of satellite imagery and DEM data," *Int. J. Remote Sens.*, vol. 26, no. 7, pp. 1505–1509, 2005.

[30] J.-Y. Rau, J.-P. Jhan, and R.-J. Rau, "Semiautomatic object-oriented landslide recognition scheme from multisensor optical imagery and DEM," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 2, pp. 1336–1349, Feb. 2013.

[31] S. Heleno, M. Matias, P. Pina, and A. J. Sousa, "Semiautomated object-based classification of rain-induced landslides with VHR multispectral images on Madeira Island," *Natural Hazards Earth Syst. Sci.*, vol. 16, no. 4, pp. 1035–1048, 2016.

[32] 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.

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

[34] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, 2017.

[35] X. X. Zhu *et al.*, "Deep learning in remote sensing: A comprehensive review and list of resources," *IEEE Geosci. Remote Sens. Mag.*, vol. 5, no. 4, pp. 8–36, Dec. 2017.

[36] D. He, Q. Shi, X. Liu, Y. Zhong, and X. Zhang, "Deep subpixel mapping based on semantic information modulated network for urban land use mapping," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 12, pp. 10628–10646, Dec. 2021.

[37] H. Guo, Q. Shi, A. Marinoni, B. Du, and L. Zhang, "Deep building footprint update network: A semi-supervised method for updating existing building footprint from bi-temporal remote sensing images," *Remote Sens. Environ.*, vol. 264, 2021, Art. no. 112589.

[38] D. He, Y. Zhong, X. Wang, and L. Zhang, "Deep convolutional neural network framework for subpixel mapping," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 11, pp. 9518–9539, Nov. 2020.

[39] A. Ding, Q. Zhang, X. Zhou, and B. Dai, "Automatic recognition of landslide based on CNN and texture change detection," in *Proc. 31st Youth Acad. Annu. Conf. Chin. Assoc. Automat.*, 2016, pp. 444–448.

[40] H. Yu, Y. Ma, L. Wang, Y. Zhai, and X. Wang, "A landslide intelligent detection method based on CNN and RSG_R," in *Proc. IEEE Int. Conf. Mechatron. Automat.*, 2017, pp. 40–44.

[41] Y. Yi and W. Zhang, "A new deep-learning-based approach for earthquake-triggered landslide detection from single-temporal Rapideye satellite imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 6166–6176, Oct. 2020.

[42] S. Qin *et al.*, "Landslide detection from open satellite imagery using distant domain transfer learning," *Remote Sens.*, vol. 13, no. 17, 2021, Art. no. 3383.

[43] A. S. Avelar, A. L. C. Netto, W. A. Lacerda, L. B. Becker, and M. B. Mendonça, "Mechanisms of the recent catastrophic landslides in the mountainous range of Rio de Janeiro, Brazil," in *Landslide Science and Practice*. Berlin, Germany: Springer, 2013, pp. 265–270.

[44] A. L. C. Netto *et al.*, "January 2011: The extreme landslide disaster in Brazil," in *Landslide Science and Practice*. Berlin, Germany: Springer, 2013, pp. 377–384.

[45] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 3431–3440.

[46] W. Sun and R. Wang, "Fully convolutional networks for semantic segmentation of very high resolution remotely sensed images combined with DSM," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 3, pp. 474–478, Mar. 2018.

[47] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Interv.*, 2015, pp. 234–241.

[48] V. Iglovikov and A. Shvets, "Ternausnet: U-net with vgg11 encoder pretrained on imagenet for image segmentation," 2018, *arXiv:1801.05746.*

[49] M. Volpi and D. Tuia, "Dense semantic labeling of subdecimeter resolution images with convolutional neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 2, pp. 881–893, Feb. 2016.

[50] X. Wei *et al.*, "Semantic pixel labelling in remote sensing images using a deep convolutional encoder–decoder model," *Remote Sens. Lett.*, vol. 9, no. 3, pp. 199–208, 2018.

[51] L. P. Soares, H. C. Dias, and C. H. Grohmann, "Landslide segmentation with u-net: Evaluating different sampling methods and patch sizes," 2020, *arXiv:2007.06672.*

[52] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder–decoder with atrous separable convolution for semantic image segmentation," in *Proc. Eur. Conf. Comput. Vis.*, 2018, pp. 801–818.

[53] S. Mallat, *A Wavelet Tour of Signal Processing*. Amsterdam, The Netherlands: Elsevier, 1999.

[54] A. R. Choudhury, R. Vanguri, S. R. Jambawalikar, and P. Kumar, "Segmentation of brain tumors using deeplabv3+," in *Proc. Int. MICCAI Brainlesion Workshop*, 2018, pp. 154–167.

[55] L.-C. Chen, G. Papandreou, F. Schroff, and H. Adam, "Rethinking atrous convolution for semantic image segmentation," 2017, *arXiv:1706.05587.*

[56] X. Meng, P. Zhang, and D. Zhang, "Decision tree for online voltage stability margin assessment using C4.5 and Relief-F algorithms," *Energies*, vol. 13, no. 15, 2020, Art. no. 3824.

[57] K. Kira and L. A. Rendell, "A practical approach to feature selection," in *Proc. Mach. Learn.*, 1992, pp. 249–256.

[58] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "Unet++: A nested u-net architecture for medical image segmentation," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Berlin, Germany: Springer, 2018, pp. 3–11.

[59] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 4700–4708.

[60] Z. Wang, Y. Zhou, S. Wang, F. Wang, and Z. Xu, "House building extraction from high resolution remote sensing image based on IEU-Net," *J. Remote Sens.*, vol. 10, 2021, Art. no. 1768.

[61] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2015, pp. 1026–1034.

[62] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.

[63] J. Konečný, J. Liu, P. Richtárik, and M. Takáč, "Mini-batch semi-stochastic gradient descent in the proximal setting," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 2, pp. 242–255, Mar. 2015.

[64] X. Li, X. Sun, Y. Meng, J. Liang, F. Wu, and J. Li, "Dice loss for data-imbalanced NLP tasks," 2019, *arXiv:1911.02855*.

[65] S. Visa and A. Ralescu, "Issues in mining imbalanced data sets-a review paper," in *Proc. 16th Midwest Artif. Intell. Cogn. Sci. Conf.*, 2005, pp. 67–73.

[66] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-net: Fully convolutional neural networks for volumetric medical image segmentation," in *Proc. 4th Int. Conf. 3D Vis.*, 2016, pp. 565–571.

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

[68] P. Zhang, C. Xu, S. Ma, X. Shao, Y. Tian, and B. Wen, "Automatic extraction of seismic landslides in large areas with complex environments based on deep learning: An example of the 2018 Iburi Earthquake, Japan," *Remote Sens.*, vol. 12, no. 23, 2020, Art. no. 3992.

[69] W. Qi, M. Wei, W. Yang, C. Xu, and C. Ma, "Automatic mapping of landslides by the ResU-Net," *Remote Sens.*, vol. 12, no. 15, 2020, Art. no. 2487.

[70] H. Lu *et al.*, "Landslides information extraction using object-oriented image analysis paradigm based on deep learning and transfer learning," *Remote Sens.*, vol. 12, no. 5, 2020, Art. no. 752.

[71] L. Bragagnolo, L. R. Rezende, R. V. da Silva, and J. M. V. Grzybowski, "Convolutional neural networks applied to semantic segmentation of landslide scars," *Catena*, vol. 201, 2021, Art. no. 105189.

[72] W. Shi, M. Zhang, H. Ke, X. Fang, Z. Zhan, and S. Chen, "Landslide recognition by deep convolutional neural network and change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 6, pp. 4654–4672, Jun. 2021.

[73] P. Liu *et al.*, "A research on landslides automatic extraction model based on the improved mask R-CNN," *ISPRS Int. J. Geo-Inf.*, vol. 10, no. 3, 2021, Art. no. 168.

[74] O. Ghorbanzadeh, S. R. Meena, H. S. S. Abadi, S. Tavakoli Piralilou, L. Zhiyong, and T. Blaschke, "Landslide mapping using two main deep-learning convolution neural network (CNN) streams combined by the Dempster–Shafer (DS) model," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 452–463, Dec. 2020.

Guosen Xu received the B.E. degree in resource exploration engineering from Hebei GEO University, Shijiazhuang, China, in 2019. He is currently working toward the M.S. degree in earth exploration and information technology at the China University of Geosciences, Wuhan, China.

His research interests include landslide detection, deep learning, and remote sensing applications.

Yi Wang (Member, IEEE) received the B.S. degree in printing engineering and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2002 and 2007, respectively.

He is currently a Professor with the Institute of Geophysics and Geomatics, China University of Geosciences (CUG), Wuhan, China. He is the Head of the Department of Geoinformatics. His research interests include remote sensing technology and application, geological and ecological environment monitoring, and data intelligent processing.

Prof. Wang is a member of Geological Society of China and Chinese Association of Automation. In 2019, he was named CUG Outstanding Young Talent.
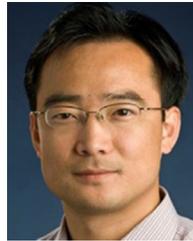
Lizhe Wang (Fellow, IEEE) received the B.E. and M.E. degrees in electrical engineering and automation from Tsinghua University, Beijing, China, in 1998 and 2001, respectively, and the D.E. degree in applied computer science from the University of Karlsruhe, Karlsruhe, Germany, in 2007.

He is currently the Dean of the School of Computer Science, China University of Geosciences, Wuhan, China. His research interests include remote sensing data processing, digital earth, and big data computing.

Dr. Wang is a Fellow of SPIE, IET, and BCS, and an Associate Editor for the *Remote Sensing* and *International Journal of Digital Earth*.

Lucas Pedrosa Soares received the B.S. degree in geology from the Institute of Geosciences, University of Brasília, Brasília, Brazil, in 2018. He is currently working toward the master's degree in remote sensing applications with deep learning with the Institute of Geosciences of the University of São Paulo, São Paulo, Brazil.

He is also a Researcher with Spatial Analysis and Modelling Lab. His research interests include landslide detection, landslide susceptibility mapping, artificial intelligence, deep and machine learning, and remote sensing applications with deep learning.

Carlos H. Grohmann received the B.S. degree in geology, and the M.Sc. and Ph.D. degrees in geotectonics from the University of Sao Paulo, Brazil, in 2002, 2004, and 2008, respectively.

In 2007, he was a Visiting Researcher with Kingston University London, Kingston-upon-Thames, U.K., and in 2009, he was employed as a Full-Time Professor with the Institute of Geosciences, University of Sao Paulo. He is currently an Associate Professor of geotechnologies with the Institute of Energy and Environment of the University of Sao Paulo, where he runs the Spatial Analysis and Modelling Lab. His research interests include digital terrain analysis (geomorphometry), remote sensing, applications of data collected from remotely piloted aircrafts in Earth and environmental sciences, LiDAR, and structure from motion-multiview stereo.

Dr. Grohmann has been a Board Member of the Brazilian Geological Society from 2014 to 2019 and of the International Society for Geomorphometry from 2015 to 2021. He is an Associate Editor for *Computers & Geosciences* and is the present Editor-in-Chief of the *Brazilian Journal of Geology*.