Recognition and Mapping of Landslide Using a Fully Convolutional DenseNet and Influencing Factors

Xiao Gao, Tao Chen, *Senior Member, IEEE*, Ruiqing Niu, and Antonio Plaza, *Fellow, IEEE*

***Abstract*—The recognition and mapping of landslide (RML) is an important task in hazard and risk research and can provide a scientific basis for the prevention and control of landslide disasters. However, traditional RML methods are inefficient, costly, and not intuitive. With the rapid development of computer vision, methods based on convolutional neural networks have attracted great attention due to their numerous advantages. However, problems such as insufficient feature extraction, excessive parameters, and slow model testing have restricted the development of this technology. This research proposes a new RML framework based on a new semantic segmentation network termed the fully convolutional DenseNet (FC-DenseNet). In this network, the features extracted from each layer are repeatedly used in a dense connection, and the parameters are controlled by a bottle-neck structure. Meanwhile, the structure of the encoder-decoder solves the problem of the slowness of model testing. Finally, the landslide influencing factors are added, which enriches the training data. To verify the effectiveness of the proposed method, we focused on several deep networks for comparison and analysis. The results show that FC-DenseNet can better recognize the boundary and interior of landslides, and there are fewer missing and excessive recognition results. The kappa value of the new method is 94.72% in Site 1, which is 6% and 4% higher than that of U-Net and ResU-Net, respectively, and 94.56% in Site 2, which is 6% and 3% higher than that of U-Net and ResU-Net, respectively, indicating that FC-DenseNet has great potential in RML applications.**

***Index Terms*—Fully convolutional DenseNet (FC-DenseNet), influencing factors, landslide recognition and mapping, remote sensing.**

I. Introduction

LANDSLIDE inventory maps are designed to record the location, distribution, and number and extent of landslides occurring in a given area [1], [2]. The recognition and mapping of landslide (RML) is an important task in hazard and risk studies and has attracted much attention from researchers in related fields in recent years. Landslide census information needs to be updated following large rainfall and seismic events. Information on past landslides is important for understanding the associated cause-effect relationships and predicting future hazards [3], [4]. The traditional RML method is mainly implemented by professionals in the field for detailed investigation and mapping, which makes it difficult to identify old landslides. Furthermore, field surveys are limited in their ability to accurately capture the extent of landslide boundaries due to sight lines being limited by slope, making them difficult and inefficient to implement on many mountain landslides. With the rapid development of remote sensing technology, it has been widely used for landslide recognition, mapping and susceptibility analysis due to its rapid, wide-coverage and all-weather data collection characteristics [5].

RML based on remote sensing is divided into three stages. Visual interpretation is the first step in extracting landslide information using remote sensing. Landslide recognition is based on features such as hue, texture, shape, location, and interrelationship of the image and relies on the experience and knowledge of the interpreter, with a large workload and long work cycles, making it difficult to meet emergency needs, and the results are sometimes unreliable [6], [7]. To improve efficiency and accuracy, many machine learning-based computer interpretation methods have been proposed for remote sensing image classification tasks, such as support vector machines (SVMs) [8], random forests (RFs) [9], and artificial neural networks (ANNs) [10]. These methods also make the landslide research further developed [11], [12]. The two main computer interpretation methods are pixel-based methods and object-oriented methods. The pixel-based method determines whether a landslide is a landslide according to the characteristics of the pixel value of a remote sensing image, which considers only the characteristics of a single pixel and does not consider the other attribute features of the landslide (shape, texture, spatial structure, interrelation, etc.) [13], [14]. However, the object-oriented recognition method is based on image segmentation, and the result depends on the selection of the segmentation scale, which is often determined by the texture characteristics of the image [15]. Some existing image segmentation algorithms cannot process complicated, large-scale remote sensing data, resulting in a low RML efficiency and a poor RML performance [16]. In recent years, deep learning has developed rapidly, and many

Manuscript received May 24, 2021; revised July 9, 2021; accepted July 27, 2021. Date of publication August 2, 2021; date of current version August 18, 2021. This work was supported in part by the National Natural Science Foundation of China under Grant 62071439 and Grant 61871259, in part by the Opening Foundation of Qilian Mountain National Park Research Center (Qinghai) under Grant GKQ2019-01, in part by the Opening Foundation of Beijing Key Laboratory of Urban Spatial Information Engineering, under Grant 20210209, and in part by the Opening Foundation of Geomatics Technology and Application Key Laboratory of Qinghai Province, under Grant QHDX-2019-01. *(Corresponding author: Tao Chen.)*

Xiao Gao and Ruiqing Niu are with the Institute of Geophysics and Geomatics, China University of Geosciences, Wuhan 430074, China (e-mail: gaoxiao@cug.edu.cn; niuruiqing@cug.edu.cn).

Tao Chen is with the Institute of Geophysics and Geomatics, China University of Geosciences, Wuhan 430074, China, and also with the Beijing Key Laboratory of Urban Spatial Information Engineering, Beijing 100038, China (e-mail: taochen@cug.edu.cn).

Antonio Plaza is with the Hyperspectral Computing Laboratory, Department of Technology of Computers and Communications, Escuela Politécnica, University of Extremadura, 10071 Cáceres, Spain (e-mail: aplaza@unex.es).

Digital Object Identifier 10.1109/JSTARS.2021.3101203

deep learning methods based on pixels and objects are used for RML development [17]–[20]. In contrast to traditional methods, the multilayer feedforward perceptron of convolutional neural networks (CNN) can automatically acquire effective feature representations of images, which allows these networks to identify the semantic features of landslide without having to manually compute complex landslide features [21].

In 1989, LeCun first trained a CNN model to classify handwritten digital images using a backpropagation approach [22], [23]. Subsequently, many scholars have successfully developed different CNN models, such as LeNet [24], AlexNet [25], GoogleNet [26], VGGNet [27], ResNet [28], and DenseNet [29], and all these traditional CNN models have achieved great success in image classification tasks. Since the first application of a CNN to landslide susceptibility analysis by Wang *et al.*, a wide variety of CNNs have been used for landslide recognition, mapping, and susceptibility analysis [30]–[34]. However, in these classification models, to classify a pixel, a block of images around a pixel is used as input into the CNN for training and prediction. This approach has several disadvantages. One is the great storage cost associated with this use. For example, the size of the image block used for each pixel is 15×15, and the window slides continuously. The window sliding pixel by pixel is sent to the network for classification, so the required storage space increases sharply according to the number and size of sliding windows. Second, the computational efficiency is low. The adjacent pixel blocks are basically repeated, and convolution is iteratively calculated for each pixel block, which is also repeated to a great extent. Third, the size of the pixel block limits the size of the sensing area. Generally, the selected pixel block is much smaller than the whole image and can only be used to extract some local features, which limits the classification performance. The application of semantic segmentation underwent a major breakthrough in 2014, when Long *et al*. at UC Berkeley proposed fully convolutional networks (FCNs) [35], which replaced the fully connected layer at the tail of traditional CNNs with a convolutional operation that traverses the entire network, and can be considered the beginning of the application of CNNs in the semantic segmentation field, enabling end-to-end, pixel-level classification of images. Since then, there have been numerous improvements to FCN storage to meet the needs of accuracy and efficiency, each with its own characteristics. Among them, more classical and outstanding performance are U-Net [36] and ResU-Net [37]. However, in general, most semantic segmentation models are still used with good results in tasks such as medical image segmentation and autonomous driving, and only a few models are used in landslide hazard research in last two years [38]–[40]. However, due to the limitations of the network structure, problems such as insufficient feature extraction capability, too many network parameters and easy overfitting of the model still exist, and the accuracy of semantic segmentation networks in RML applications needs to be further improved.

To address the above issues, this article proposes a new RML framework based on fully convolutional DenseNet (FC-DenseNet) that can accurately and reliably recognize and map landslides caused by heavy rainfall and earthquakes from monotemporal Landsat-8 OLI images. The FC-DenseNet supports feature reuse, enhanced feature transfer and reduced network parameters. In RML, the deeper FC-DenseNet not only exploits the potential of CNN, but also reduces computational cost due to its unique structure. In addition, FC-DenseNet is less prone to overfitting and is suitable for RML with unbalanced samples. The two main contributions of this article are as follows: 1) exploring the performance of FC-DenseNet in RML; and 2) adding landslide influencing factors to improve the performance of FC-DenseNet in RML.

The rest of this article is organized as follows. Section II provides a basic overview of the study area and the dataset used in the experiments. Section III presents the experimental steps of the new framework. Section IV analyses the experimental results in both qualitative and quantitative terms. Section V discusses the complexity of the model, the importance of the influencing factors and the limitations of this article, as well as future work. Section VI concludes the article.

TABLE I
OVERVIEW OF STUDY AREA, DISASTER EVENTS, AND LANDSAT-8 OLI IMAGES

	Site 1	**Site 2**
Size of study area	116.29km^2	44.26 km^2
Date of Event	before 2013	2017/08/08
Center location	31.00°N,110.64°E	33.20°N,103.82°E
Optical image	Landsat-8 OLI	
Spatial resolution	15m	15m
Acquired date	2013/09/15	2018/04/09
Spectral bands	Coastal, Blue, Green, Red, Near Infrared, Short-wave infrared 1 and Short-wave infrared 2	

II. STUDY SITES AND DATA

To verify the performance of the proposed method, two typical areas with severe landslide hazards, i.e., Zigui County in Hubei Province and Jiuzhaigou County in Sichuan Province, China, were selected for the experiments (see Table I). The environmental background of these two regions is diverse, and their geological structure is complex and active. The causes and development mechanisms of landslide also differ and are typical, representative, and sufficient to demonstrate the generalization capacity of the proposed method.

A. Study Area

Site 1: Zigui—Heavy rainfall from May to September

Site 1 is located in Zigui County of the Three Gorges Reservoir Area in the middle and lower reaches of the Yangtze River in southwestern Hubei Province, China, and belongs to the subtropical monsoon climate zone. Due to the influence of the monsoon, the rainy season in Site 1 lasts from May to September each year, and most of the yearly rainfall in the area is concentrated in these five months, exceeding 70% of the total annual rainfall. Site 1 is susceptible to land-sliding due to its hot and humid climate and frequent rainfall [38]. In addition, the topographic and geomorphological structure of Site 1 is complex, and all the

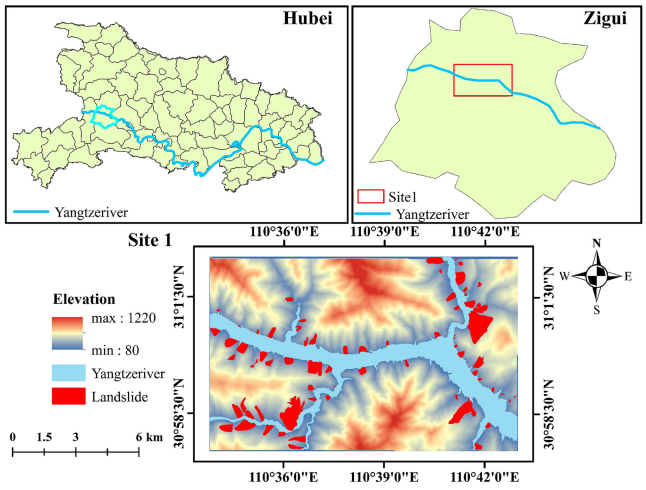


Fig. 1. Geo-location and topography of Site1.

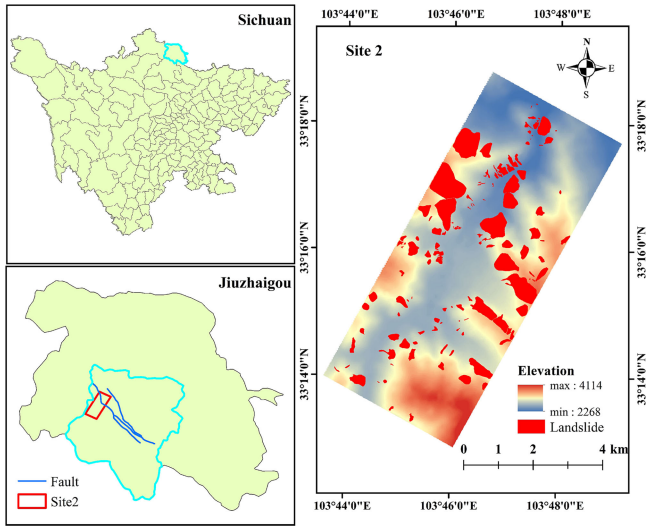


Fig. 2. Geo-location and topography of Site2.

study areas basically belong to mountainous areas. There are many kinds of stratum lithologies, and the geological structure is extremely unstable. These unstable geological structures often more easily breed many types of geological disasters, especially landslide disasters [41], [42]. In addition, the economic development of Site 1 has been relatively rapid in recent years, which has been followed by an enhancement in the scale and intensity of major human engineering activities. The action of human engineering activities, such as immigration and relocation, Three Gorges Reservoir construction, deforestation and mining, very readily induces slope instability and triggers landslide disasters [43], [44]. In short, due to its unique geographical environment, the study area is threatened by various inevitable geological disasters (especially landslide disasters). The size of Site 1 in this experiment is 616 × 976 pixels, the spatial resolution is 15 m, and the total area is about 135 km^2. It contains a total of 74 landslide points, with a landslide area of 6.82 km^2. The geo-location and topography of Site 1 are shown in Fig. 1.

Site 2: Jiuzhaigou—Ms 7.0 Earthquake on August 8, 2017

A magnitude 7.0 earthquake struck Jiuzhaigou County, Sichuan Province, on August 8, 2017. Jiuzhaigou is located in Jiuzhaigou County, Aba Tibetan and Qiang Autonomous Prefecture, Sichuan Province, in the transition zone between the eastern edge of the Qinghai-Tibet Plateau and the Sichuan Basin. It is a valley with a depth of more than 50 km. In this area, the terrain is high in the south and low in the north, with deep valleys and large differences in elevation. The northern edge of the region is only 2000 m above sea level in Jiuzhaigou, while the central peaks are above 4000 m and the southern edge is above 4500 m. The geological background of Jiuzhaigou is complex, with an extensive distribution of carbonate rocks, well-developed folds and faults, strong neotectonic movement and large crustal uplift, resulting in a variety of landforms [45], [46]. The epicenter of this earthquake is located in the town of Zhangzha in Jiuzhaigou County at 33.20°N, 103.82°E, with a depth of 20 km. The earthquake caused 25 deaths, 525 injuries, and six lost contacts; 176 492 people (including tourists) were affected, and 73 671 houses were damaged. The earthquake triggered more than 4800 landslides, causing substantial damage to scenic spots, roads and other facilities in Jiuzhaigou [47]–[49]. The Site 2's size is 733 × 585 pixels with the same spatial resolution with Site 1, and the area of Site 2 is about 96 km^2, which contains a total of 124 landslides and its area is 6.08 km^2. The geo-location and topography of Site 2 are shown in Fig. 2.

B. Optical Images and Landslide Influencing Factors

The optical remote sensing data used in this study mainly include Landsat-8 OLI images (2013/09/15, 2018/04/09) and the Global Digital Elevation Model.[1] A series of preprocessing steps (such as correction, fusion, etc.) should be performed on the multispectral remote sensing images to better interpret it. To obtain more detailed information on the surface vegetation and water body coverage in the study area, bands 4 and 5 of the Landsat-8 OLI image were used to calculate the normalized difference vegetation index (NDVI), and bands 3 and 6 were used to calculate the modified normalized difference water index (MNDWI) [50].

There is a close relationship between the occurrence of a landslide event and its influencing factors [51], [52]. In this study, taking into account the causes and distributions of landslide hazards across different study areas and their local topographical features, different landslide impact factors are selected for addition to the deep learning network as input data to improve the learning of landslide features and RML. According to previous research experience [53], the landslide influencing factors selected in this article are the elevation, aspect, slope, curvature, NDVI, MNDWI, distance to river, distance to fault, and lithology of the areas. Landslide influencing factors were extracted from geological maps, topographic maps, and multisource remote sensing images of the study area, while the landslide samples were annotated from remote sensing images combined with field surveys, which can be used as actual labels for later verification

[1][Online]. Available: http://www.gscloud.cn/

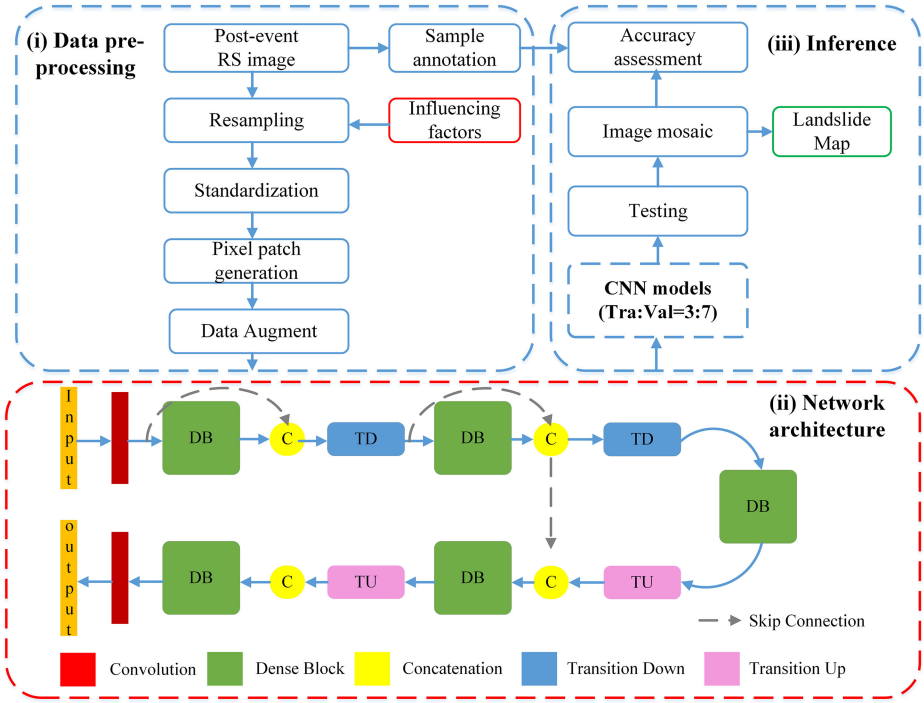


Fig. 3. Framework of landslide recognition and mapping method based on FC-DenseNet.

TABLE II
OVERVIEW OF LANDSLIDE INFLUENCE FACTORS IN THE TWO STUDY AREAS

Factors	Site 1	Site 2
Elevation (m)	(80, 1220)	(2268, 4121)
Slope (°)	(0, 62.69)	(0, 67.57)
Aspect	Flat, North, West, South, East, Southeast, Northwest, Southwest, Northeast	
Curvature (°/100m)	(-7.85, 5.47)	(-6.33, 5.67)
NDVI	(-0.60, 0.68)	(-0.08, 0.29)
MNDWI	(0.45, 0.99)	(0.35, 0.99)
Distance to river(m)	(0, 3063)	(160, 3032)
Distance to fault(m)	(0, 8754)	(0, 6860)
Lithology	Soft rock, Soft and hard, Hard rock	

of landslide recognition results. Table II shows detailed information on the landslide influencing factors of the two study areas and all landslide influencing factor maps were prepared using ArcGIS 10.3.

III. METHODOLOGY

As shown in Fig. 3, the whole experimental process consists of three parts. First, a data preprocessing workflow for image processing and landslide training dataset preparation is designed to generate training sample blocks. Data preprocessing mainly includes optical image processing, sample annotation and landslide influencing factors extraction. Training samples are generated through resampling, standardization, layer overlay, and cutting operations in this stage. The semantic segmentation network selected in this article is FC-DenseNet, which is applied to RML for the first time by adjusting the number of layers in DenseLayer. The number of DenseBlocks and the size of growth rate can make the network performance reach the optimal to the landslide characteristics better. The main points of this network are as follows: first, each layer of the network is directly connected with its all-front layer to improve the utilization of features; second, each layer in the network is designed to be very narrow, i.e., the number of convolution output channels is usually very small, which maximizes the utilization of resources and reduces the amount of computation needed without reducing the accuracy. Finally, in the inference process, the trained model is used to test the entire study area, and the landslide distribution map is drawn. In addition, to assess the performance of the method, we used six evaluation indicators for comparison with the remaining four common semantic segmentation networks: FCN-8s [35], SegNet [54], U-Net [36], and ResU-Net [37].

A. Data Preprocessing

The data preprocessing in this study mainly includes resampling, image standardization, pixel patch generation, and data augment (part i in Fig. 3). First, the optical remote sensing image data and factors data of the study area are cropped according to the selected vector range. As the resolution of the various types of geographical data varies, the input data need to be resampled to make the number of rows and columns consistent. The resampling method chosen in this article is a bilinear interpolation method that uses the image element values of four nearby points, which are given different weights according to their distance from the interpolation point, for linear interpolation [55]. Second, to facilitate the training of the deep learning network, the input data should be normalized. In this study, to make the input data of different dimensions have the same data distribution and further accelerate the convergence

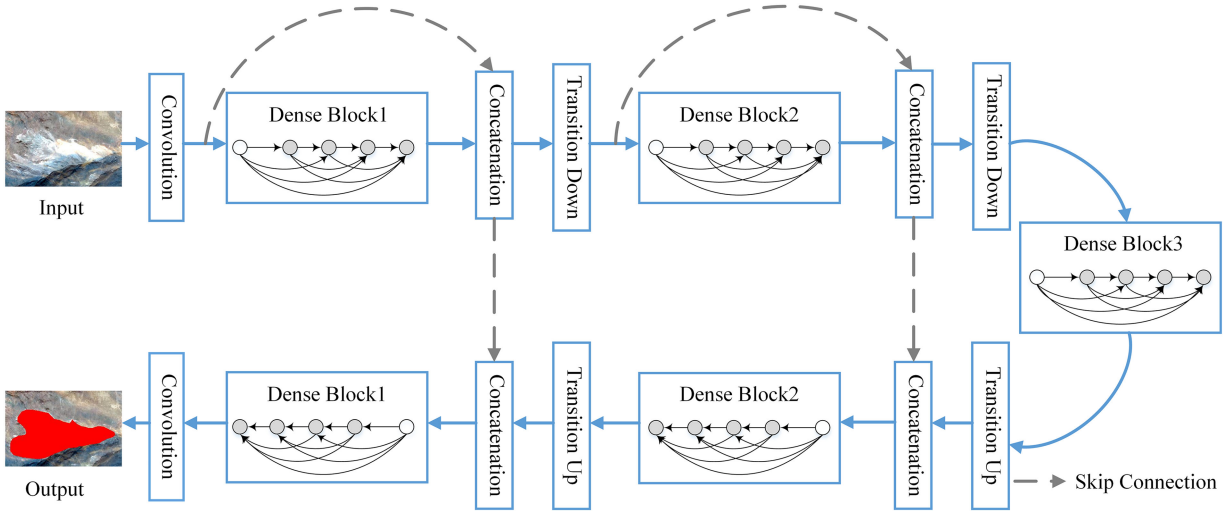


Fig. 4. Architecture of FC-DenseNet.

speed of the model, the Z-score method is used to normalize each channel of the input data and unify the data distribution so that the mean is 0 and the variance is 1 [56]. Finally, the layers are stacked as blocks of $m \times n \times c$ and fed into the deep learning network for training, where m and n represent the number of rows and columns, and c represents the number of channels.

Although deep learning has great advantages in landslide research [57], [58], it usually requires a large number of training samples. However, the representative high-quality training samples still have defects, and in the natural state, the ratio of landslide and nonlandslide is seriously unbalanced, so this problem is especially serious in remote sensing image landslide detection applications [59], [60]. Inspired by pioneering research in the field, this article employs a data augmentation strategy to generate more training samples [61], [62]. To boost the proportion of positive samples (e.g., landslide), this experiment randomly selected landslide points as seed pixels in the training area and scanned blocks of $m \times n \times c$ using a sliding window size of 64×64 and a step size of 32 to generate a sufficient number of sample blocks. These sample blocks were divided into training and validation data sets with a 3:7 ratio; 30% were used to train the deep learning model by learning the landslide features and 70% were used to validate the performance of the model.

B. Fully Convolutional DenseNet Architecture

FC-DenseNet uses DenseNet for semantic segmentation according to the FCN framework (part ii in Fig. 3). DenseNet breaks away from the stereotyped thinking of deepening the network layer (ResNet) and widening the network structure (inception) to improve the network's performance [28], [63]. From the point of view of features, through feature reuse and bypass setting, the number of network parameters is greatly reduced, and the vanishing gradient problem is alleviated to a certain extent. In deep learning networks, as the depth of the network deepens, the gradient disappearance problem becomes increasingly obvious, and many papers have proposed solutions to this problem, such as ResNet [28], highway networking [64], stochastic depth [65], and FractalNets [66]. Although the network structures of these algorithms are different, the core lies in the creation of short paths from early layers to later layers.

As shown in Fig. 4, the semantic segmentation network FC-DenseNet is used for the first time in this article to recognize landslide from postdisaster monotemporal Landsat-8 OLI imagery and landslide influencing factors. The network consists of three building blocks: dense block (DB), transition down (TD), and transition up (TU).

DB is the main component module of DenseNet and connects all layers directly while ensuring maximum information transfer between layers in the network. In traditional CNNs, if you have L layers, there will be L connections, but in DB, there will be L (L + 1)/2 connections (as shown in Fig. 5). In short, the input of each layer comes from the output of all previous layers. The formula is as follows:

$$xl = Fl\ ([x0,\ x1,\ \ldots,\ xl-1]) \tag{1}$$

where $[x0, x1 \ldots, xl - 1]$ refers to the dense connectivity of the feature maps in layers 0 to $l - 1$. $Fl\ (\cdot)$ denotes a nonlinear transformation function consisting of a combination of three operations, namely, batch normalization, rectified linear unit (ReLU) classification and 3×3 Convolution (Conv) [67], [68].

Each layer in the DB outputs k feature mappings after $Fl\ (\cdot)$, where k refers to the growth-rate in DenseNet and is a hyperparameter. Generally, a better performance can be obtained by using a smaller k. Assuming that the number of channels in the input layer feature map is $k0$, the number of channels in layer l is $k0 + k(l - 1)$. As the number of layers increases, although k is set to be very small, the output of DB will be very large, which results from feature reuse. To solve the problem of excessive input in the deep layer, the bottleneck layer can be used inside DB to reduce the amount of calculation, and that is 1×1 Conv is added to the original structure, which is called DenseNet-B (as shown in Fig. 6). A 1×1 Conv gets $4 * k$ feature maps. Its function is to compress the feature layers to improve the computational efficiency.

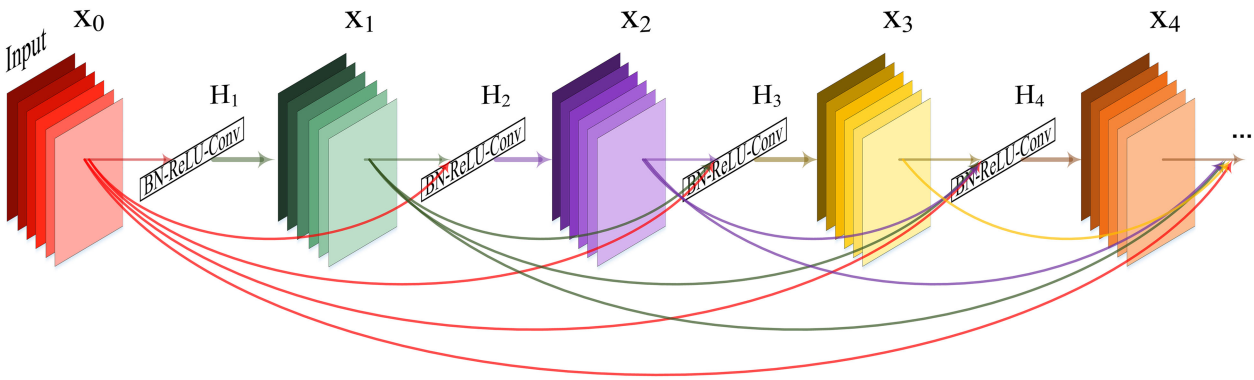


Fig. 5. Architecture of Dense block.

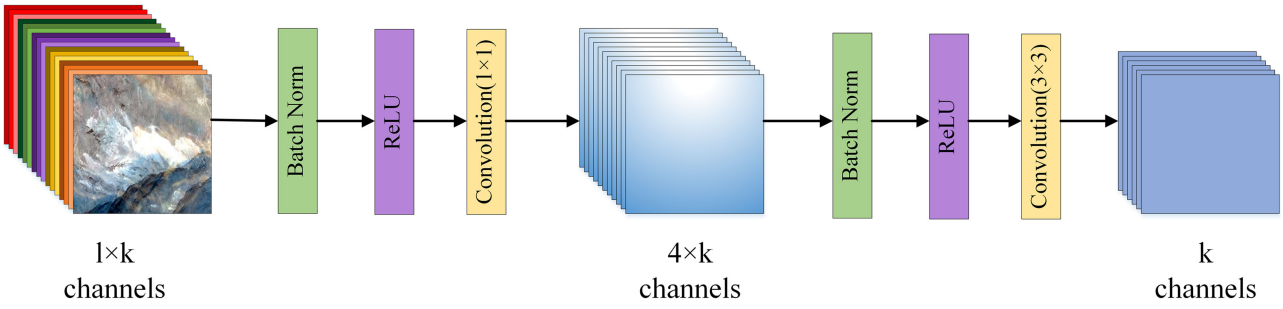


Fig. 6. Architecture of DenseNet-B.

TABLE III
THREE BUILDING BLOCKS OF FC-DENSENET

Dense Layer (k=12)	TD (θ=0.5)	TU
Batch Normalization ReLU	Batch Normalization ReLU	Batch Normalization ReLU
1×1 Convolution	1×1 Convolution	3×3 Convolution
Batch Normalization	Dropout p=0.2	2×2 Up Sampling
ReLU	2×2 Max Pooling	——
3×3 Convolution	——	
Dropout p=0.2	——	

TABLE IV
ARCHITECTURE DETAILS OF FC-DENSENET MODEL

Input, m=7/17
3×3 Convolution, m = 48
DB (4 layers) + TD, m = 72
DB (5 layers) + TD, m = 102
DB (6 layers), m = 174
TU + DB (5 layers), m = 351
TU + DB (4 layers), m = 367
1×1 Convolution, m = 1
Sigmoid

TD consists of a convolution layer with a core size of 1×1 and a maximum pool layer with a stride of 2, which can be used as a compression model. If the output feature layer of DB is m, the transition layer can be compressed by compression parameter θ. Usually, the value of θ is less than 1, which can compress the channel number of the feature map and reduce the network parameter. Here, another $1*1$ convolution operation and a compression parameter θ are added between each two DBs, which is called DenseNet-C. The network containing both bottleneck layers and compression layers is called DenseNet-BC, which not only has fewer parameters but also saves memory and avoids overfitting compared to a DenseNet of the same depth.

TU consists of a convolution layer with a core size of 3×3 and an up-sampling layer with a stride of 2. In Table III, we define the architecture details of DenseLayer, TD, and TU. In this table, dropout refers to randomly hiding part of the units with a certain probability (e.g., 0.2) each time during the training process, i.e., setting the activation function of that part of the neuron to 0, so as to avoid overfitting the model. In Table IV , we define the architecture and parameters of the FC-DenseNet network used in this study. The last layer is the convolution layer with a kernel size of 1×1, followed by sigmoid nonlinear mapping to provide each class distribution of each pixel.

C. Inference

In the inference stage (Part III in Fig. 3), the image is first cut into a series of overlapping patches, and then the class of each patch is predicted independently. Theoretically, this inference is not limited by the size of the area. Specifically, we use the sliding window algorithm to cut the test image into a series of 64×64 patches. It is worth noting that the sliding window algorithm here is different from the process of generating training samples. Here, the sliding window slides from the top left to the bottom right of the whole study area until it spreads throughout

TABLE V
WORKSTATION CONFIGURATION

Item	Parameter
System	Ubuntu 18.04.5 LTS
CPU	Intel Xeon(R) Silver 4210R CPU @2.4GHZ×40
GPU	GeForce RTX 3090/PCIe/SSE2
RAM	128GB

TABLE VI
LANDSLIDE RECOGNITION CONFUSION MATRIX

True Value / Predicted Value	Landslide	Non-landslide
Landslide	True Positive (TP)	False Positive (FP)
Non-landslide	False Negative (FN)	True Negative (TN)

the whole study area. Then, the patches are classified into binary labels (1 for landslide and 0 for nonlandslide) using the trained model. Finally, an image mosaic is carried out to completely cover the study area and obtain the landslide distribution map. In addition, to better compare the performance of different models, six different evaluation indexes are selected to quantitatively evaluate the results.

IV. RESULTS AND ACCURACY ASSESSMENT

A. Experiment Setting

The experiment environment for this manuscript is as shown in Table V. All the experimental codes are written in Python language under the framework of TensorFlow deep learning.

First, as landslide recognition is a binary classification problem, the binary cross-entropy loss function is chosen as the learning object during the training process. Second, the model uses the positive unit ball initialization method to initialize the weights of the network such that the sum of the weights of the inputs of each neuron is 1. In this way, it can effectively prevent the initialization of the weights from being too large to prevent the sigmoid activation function of the output layer from saturating too fast and the gradient from disappearing. Finally, the model uses the Adam optimizer to optimize the weights of the network to speed up the convergence of the model, and the initial learning rate is set to 0.0001. In addition, to prevent overfitting, the model adds a dropout layer to DB, with drop_rate set to 0.2. Due to computational limitations, the batch size is set to 2 after considering the tradeoff between time cost and model accuracy. After 100 iterations, the model was roughly stable. In all our experiments, we trained the network from scratch without involving any pretrained models.

B. Evaluation Metrics and Comparison Models

To quantitatively evaluate the performance of FC-DenseNet, this article selects six commonly used evaluation metrics based on the binary confusion matrix, namely, overall accuracy, precision, recall, F1-score, Kappa coefficient, and mean intersection over union (MIoU) index [69], [70]. Overall accuracy is commonly used evaluation index and is generally defined as the proportion of the number of samples with correct classification to the total number of samples. However, the index is not applicable in the dataset with unbalanced samples. In binary classification problems such as landslide recognition, precision, and recall are usually the most commonly used performance evaluation indicators. Precision refers to the proportion of positive samples correctly predicted to the total number of positive samples predicted. Recall refers to the proportion of positive samples correctly predicted to all positive samples. The formulas are as follows:

$$\text{Overall accuracy} = \frac{\text{TP}+\text{TN}}{\text{TP}+\text{FP}+\text{FN}+\text{TN}} \tag{2}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP}+\text{FP}} \tag{3}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP}+\text{FN}} \tag{4}$$

where TP, FP, FN, and TN come from the confusion matrix (Table VI). TP is the number of correctly predicted positive samples, TN is the number of correctly predicted negative samples, FN is the number of positive samples predicted to be negative, and FP is the number of negative samples predicted to be positive.

Since precision and recall are a pair of contradictory quantities, to better evaluate the performance of the classifier, the F1-score is generally used as the evaluation standard to measure the comprehensive performance of the classifier. The F1-score is the harmonic average of precision and recall. The higher the F1-score is, the better the performance of the classifier. The formula is as follows:

$$F1-\text{Score} = 2\times\frac{\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}}. \tag{5}$$

The Kappa coefficient is an index used to measure classification accuracy, which can quantitatively evaluate the consistency between classification results and real labels. Its value usually falls between 0 and 1, and can be divided into 5 groups to indicate different levels of consistency (0.0–0.2: slight, 0.2–0.4: fair, 0.4–0.6: moderate, 0.6–0.8: substantial, 0.8–1.0: almost perfect). Therefore, when its value is greater than 0.8, the consistency can be considered to be good [71], [72]. The formula is as follows:

$$\text{Kappa} = \frac{p_0-p_e}{1-p_e} \tag{6}$$

$$pe = \frac{(\text{TP}+\text{FN})(\text{TP}+\text{FP})+(\text{FP}+\text{TN})(\text{FN}+\text{TN})}{n^2} \tag{7}$$

where p_0 refers to the overall accuracy and n represents the total number of samples.

MIoU is the standard measure of semantic segmentation. It starts by finding the intersection ratio of each category on the two sets and then averaging the sets. In semantic segmentation, these two sets are the ground truth data and the predicted segmentation data. Ideally, the true value and the predicted value are exactly the same, i.e., the MIoU value is 1. In this study, the formula is

as follows:

$$\mathrm{MIoU} = \frac{1}{2}\left(\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}+\mathrm{FP}} + \frac{\mathrm{TN}}{\mathrm{TN}+\mathrm{FN}+\mathrm{FP}}\right). \quad (8)$$

To verify the effectiveness of FC-DenseNet, the two selected datasets are trained on five semantic segmentation networks, and the results are compared and analyzed. Among them, FCN is the pioneer of semantic segmentation, mainly for natural image semantic segmentation, multimodal medical image analysis, and multispectral satellite image segmentation. The encoder part of FCN is based on the AlexNet model. It transforms the fully connected layers in a traditional CNN into convolutional layers and upsamples the feature map of the last convolutional layer so that it recovers to the same size as the input image, thus producing a prediction for each pixel. The spatial information in the original input image is retained while the final pixel-by-pixel classification is performed on the upsampled feature map. However, the results obtained in this way are not sufficiently precise, and some details cannot be recovered. SegNet is a deep network for semantic segmentation of images proposed by researchers at Cambridge to address autonomous driving or intelligent robotics [54]. SegNet and FCN have similar ideas, but the technologies used by the encoder and decoder are different. Their encoder parts use the first 13 layers of the VGG16 convolutional network, with each encoder layer corresponding to a decoder layer that performs nonlinear upsampling using the pooling index calculated in the maximum pooling step of the corresponding encoder. The U-Net model was proposed in 2015 and originally used for medical image segmentation [36]. It is a symmetrical U-shaped structure, which is innovative in that there is usually a shortcut connection between the encoder and the decoder, fusing features at the same scale and therefore helping the decoder to better repair the details of the target. ResU-Net is based on an improvement of U-Net, which adds residual units to the upsampling and downsampling process, solving the problem that traditional U-Net cannot adequately extract shallow structural features [38]. U-Net and ResU-Net have achieved good results in remote sensing-related tasks of landslide disasters. Note that all of the above networks are trained from scratch using the same training dataset.

C. Accuracy Assessment and Comparison

To demonstrate the performance of FC-DenseNet in regional RML, we compared five semantic segmentation models in two separate study areas. Among them, FCN-8s is the simplest semantic segmentation model, which can simply extract the shallow semantic features of landslide for RML, but the decoding process will lose a large number of features, making the segmentation results not fine enough. SegNet deepens the network to extract deeper landslide features and the decoding part by adding pooling indexes to reduce the loss of features, but there seems to be no notable improvement in the results. The depth of the U-Net network is further deepened, and feature fusion is added to the encoder and decoder sections through a fully symmetrical structure, which allows the same scale of spatial features extracted by the encoder to be fused in the decoder section. ResU-Net takes into account the model overfitting problem caused by the deepening of the U-Net encoder part of the network, thus adding a kip connection that allows the fusion of shallow features while the network deepens further to extract deeper features. However, FC-DenseNet extracts features at each layer while fusing them with features extracted from all previous layers, compressing the parameters and reducing the computation through the bottleneck and TD settings while ensuring maximum information transfer between layers. Note that all the training data of the five models, as well as the super-parameters and variable settings of the network, are consistent.

Site 1: Zigui—Heavy rainfall from May to September

Fig. 7(a) and (b) shows the results for the FCN-8s and SegNet models, respectively. From these two figures, we can see that many small and medium landslides are not identified, and the identified landslide boundaries is not consistent with the actual range of landslides. From the recognition of the landslide boundary, the boundary of FCN-8s is too regular, showing a grid shape. This is due to the simple structure of the FCN-8s network and its incomplete landslide feature extraction. At the same time, this also highlights the shortcomings of pixel-based semantic segmentation methods of poor boundary segmentation and serious salt and pepper noise. In contrast, the boundary of SegNet is too smooth, which may be due to the function of adding a pool index to SegNet, but there are too many unrecognized landslides. In addition, both models have fewer overecognized landslides, but this is not a desired result in the practical application of landslide hazard recognition. Fig. 7(c) and (d) shows the results for the U-Net and ResU-Net models. It is clear from the plots that the recognition results of these two models are considerably better than those of Fig. 7(a) and (b). Overall, almost all of the large and medium-sized landslides are well identified, with only a few small landslides not being fully identified. In terms of boundaries, the boundaries of most landslides are identified, although they do not fit perfectly with the real landslide boundaries. Comparing the recognition results of several large landslides, it can be seen that ResU-Net produces better results for the interior of the landslide. However, both models still have some pretzel noise present.

Fig. 7(e) shows a plot of the results for the FC-DenseNet model. The FC-DenseNet recognition results give a closer comparison of small and medium-sized landslides, which shows that many small landslides that are not recognized by the ResU-Net model are recognized, and the FC-DenseNet model also outperforms ResU-Net in the recognition results within many small- to medium-sized landslides, which is a major indication of the improved accuracy of the model. In addition, it is evident that the FC-DenseNet model exhibits some over-recognition errors, which may be a result of misrecognition features such as bare rock and bare soil, which are similar to the shallow features of landslide in the form of landslide.

It is worth noting that the RML work described above was based on seven multispectral bands of postdisaster monotemporal Landsat-8 OLI imagery. To a large extent, RML was carried out by the spectral features of landslide as well as by their spatial features. However, landslide are gravity features on slopes, the

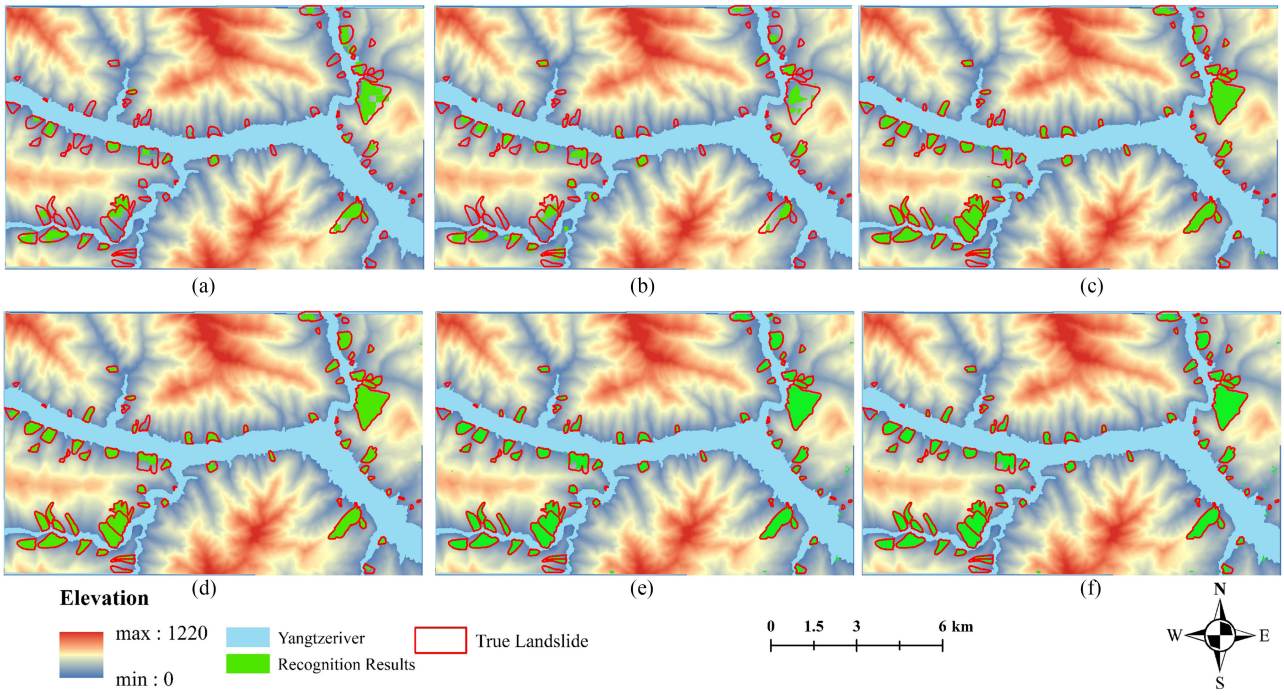


Fig. 7. RML results by using different methods in Site 1. (a) FCN. (b) SegNet. (c) U-Net. (d) ResU-Net. (e) FC-DenseNet. (f) FC-DenseNet + Influencing factors.

TABLE VII
SITE 1: QUALITATIVE EVALUATION OF SEGMENTATION RESULTS BY DIFFERENT METHODS

	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Kappa (%)	MIoU (%)
FCN-8s	95.83	82.87	37.74	51.86	50.00	65.37
SegNet	95.52	79.11	33.62	47.19	45.24	63.15
U-Net	98.82	95.49	84.19	89.48	88.87	89.87
ResU-Net	98.97	**96.84**	85.45	90.79	90.25	91.02
FC-DenseNet	99.08	92.72	91.82	92.27	91.78	92.34
FC-DenseNet +Factors	**99.42**	96.34	**93.77**	**95.04**	**94.72**	**94.96**

Note: Bold and underlined values mean the highest number of the corresponding indicator.

formation of which is closely related to topography, with certain slope and lithological conditions required [51]–[53]. Therefore, the inclusion of landslide influencing factors in the construction of landslide samples can not only distinguish between bare rock areas, roads, houses, and other features, thus improving the accuracy of RML, but it also be used to identify potential landslides. Fig. 7(e) and (f) both show the results of FC-DenseNet using different training samples. Fig. 7(e) shows the results of landslide recognition using remote sensing imagery alone, while Fig. 7(f) shows the results of landslide recognition using a combination of remote sensing imagery and landslide influencing factors. Overall, both of them have better identification results, but in terms of the more difficult identification of small- to medium-sized landslides, Fig. 7(f) has considerably few missed recognition and over-recognition instances than Fig. 7(e), resulting in better recognition results.

To quantitatively analyze the performance of the models, we calculated six evaluation factors based on the confusion matrix. As seen in Table VII, the kappa coefficients of both FCN-8s and SegNet are below 0.8, which means they do not meet the consistency requirement [71]. In practical applications, they can be used at most for landslide location but not for landslide boundary outlining [72]. In contrast, U-Net, ResU-Net and FC-DenseNet all obtained a better recognition accuracy. The overall accuracy of FC-DenseNet is 0.9908, the precision is 0.9272, the recall is 0.9182, the F1-score is 0.9227, the Kappa value is 0.9178, and the MIoU is 0.9234. The remaining evaluation metrics of FC-DenseNet are highest, excluding its precision. Compared to U-Net, F1-score, Kappa, and MIoU showed improvements of 2.97%, 2.91%, and 2.47%, respectively.

In addition, we can see that the inclusion of the landslide influencing factors in FC-DenseNet model-based landslide recognition led to an improvement in all six of these indicators, with an overall accuracy increase of 0.034%, a precision increase of 3.62%, a recall increase of 1.95%, a F1-score increase of 2.77%, a Kappa value increase of 2.94%, and a MIoU increase of 2.62%. Of these, precision improved the most due to the inclusion of the landslide influencing factors, which better differentiates landslide from bare rock, roads, houses, and other features.

Site 2: Jiuzhaigou—Ms 7.0 Earthquake on August 8, 2017

Although the landslide features in the Jiuzhaigou area are different from those in Zigui, most of the landslides

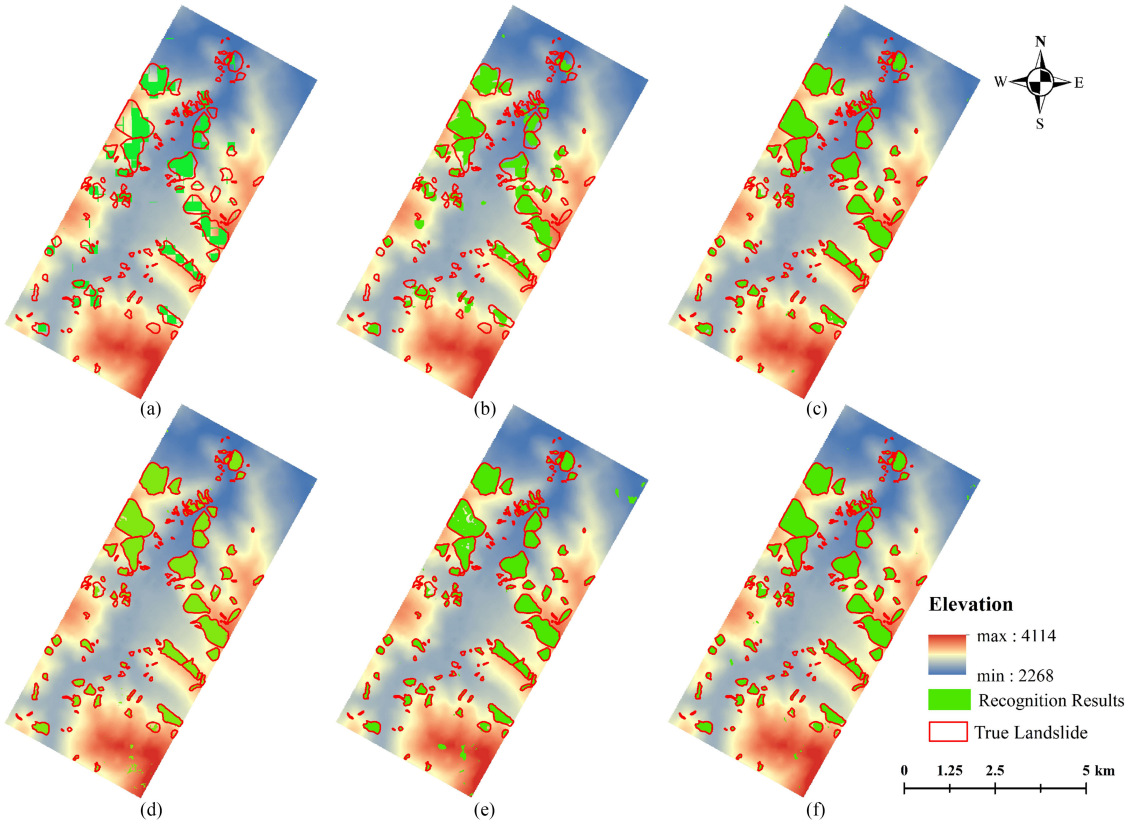


Fig .8. RML results by using different methods in Site 2. (a) FCN. (b) SegNet. (c) U-Net. (d) ResU-Net. (e) FC-DenseNet. (f) FC-DenseNet + Influencing factors.

TABLE VIII
SITE 2: QUALITATIVE EVALUATION OF SEGMENTATION RESULTS BY DIFFERENT METHODS

	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	Kappa (%)	MIoU (%)
FCN-8s	90.99	76.60	49.56	60.18	55.36	66.69
SegNet	91.65	75.84	57.61	65.48	60.83	69.81
U-Net	97.45	**96.91**	84.14	90.07	88.62	89.53
ResU-Net	98.10	96.46	89.49	92.85	91.76	92.24
FC-DenseNet	98.11	93.69	92.51	93.09	92.00	92.46
FC-DenseNet +Factors	**98.71**	94.91	**95.72**	**95.31**	**94.56**	**94.78**

Note: Bold and underlined values mean the highest number of the corresponding indicator.

in Jiuzhaigou are large seismic landslides. However, Fig. 8 shows that the RML results of the five models are similar to those of Fig. 7. Although the overall characteristics of Fig. 8(a) and (b) are similar to those of Fig. 7(a) and (b), the overall identification is better than that of the Zigui area. Most of the landslides are irregular, and the boundaries of the identification results are intricate, which is particularly the case in Fig. 8(b). As seen in Fig. 8(c), (d), and (e), the landslide boundaries are roughly outlined, but the recognition of small landslides in Fig. 8(c) is still low, while the over-recognition in Fig. 8(d) is severe and is still weaker than FC-DenseNet for small landslides. As shown in Fig. 8(e), FC-DenseNet is better for all landslide models. As shown in Fig. 8(e), FC-DenseNet is better at recognizing all types of landslides. However, there are more over-recognition errors.

To further improve the RML results of FC-DenseNet, we added landslide influencing factors to the training samples. Fig. 8(e) shows the landslide recognition results using remote sensing imagery only, while Fig. 8(f) shows the results of landslide recognition using a combination of remote sensing imagery and landslide influencing factors. Overall, both have better recognition results, but in terms of the more difficult recognition of small- to medium-sized landslides, Fig. 8(f) has notably misrecognized and over-recognized more events than Fig. 8(e), resulting in better recognition results.

From Table VIII, it can be seen that the kappa coefficients of FCN-8s and SegNet in the Jiuzhaigou area were both below 0.8, e.g., they did not meet the consistency requirement [71]. In contrast, U-Net, ResU-Net, and FC-DenseNet all obtained a better recognition accuracy. The overall accuracy of FC-DenseNet is

TABLE IX
COMPARISONS OF MODEL COMPLEXITY

Models	Parameters	RAM usage (M)	Test Time(s)
FCN	6,300,549	73	6.2
SegNet	29,461,253	338	8.5
U-Net	7,700,741	89	7.4
ResU-Net	9,874,437	114	7.8
FC-DenseNet	2,961,676	35	11.4

0.9811, the precision is 0.9369, the recall is 0.9251, the F1-score is 0.9309, the Kappa is 0.9200, and the MIoU is 0.9246. With the exception of precision, the rest of the metrics for FC-DenseNet are highest. Compared to U-Net, F1-score, Kappa and MIoU have improvements of 3.02%, 3.38%, and 2.93%.

In addition, we can see that the inclusion of the landslide influencing factors in the FC-DenseNet model-based RML led to an improvement in all six of these indicators, with the overall accuracy improving by 0.06%, the precision improving by 1.22%, the recall improving by 3.21%, the F1-score improving by 2.22%, the Kappa value improving by 2.56% and the MIoU score improving by 2.32%. This is also due to the inclusion of the landslide influencing factors, which better differentiates landslide from bare rock areas, roads, houses, and other similar features.

V. DISCUSSION

Using optical remote sensing images for RML is an important but arduous and challenging task. In this study, a new deep learning framework is used to recognize and map landslide caused by earthquakes or precipitation from optical remote sensing images. Experiments are carried out in two landslide disaster areas, and the results show that the method has a good performance in practical applications. However, some key issues still should be discussed.

A. Comparisons of Model Complexity

Accurate and rapid RML is key to disaster monitoring and emergency response. For deep learning models, the complexity of their networks has a considerable impact on the time taken for RML. Here, we briefly compare the complexity of five deep models, i.e., FCN-8s, SegNet, U-Net, ResU-Net, and FC-DenseNet, using three metrics: the number of parameters, the RAM usage of the model and the full-area test time (take Site 1 as an example).

As shown in Table IX, SegNet has the largest number of parameters and greatest RAM usage, far exceeding that of the remaining four models, which means that SegNet requires more computational resources. FCN-8s has the shortest full-area test time and a relatively smaller number of parameters and a lesser RAM usage, but this is in exchange for sacrificing most of the image feature information, so it will also produce the worst recognition results. The number of parameters and RAM usage of U-Net is slightly smaller than that of ResU-Net because ResU-Net adds skip connections to the U-Net structure, increasing the number of parameters and reducing the loss of features. In addition, the test times of the two models are similar.

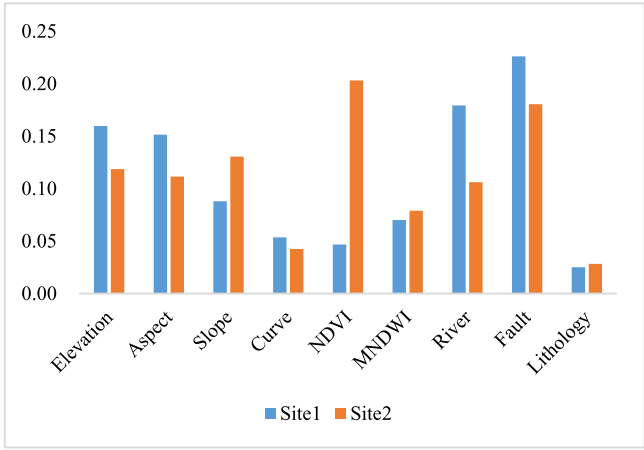


Fig. 9. IGR of the landslide influencing factors.

However, the FC-DenseNet model has a much smaller number of parameters and RAM usage than the other four models due to its unique network connectivity, and although the testing time for this model is slightly higher than that of the other models, this short time cost is acceptable from a practical application perspective.

In summary, U-Net, ResU-Net, and FC-DenseNet all meet the minimum accuracy requirements for RML (Kappa coefficients greater than 0.8), on the basis of which lightweight deep neural network models such as FC-DenseNet may be more popular.

B. Importance of Landslide Influencing Factors

To verify the effects of various factors on the performance of the model, we analyzed the importance of 9 landslide influencing factors in the study area for RML before modeling. In this study, landslide influencing factors was ranked in importance by calculating the Gini index [73]. An array of size 9 was eventually generated, with each element having a positive value and summing to 1. The higher the value of an element, the greater the contribution of the corresponding feature to the model.

The distribution of the ranked importance of each factor in the two study areas is shown in Fig. 9. First, the importance of the distance to fault factor is high in both sites because distance to fault is a causal factor, while the lithology and curvature factors are both ranked low in importance, probably because the sites selected for this experiment are relatively small and therefore the variation in lithology type and curvature within the study area is not significant. Second, for Site 1, which is reservoir landslide, hydrological categories such as distance to river and MNDWI are important, whereas in contrast, this category is less important in Site 2, which is mostly seismic landslide. Furthermore, the NDVI factor does not play a very important role in modelling as the landslides in Site 1 are mostly historical and the landslide restoration has been completed, whereas the landslides in Site 2 are mostly new and the destruction of vegetation is easily visible in the images, making the NDVI factor particularly important in the modelling. Finally, the occurrence of all landslides causes deformation of the ground surface and therefore topographic

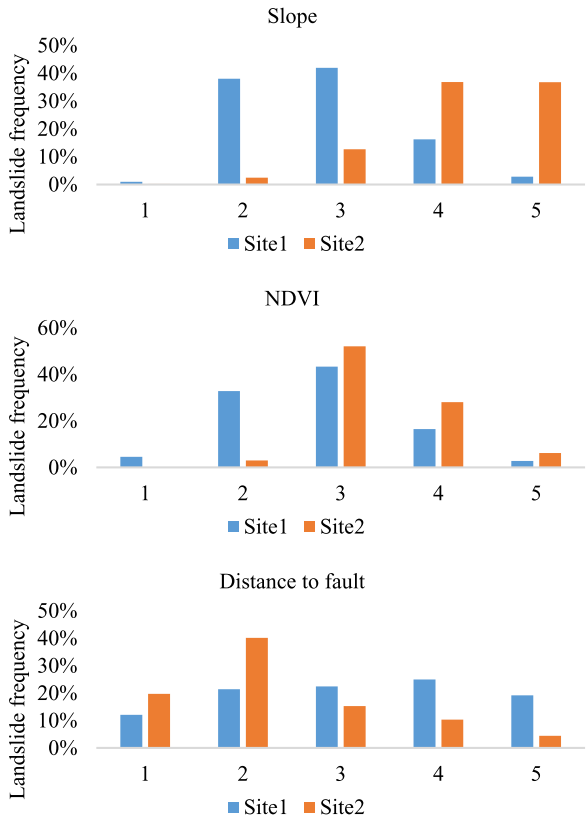


Fig. 10. Landslide frequency in factors intervals.

factors such as elevation, slope and slope direction all play a very important role in RML modelling as well.

C. Limitations and Future Work

First, a good deep learning model requires a large amount of real data for training, so the labeling of samples becomes the biggest problem. In this article, the experimental results qualitatively and quantitatively demonstrate the excellent performance of the inclusion of the new method in existing methods, but this experiment was carried out in a relatively small area. It is difficult to evaluate the performance of the training samples for the new method and determine whether the new model is suitable for RML in a larger area. Furthermore, the training data and verification data in this article are randomly selected from the same large area; the topographic features and land use types of the training data and the verification data have small spatial differences, and the optical image features are also similar. However, in practical applications, the features of landslide caused by different natural events differ, so the transferability of the model is also an urgent problem that needs to be addressed. In-depth analysis of this problem is beyond the scope of this article and is a direction of future research.

Second, our article treats the problem of RML as a pixel segmentation problem in the field of computer vision, where one category is the landslide and the other is the background. Although landslides vary in their characteristics on remotely sensed imagery, the distinction from the background is relatively clear. Of course, there are features such as bare soil, rocks, and mountain roads that have similar spectral features to landslides, which can affect the recognition result of landslide, and the inclusion of landslide influencing factors can better distinguish these features. As shown in Fig. 10, we randomly classified three landslide influencing factors into five categories in ArcGIS 10.3 according to the natural breakpoint method. We can see that the distribution of landslides in different intervals of slope, NDVI and the distance to fault is distinguishable in two sites. As long as the distribution is different, CNN can use this difference to learn the features of the landslide in different area, so as to distinguish it from the background and help improve the performance of the model. Thus, from the validation results for the two sites in this article, we can see that the generalization ability of the model is better verified within the same area where the landslides are similarly distributed.

Finally, from the perspective of the data source, this article uses a single-temporal Landsat-8 OLI image, which produces low-to-medium resolution imagery. With the development of multimodal remote sensing data, an increasing number of remote sensing images can be used in RML, but this also introduces many problems in remote sensing image processing, such as fusion, mosaicking, and the removal of cloud images. Different resolution images have varying effects on different scales and different types of landslides, which are problems we need to further explore.

VI. Conclusion

In this article, a new semantic segmentation network has been developed to recognize landslide caused by precipitation and earthquakes from single temporal Landsat-8 OLI images for the first time in the literature. In addition, a new landslide sample database is established by adding landslide influencing factors. First, the training samples are generated by processing the image and factors data, and they are then sent to the FC-DenseNet network to learn the various feature of landslide and test the whole area to obtain the regional landslide distribution map. To demonstrate the effectiveness and generalization ability of the proposed model, we conducted experiments in the Three Gorges Reservoir area and Jiuzhaigou area in China, where landslide occur frequently, and compared them with four other semantic segmentation models. The experimental results show that FC-DenseNet can better recognize the boundary and interior of landslide, and there are fewer missing and excessive recognition results. The quantitative results show that in Site 1, the optimal kappa value of the model reaches 94.72%, which is approximately 6% and 4% higher than that of U-Net and ResU-Net, respectively. In Site 2, the optimal kappa value of the model reaches 94.56%, which is approximately 6% and 3% higher than U-Net and ResU-Net, respectively. In summary, this method has good practicability in the application of RML.

References

[1] A. C. Mondini *et al.*, "Semi-automatic recognition and mapping of rainfall induced shallow landslides using optical satellite images," *Remote Sens. Environ.*, vol. 115, no. 7, pp. 1743–1757, Jul. 2011.

[2] M. C. Guzzetti *et al.*, "Comparing landslide maps: A case study in the Upper Tiber river basin, Central Italy," *Environ. Manage.*, vol. 25, no. 3, pp. 247–263, Mar. 2000.

[3] T. R. Martha *et al.*, "Landslide hazard and risk assessment using semi-automatically created landslide inventories," *Geomorphology*, vol. 184, pp. 139–150, Feb. 2013.

[4] F. Guzzetti *et al.*, "Landslide inventory maps: New tools for an old problem," *Earth-Sci. Rev.*, vol. 112, no. 1–2, pp. 42–66, Apr. 2012.

[5] C. Zhao and L. Zhong, "Remote sensing of landslides—A review," *Remote Sens.*, vol. 10, no. 2, Feb. 2018, Art. no. 279.

[6] X. Chong, "Preparation of earthquake-triggered landslide inventory maps using remote sensing and GIS technologies: Principles and case studies," *Geoence Front.*, vol. 6, no. 6, pp. 825–836, Nov. 2015.

[7] M. Galli *et al.*, "Comparing landslide inventory maps," *Geomorphology*, vol. 94, no. 3/4, pp. 268–289, Feb. 2008.

[8] M. Marjanovi *et al.*, "Landslide assessment of the Stara basin (Croatia) using machine learning algorithms," *Acta Geotechnica Slovenica*, vol. 8, no. 2, pp. 45–55, Jan. 2011.

[9] J. N. Goetz *et al.*, "Evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling," *Comput. Geosci.*, vol. 81, pp. 1–11, Aug. 2015.

[10] X. Luo *et al.*, "Researches on remote sensing images processing using deep learning methods," *J. Quanzhou Normal Univ.*, vol. 35, no. 6, pp. 40–46, Dec. 2017.

[11] T. Chen, R. Niu, and X. Jia, "A comparison of information value and logistic regression models in landslide susceptibility mapping by using GIS," *Environ. Earth Sci.*, vol. 75, no. 10, May 2016, Art. no. 867.

[12] T. Chen *et al.*, "Landslide spatial susceptibility mapping by using GIS and remote sensing techniques: A case study in Zigui County, the Three Gorges Reservoir, China," *Environ. Earth Sci.*, vol. 73, no. 9, pp. 5571–5583, May 2015.

[13] Z. Lv, T. Liu, J. A. Benediktsson, and N. Falco, "Land cover change detection techniques: Very-high-resolution optical images: A review," *IEEE Geosci. Remote Sens. Mag.*, to be published, doi: 10.1109/MGRS.2021.3088865.

[14] T. Huang *et al.*, "Research on landslides extraction based on the Wenchuan earthquake in GF-1 remote sensing image," *Bull. Surveying Mapping*, no. 2, pp. 67–71+82, 2018.

[15] Y. U. Huan *et al.*, "Optimal segmentation scale selection for object-oriented remote sensing image classification," *J. Image Graph.*, vol. 15, no. 2, pp. 352–360, 2010.

[16] T. Chen, T. John, and R. Niu, "Object-oriented landslide mapping using ZY-3 satellite imagery, random forest and mathematical morphology, for the Three-Gorges Reservoir, China," *Remote Sens.*, vol. 9, no. 4, Apr. 2017, Art. no. 333.

[17] H. Cai, T. Chen, R. Niu, and A. J. Plaza, "Landslide detection using densely connected convolutional networks and environmental conditions," *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 5235–5247, 2021.

[18] S. Chen, Z. Miao, L. Wu, and Y. He, "Application of an incomplete landslide inventory and one class classifier to earthquake-induced landslide susceptibility mapping," *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 1649–1660, 2020.

[19] A. Mohan *et al.*, "Review on remote sensing methods for landslide detection using machine and deep learning," *Trans. Emerg. Telecommun. Technol.*, vol. 32, Jun. 2020, Art. no. e3998.

[20] C. Lissak *et al.*, "Remote sensing for assessing landslides and associated hazards," *Surv. Geophys.*, vol. 41, no. 6, pp. 1391–1435, Sep. 2020.

[21] 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. Autom.*, Nov. 2016, pp. 444–448.

[22] Y. Lecun and L. Bottou, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.

[23] Y. Bengio and Y. Lecun, "Pattern recognition and neural networks," in *Handbook of Brain Theory & Neural Networks*. Cambridge, MA, USA: MIT Press, Nov. 1996.

[24] Y. Lecun, "LeNet-5, convolutional neural networks," vol. 20, no. 5, p. 14, 2015. [Online]. Available: http://yann.lecun.com/exdb/lenet

[25] T. Technicolor *et al.*, "ImageNet classification with deep convolutional neural networks," *ILSVRC*, vol. 60, no. 6, pp. 84–90, Jun. 2017.

[26] C. Szegedy *et al.*, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 1–9.

[27] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *Comput. Sci.*, 2014, *arXiv: 1409.1556*.

[28] K. He *et al.*, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 770–778.

[29] 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. 2261–2269.

[30] Y. Wang, Z. Fang, and H. Hong, "Comparison of convolutional neural networks for landslide susceptibility mapping in Yanshan County, China," *Sci. Total Environ.*, vol. 666, pp. 975–993, May 2019.

[31] S. L. Ullo *et al.*, "A new mask R-CNN based method for improved landslide detection," *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 3799–3810, 2021.

[32] A. Pttn *et al.*, "Evaluation of deep learning algorithms for national scale landslide susceptibility mapping of Iran," *Geosci. Front.*, vol. 12, no. 2, pp. 505–519, Mar. 2021.

[33] U. Ozturk *et al.*, "How robust are landslide susceptibility estimates?," *Landslides*, vol. 18, no. 2, pp. 681–695, Aug. 2020.

[34] Z. Su *et al.*, "Deep convolutional neural network–based pixel-wise landslide inventory mapping," *Landslides*, vol. 18, no. 4, pp. 1421–1443, Oct. 2020.

[35] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 4, pp. 640–651, Apr. 2015.

[36] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," *Lecture Notes Comput. Sci.*, vol. 9351, pp. 234–241, 2015.

[37] X. Xiao, L. Shen and Z. Luo, "Weighted Res-UNet for high-quality retina vessel segmentation," in *Proc. 9th Int. Conf. Inf. Technol. Med. Educ.*, Oct. 2018, pp. 327–331.

[38] P. Liu *et al.*, "Research on post-earthquake landslide extraction algorithm based on improved U-Net model," *Remote Sens.*, vol. 12, no. 5, Mar. 2020, Art. no. 894.

[39] Y. Chen *et al.*, "Mapping post-earthquake landslide susceptibility: A U-Net like approach," *Remote Sens.*, vol. 12, no. 17, 2020, Art. no. 2767.

[40] 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*.

[41] C. Wu, J. Qiao, and M. Wang, "Landslides and slope aspect in the Three Gorges Reservoir area based on GIS and information value model," *Wushan Univ. J. Natural Sci.*, vol. 11, no. 4, pp. 773–779, 2006.

[42] J. Qiao *et al.*, "Bottom factors applied to the zoning study of the risk levels of landslides in the Three Gorges Reservoir area," *J. Mountain Sci.*, vol. 24, no. 5, pp. 569–573, 2006.

[43] L. Liu *et al.*, "Research on the distribution of major metal elements in the typical landslide soil of the Three Gorges Reservoir area," *J. Earth Sci.*, vol. 23, no. 2, pp. 207–212, Apr. 2012.

[44] T. Chen *et al.*, "Mapping landslide susceptibility at the Three Gorges Reservoir, China, using gradient boosting decision tree, random forest and information value models," *J. Mountain Sci.*, vol. 17, no. 3, pp. 670–685, Mar. 2020.

[45] Q. Li *et al.*, "Recognition of earthquake-induced landslide and spatial distribution patterns triggered by the Jiuzhaigou earthquake in August 8, 2017," *Remote Sens.*, vol. 23, no. 4, pp. 785–795, 2019.

[46] J. Wang *et al.*, "Earthquake-triggered landslides affecting a UNESCO natural site: The 2017 Jiuzhaigou earthquake in the World National Park, China," *J. Mountain Sci.*, vol. 15, no. 7, pp. 1412–1428, Jul. 2018.

[47] I. Yilmaz, "Comparison of landslide susceptibility mapping methodologies for Koyulhisar, Turkey: Conditional probability, logistic regression, artificial neural networks, and support vector machine," *Environ. Earth Sci.*, vol. 61, no. 4, pp. 821–836, Aug. 2010.

[48] Y. Tian *et al.*, "Inventory and spatial distribution of landslides triggered by the 8th August 2017 MW 6.5 Jiuzhaigou earthquake, China," *J. Earth Sci.*, vol. 30, no. 1, pp. 206–217, Feb. 2019.

[49] M. Chang *et al.*, "The spatial distribution characteristics of coseismic landslides triggered by the Ms7.0 Lushan earthquake and Ms7.0 Jiuzhaigou earthquake in Southwest China," *Environ. Sci. Pollut. Res.*, vol. 28, no. 16, pp. 20549–20569, Apr. 2021.

[50] P. Sun *et al.*, "A method to enhance information of water cover based on feature space of NDVI and MNDWI," *J. Hubei Univ. (Natural Sci.)*, vol. 40, no. 6, pp. 29–34, 2018.

[51] Y. Q. Wang, "The sensitivity analysis and control measures to the stability influencing factors landslide of Qianjiang East Bus Station in Chongqing," *J. Changchun Inst. Technol. (Natural Sci. Ed.)*, vol. 13, no. 4, pp. 79–85, 2012.

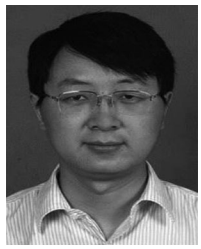
[52] S. Lee and N. T. Dan, "Probabilistic landslide susceptibility mapping in the Lai Chau province of Vietnam: Focus on the relationship between tectonic fractures and landslides," *Environ. Geol.*, vol. 48, no. 6, pp. 778–787, Sep. 2005.
[53] W. Yu *et al.*, "Potential influencing factors analysis and safety evaluation on the landslide of high and steep slope," *J. Univ. Sci. Technol. Beijing*, vol. 30, no. 3, pp. 227–232, 2008.
[54] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 12, pp. 2481–2495, Dec. 2017.
[55] K. Bailey, "A novel approach to real-time bilinear interpolation," in *Proc. Int. Symp. Electron. Des. Test Appl. DELTA*, 2004, pp. 126–131.
[56] Z. Lv, T. Liu, C. Shi, and J. A. Benediktsson, "Local histogram-based analysis for detecting land cover change using VHR remote sensing images," *IEEE Geosci. Remote Sens. Lett.*, vol. 18, no. 7, pp. 1284–1287, Jul. 2021.
[57] O. Ghorbanzadeh *et al.*, "Evaluation of different machine learning methods and deep-learning convolutional neural networks for landslide detection," *Remote Sens.*, vo1. 11, no. 2, Jan. 2019, Art. no. 196.
[58] Y. Yi and W. Zhang, "A new deep-learning-based approach for earthquake-triggered landslide detection from single-temporal rapid-eye satellite imagery," *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 6166–6176, 2020.
[59] L. Ying and L. Wua, "Geological disaster recognition on optical remote sensing images using deep learning," *Procedia Comput. Sci.*, vol. 91, pp. 566–575, 2016.
[60] C. Ye *et al.*, "Landslide detection of hyperspectral remote sensing data based on deep learning with constrains," *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, vol. 12, no. 12, pp. 5047–5060, Dec. 2019.
[61] L. Perez and J. Wang, "The effectiveness of data augmentation in image classification using deep learning," Dec. 2017, *arXiv:1712.04621*.
[62] E. D. Cubuk, B. Zoph, D. Mane, V. Vasudevan, and Q. V. Le, "AutoAugment: Learning augmentation strategies from data," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 113–123.
[63] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.
[64] O. K. Oyedotun, A. Shabayek, D. Aouada, and B. Ottersten, "Highway network block with gates constraints for training very deep networks," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Workshops*, 2018, pp. 1739–1748.
[65] H. Gao *et al.*, "Deep networks with stochastic depth," *Lecture Notes Comput. Sci.*, vol. 9908, pp. 646–661, 2016.
[66] G. Larsson, M. Maire, and G. Shakhnarovich, "FractalNet: Ultra-deep neural networks without residuals," May 2016, *arXiv:1605.0748*.
[67] J. Ge *et al.*, "BNReLU: Combine batch normalization and rectified linear unit to reduce hardware overhead," in *Proc. Int. Conf. ASIC*, 2019, pp. 1–4.
[68] S. H. Wang *et al.*, "Alcoholism recognition via convolutional neural network based on parametric ReLU, dropout, and batch normalization," *Neural Comput. Appl.*, vol. 32, no. 32, pp. 665–680, Dec. 2018.
[69] M. F. Ahmed and J. D. Rogers, "Regional level landslide inventory maps of the Shyok river watershed, Northern Pakistan," *Bull. Eng. Geol. Environ.*, vol. 75, no. 2, pp. 563–574, May. 2016.
[70] O. Ghorbanzadeh, S. R. Meena, H. Shahabi, S. T. Piralilou, and L. Zhiyong, "Landslide mapping using two main deep-learning convolution neural network (CNN) streams combined by the dempster—Shafer (DS) model," *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 452–463, 2021.
[71] X. Sun *et al.*, "Landslide susceptibility mapping along the upper Jinsha river, South-Western China: A comparison of hydrological and curvature watershed methods for slope unit classification," *Bull. Eng. Geol. Environ.*, vol. 79, no. 9, pp. 4657–4670, Nov. 2020.
[72] W. Tang *et al.*, "Kappa coefficient: A popular measure of rater agreement," *Gen. Psychiatry*, vol. 27, no. 1, pp. 62–67, 2015.

Xiao Gao received the B.S. degree in geographic information science from Henan University, Kaifeng, China, in 2019. She is currently working toward the M.S. degree in earth exploration and information technology with the China University of Geosciences, Wuhan, China.

Her current main research interests include data mining of geographic information based on deep learning.

Tao Chen (Senior Member, IEEE) received the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2008.

From 2015 to 2016, he was a Visiting Scholar with the University of New South Wales, Sydney, Australia. He has authored or coauthored more than 40 scientific papers including IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, REMOTE SENSING, ENVIRONMENTAL EARTH SCIENCES, and *Environmental Science and Pollution Research* and guest edited two journal special issues. He is currently an Associate Professor with the Institute of Geophysics and Geomatics, China University of Geosciences, Wuhan. His research interests include image processing, machine learning, and geological remote sensing.



Ruiqing Niu received the Ph.D. degree in earth exploration and information technology from the China University of Geosciences, Wuhan, China, in 2005.

He is currently a Full Professor with the Institute of Geophysics and Geomatics, China University of Geosciences. His research interests include remote sensing, geographic information system, and engineering geology.

Antonio Plaza (Fellow, IEEE) received the M.Sc. and Ph.D. degrees in computer engineering from the Department of Technology of Computers and Communications, University of Extremadura, Badajoz, Spain, in 1999 and 2002, respectively.

He is currently a Full Professor and the Head with the Hyperspectral Computing Laboratory, Department of Technology of Computers and Communications, University of Extremadura. He has authored more than 600 publications and guest edited ten journal special issues. He has reviewed more than 500 manuscripts for over 50 different journals.

Dr. Plaza was the Editor-in-Chief for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING from 2013 to 2017. He was included in the Highly Cited Researchers List (Clarivate Analytics) from 2018 to 2020.