



Refined glacial lake extraction in a high-Asia region by deep neural network and superpixel-based conditional random field methods

Yungang Cao¹, Rumeng Pan¹, Meng Pan², Ruodan Lei¹, Puying Du¹, and Xueqin Bai³

¹Faculty of Geosciences and Environmental Engineering, Southwest Jiaotong University, Chengdu, 611756, China

²College of Civil Engineering, Shangqiu Institute of Technology, Shangqiu, 476000, China

³Chongqing Vocational Institute of Engineering, Chongqing, 402260, China

Correspondence: Yungang Cao (yungang@swjtu.cn) and Xueqin Bai (baixueqin1998@163.com)

Received: 24 December 2022 – Discussion started: 15 February 2023

Revised: 25 October 2023 – Accepted: 20 November 2023 – Published: 8 January 2024

Abstract. Remote sensing extraction of glacial lakes is an effective way of monitoring water body distribution and outburst events. At present, the lack of glacial lake datasets and the edge recognition problem of semantic segmentation networks lead to poor accuracy and inaccurate outlines of glacial lakes. Therefore, this study constructed a high-resolution dataset containing seven types of glacial lakes and proposed a refined glacial lake extraction method, which combines the LinkNet50 network for rough extraction and simple linear iterative clustering (SLIC) dense conditional random field (DenseCRF) for optimization. The results show that (1) with Google Earth images of 0.52 m resolution in the study area, the recall, precision, F1 score, and intersection over union (IoU) of glacial lake extraction based on the proposed method are 96.52 %, 92.49 %, 94.46 %, and 90.69 %, respectively, and (2) with the Google Earth images of 2.11 m resolution in the Qomolangma National Nature Reserve, 2300 glacial lakes with a total area of 65.17 km² were detected by the proposed method. The area of the minimum glacial lake that can be extracted is 160 m² (less than 6 × 6 pixels). This method has advantages in small glacial lake extraction and refined outline detection, which can be applied to extracting glacial lakes in the high-Asia region with high-resolution images.

Asia (Yao et al., 2017). Glacial lakes have a strong relationship to ongoing climate change (Pandey et al., 2021). Climate warming, continuous glacier retreat and ablation of differences in the debris cover have led to the formation of a large number of glacial lakes and the continuous expansion of glacial lake areas (Nie et al., 2017; Chen et al., 2021). In the past 30 years, the number of glacial lakes in high Asia has increased by 17.4 %, the total area has increased by 17.3 %, and the glacial lake area in the whole region expanded by 0.58 % a⁻¹ (M. M. Zhang et al., 2022). The rapid change of glacial lakes may increase the possibility of the occurrence of glacial lake outburst floods (GLOFs) (Zhong et al., 2021). The risk of GLOFs in high Asia is the highest (Taylor et al., 2023). This may threaten the lives and property of 30 surrounding residents and downstream infrastructures (Song et al., 2016; Begam et al., 2018; Nie et al., 2023). For example, the GLOF in Tibet on 26 June 2020 led to the destruction of 43.9 km of roads and eight bridges, as well as the flooding of 19.98 ha of farmland (Zheng et al., 2021). Therefore, continuous dynamic monitoring of glacial lakes is essential to studies on climate change, water resource distribution, and disaster warnings. However, many small and unevenly distributed glacial lakes are ignored; these glacial lakes usually have a high risk of outburst (T. G. Zhang et al., 2022). With the help of high-resolution images and deep neural networks, this study focuses on the extraction of glacial lakes at different spatial scales, especially improving the extraction accuracy of small glacial lakes.

For the glacial lake outline delineation, there are mainly manual digitization methods, semi-automatic methods, and automatic methods. The manual digitization method has

1 Introduction

Glacial lakes are natural water bodies mainly supplied by glacier meltwater or formed by water accumulation in moraine ridge depressions and are densely distributed in high

achieved good results (Yang et al., 2019). Nevertheless, it costs lots of time and resources, which is challenging to meet the needs for large-scale glacial lake identification. For the semi-automatic extraction of glacial lakes, current studies are mainly based on water body indices (Li et al., 2020) and machine learning (Jain et al., 2015; Veh et al., 2018). In 2015, Jain et al. (2015) used the support vector machine (SVM) to detect glacial lakes in Bhutan, Himalayas. In 2018, Veh et al. (2018) trained a random forest (RF) classifier based on Landsat data and detected glacial lake outbursts through change detection technology. Due to the different elements contained in the water body of the glacial lake and the depth of the glacial lake, the glacial lake has a large intra-class heterogeneity, and different spectral information is displayed on the optical remote sensing image (Zhao et al., 2018). The semi-automatic extraction methods are manually dependent and regionally restrictive, limiting the promotion and application on the global/hemispheric scale.

For automatic extraction methods, there are mainly image segmentation methods (M. M. Zhang et al., 2018) and edge detection algorithms (Cordeiro et al., 2021). These methods establish fixed models or rules and then execute them automatically, completing extraction work without manual intervention. The edge detection algorithm is one of the most classical and advanced image edge detection algorithms (Chen, 2021). Threshold and simplified C-V (TSCV) based on image segmentation technology has a better effect (Zhao et al., 2018), which could overcome the impact of spectral heterogeneity. However, it is only applicable to remote sensing images with multiple bands (especially SWIR) such as Landsat images, while most high-resolution images have only four bands of red, green, blue, and near-infrared.

Except for the traditional automatic image segmentation methods, with the development of computer vision technology, some image semantic segmentation networks have been successfully applied in water body recognition (Chen et al., 2018; Talal et al., 2018; Wang et al., 2019; S. D. Wang et al., 2022). Based on PlanetScope imagery, Qayyum et al. (2020) used the pre-trained EfficientNet as the backbone of the U-Net to map glacial lakes, which achieved a better result than the original U-Net, RF, and SVM classifiers in high-resolution glacial lake extraction. Given that the area of the glacial lake is much smaller compared with the background, the skip connection structure may transfer a large amount of redundant background information from the low level to the high level, thereby reducing the efficiency of utilizing low-level features. He et al. (2021) added a space attention mechanism into the skip connection of U-Net to focus on glacial lakes. J. Wang et al. (2022) proposed NAU-Net with the normalized difference water index (NDWI) as the spatial attention, which guided the network to pay more attention to the glacial lake information of low-level features and solved the problem of the area difference between the area occupied by positive and negative samples. However, for high-resolution Google Earth images, there are more problems with complex

spectral and texture features that lead to the large intra-class variance of glacial lakes. Therefore, based on high-spatial-resolution data, Wang et al. (2020) extracted lakes on the Tibetan Plateau with a more complex network (MSLWENet). Although the texture of the water body was complex, resulting in more noise in the segmentation, the study showed that the deeper network achieved better performance than U-Net, DeepLab V3+ (Li et al., 2019).

For end-to-end semantic segmentation networks, the network is vulnerable to negative samples because glacial lakes have small areas, and part of the spatial information is difficult to recover during upsampling (Song et al., 2019). Besides, high-resolution images provide not only rich spectral information of glacial lakes but also contain a lot of noise information and a deep network is needed. Considering the characteristics of high-resolution images and the limitations of semantic segmentation networks, this study proposed an automatic method for the refined glacial lake extraction. The main contributions of this study are as follows:

1. a glacial lake dataset with abundant glacial lake types and sufficient samples was constructed in this study;
2. to alleviate the negative impact of unbalanced positive and negative samples on the network extraction for glacial lake features, a two-step constrained loss function and training strategy were proposed with ResNet50 as the backbone;
3. simple linear iterative clustering (SLIC) and dense conditional random field (DenseCRF) were combined for post-processing to reduce the noise of segmentation results and optimize glacial lake outlines.

2 Study area and data

2.1 Study area

The study area is the Mount Qomolangma area (27°08′09″–29°19′14″ N, 84°25′16″–88°23′12″ E; also known as Mount Everest), which is the southwestern part of the Tibetan Plateau. Mount Qomolangma is located on the border between China and Nepal. The blue rectangle area (Fig. 1) is the study area for the glacial lake extraction in this study. The glaciers in the study area are cirque glaciers, which are distributed in depressions near the snow line (Ke et al., 2016). The annual precipitation in the area is less than 500 mm (Qi et al., 2013). There are no large rivers in the study area. The water supply of glacial lakes mainly relies on the melting water of ice and snow. Small streams developed by glacial lakes and glaciers in the study area are also marked in Fig. 1. According to the classification system (Yao et al., 2017), the glacial lakes are mainly moraine-dammed lakes and glacial erosion lakes (cirque lakes), while the small glacial lakes are mainly moraine thaw lakes, accounting for the largest proportion in number.

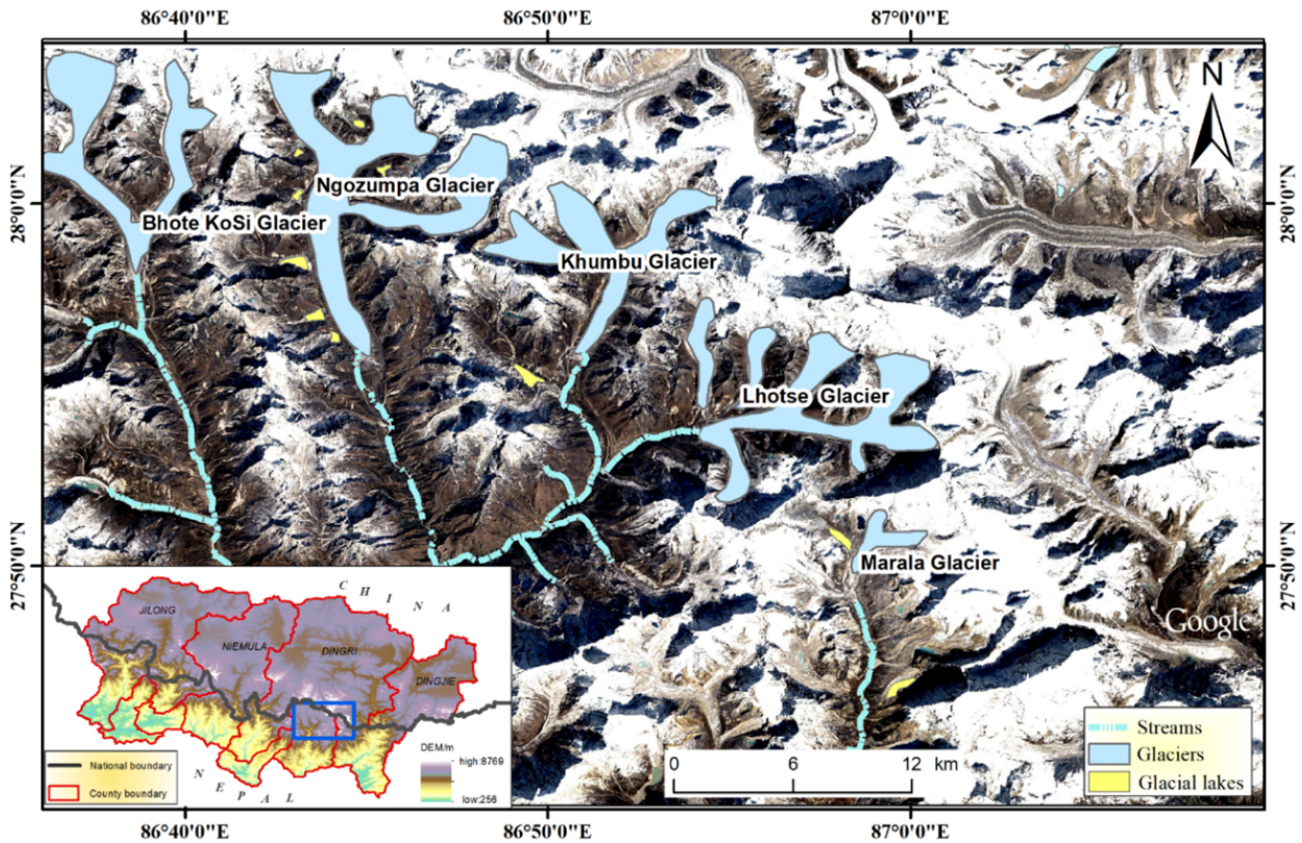


Figure 1. Location and topography information of the study area. All images © Google Earth 2020.

2.2 Data

2.2.1 Dataset and preprocessing

The global glacial lakes are mainly distributed in mountainous areas with many glaciers, including the Himalayas in Asia, the Buenos Aires mountains in South America (Bourgeois et al., 2016), the Alaska mountains in North America (Rick et al., 2022), and the Alps in Europe (Huggel et al., 2002). Among them, Mount Qomolangma has different kinds of glacial lakes, such as the glacial erosion lake and the moraine-dammed lake, including seven types of glacial lakes according to the classification system summarized by Yao et al. (2017) (Fig. 2). Due to different development environments, the morphology of glacial lakes may differ in remote sensing images (Zhao et al., 2018). Collecting more samples of different types is of great help to enhance the stability and universality of the model (He et al., 2021). For the sake of increasing the diversity of the training dataset, except for the high-Asia region, this study also collected some glacial lake samples from other continents.

Google Earth imagery is a composite of a vast array of satellite and aerial photographs. These images are sourced from a variety of providers and platforms that are responsible for satellite launches. The primary contributors of high-

resolution imagery include Maxar Technologies, the Centre National d'Études Spatiales (CNES), and Airbus. They provide IKONOS, QuickBird, GeoEye, WorldView, SPOT, and Pleiades imagery. Given that the sources of images vary across different regions, there is not a consistent time frame for image acquisition or a fixed spatial resolution. In Google Earth imagery, spatial resolution is categorized by levels – the higher the level, the greater the spatial resolution. In the data preprocessing, 14 to 19 levels of Google Earth images were chosen for the glacial lake dataset, and the image resolution covers the range of 5 to 0.14 m. When labeling, the glacial lakes were manually outlined with the help of ENVI 5.3 (other software, such as LabelMe and ArcGIS, can also do labeling), and every pixel in the image was labeled as 1 for glacial lakes or 0 for the background. When training the deep learning model, the images that were inputted into the network needed to be processed into image tiles for the limitation of the computer's memory capacity. After many experiments, it was more appropriate to divide the input images into non-overlapping image tiles of size 256×256 . Image tiles that do not contain glacial lakes were removed to alleviate the problem of large background areas. Data augmentation operations were carried out to increase the number of samples, like image rotation. Finally, a total of 15 376 samples with a size of 256×256 were obtained, out of which

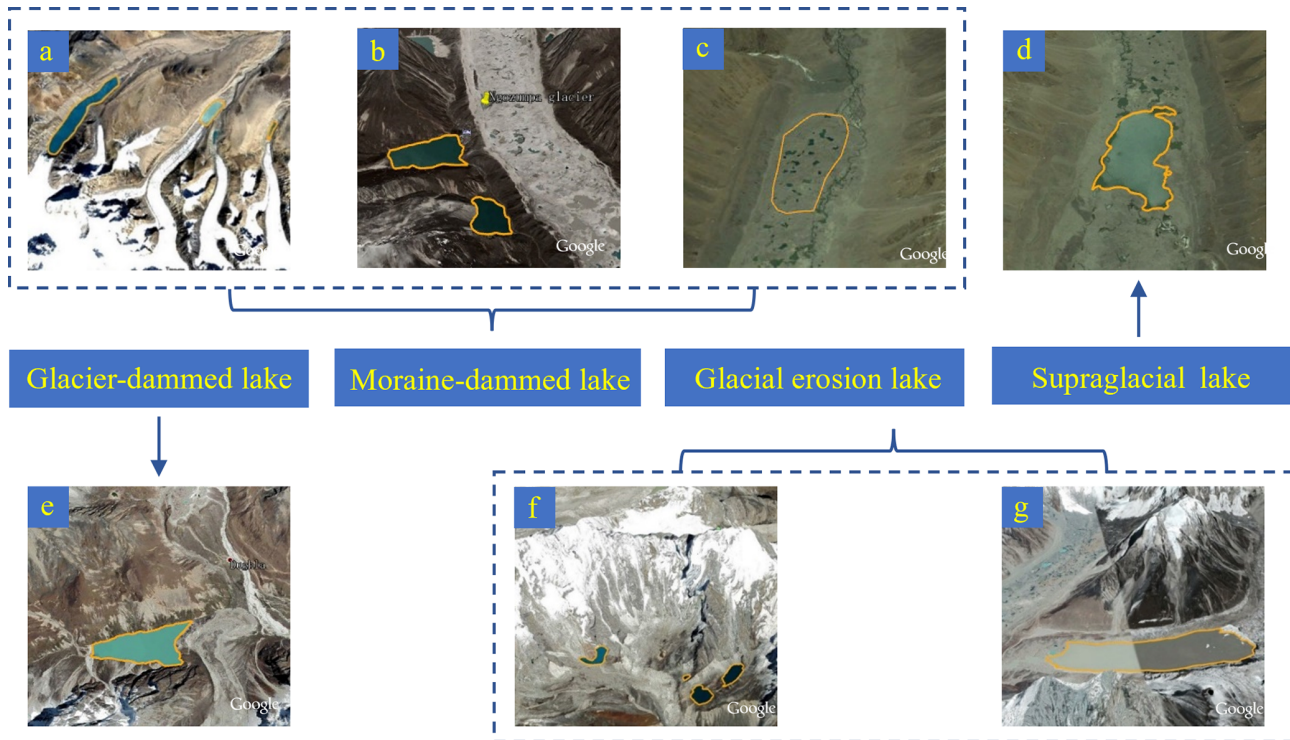


Figure 2. Seven types of glacial lakes in the training dataset. All images © Google Earth 2020. (a) Terminal moraine-dammed lake, (b) side moraine-dammed lake, (c) moraine thaw lake, (d) supraglacial lake, (e) glacial-dammed lake, (f) cirque lake, and (g) glacier valley lake.

20 % of image tiles were selected as validation data randomly (Table 1).

2.2.2 Other datasets

Except for the training dataset, other data products were also used to assist in completing this study (Table 2). The second glacier inventory dataset of China was used to delineate the distribution area of glacial lakes (Liu et al., 2012). In Sect. 4.3, the 30 m glacial lake inventory in western China based on Landsat TM/ETM+/OLI data (Wang, 2018) and three global land cover products (Gong et al., 2019) based on Sentinel-2 images were used for comparison with the glacial lakes extracted in this study.

3 Methods

In the glacial lake extraction method, based on Google Earth images, this study used the semantic segmentation framework to achieve rough extraction of the glacial lake first (“output1” in Fig. 3). In the pixel-based semantic segmentation, the outline of the glacial lake is not refined enough, which does not fit the actual smooth edge of the glacial lake. The simple linear iterative clustering (SLIC) algorithm could fuse the rough result of semantic segmentation with the edge information of superpixel segmentation to enhance the integrity of the glacial lake and improve the edge segmentation

(“output2” in Fig. 3). The dense conditional random field (DenseCRF) uses the constraint relationship between pixels to encourage similar pixels to be assigned the same label, while pixels with large differences are assigned different labels to obtain accurate glacial lake outlines (“output3” in Fig. 3). In this paper, after semantic segmentation, two-level optimization combined SLIC and DenseCRF and was used to achieve refined extraction of glacial lake outlines. These two optimization methods can also be used separately to implement single-level optimization.

3.1 Rough extraction of glacial lake information based on the semantic segmentation network

3.1.1 Deep residual network LinkNet50

The LinkNet network (Chaurasia and Culurciello, 2017) uses ResNet18 (He et al., 2016) as the backbone of the U-Net (Sathananthavathi and Indumathi, 2021). In the LinkNet network, the size of each layer feature map corresponding to encoder and decoder is the same, the addition method is used to combine the features, and the shallow features are re-learned without increasing the parameters, so that the spatial information of the glacial lake can be effectively restored, which has a lightweight structure and fast calculation speed. Good results have also been achieved in identifying glacial lakes, but the effect is unsatisfactory in areas cov-

Table 1. Details of the glacial lake training dataset based on Google Earth images in this study. All images © Google Earth 2020.

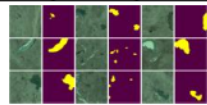
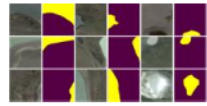


Area	Image source	Image time	Data level	Number	Sample examples
Himalaya, Northern Tibet Plateau, Asia	Pleiades 1,2; SPOT 5,6,7	2010-2020	Level14-18 (4.45-0.28m)	5494	
Buenos Aires Mountains, South America	Pleiades 1,2; QuickBird	2004-2016	Level 17,18 (0.79m, 0.4m)	3397	
Alaska Mountains, North America	WorldView 3	2013-2017	Level 18,19 (0.28m, 0.14m)	5519	
the Alps, Europe	Geoeye 1; WorldView 3	2015-2020	Level 18,19 (0.41m, 0.21m)	966	

Table 2. Other datasets used in this study.

Dataset	Temporal coverage	Res. m ⁻¹	Source
The second glacier inventory dataset of China (V1.0)	2006–2011	–	National Tibetan Plateau/Third Pole Environment Data Center (TPDC) http://www.tpdc.ac.cn/en/ (last access: 23 May 2022)
Inventory data of glacial lake in west China	2015	30	National Tibetan Plateau/Third Pole Environment Data Center (TPDC) http://www.tpdc.ac.cn/en/ (last access: 1 May 2022)
FROM-GLC10	2017	10	Tsinghua University (THU) http://data.ess.tsinghua.edu.cn/ (last access: 25 May 2022)
ESA WorldCover	2020	10	European Space Agency (ESA) https://viewer.esa-worldcover.org/worldcover (last access: 25 May 2022)
Esri Land Cover	2020	10	Environmental Systems Research Institute (ESRI) https://livingatlas.arcgis.com/landcover/ (last access: 25 May 2022)

ered with ice and snow or with a small number of glacial lakes. Moreover, high-resolution images in the dataset built in this study have complex spatial and spectral information, and deeper networks are more beneficial for high-level feature extraction (Li et al., 2020; Wang et al., 2021). Meanwhile, LinkNet50 has achieved good results in road detection based on high-resolution images (Li and Liu, 2022). Therefore, to obtain more useful features to distinguish glacial lakes from the background, this study used a deep residual network (ResNet50) instead of ResNet18 as the backbone of U-Net.

As shown in Table 3, ResNet of different depths contains five stages, and the output results (feature images) of the second to fifth stages are Res2, Res3, Res4, and Res5. In ResNet50, their sizes (width × height × channel) are 64 × 64 × 256, 32 × 32 × 512, 16 × 16 × 1024, and 8 × 8 × 2048, respectively. On the right side in Fig. 4, ResNet50 is used in the encoder of LinkNet50 for feature extraction to obtain high-level features. The input of each encoder layer is also bypassed to the output of the corresponding decoder layer. On the right side, the decoder uses the residual structure to combine low-level and high-level features and recover the detailed information of the image lost by the downsampling.

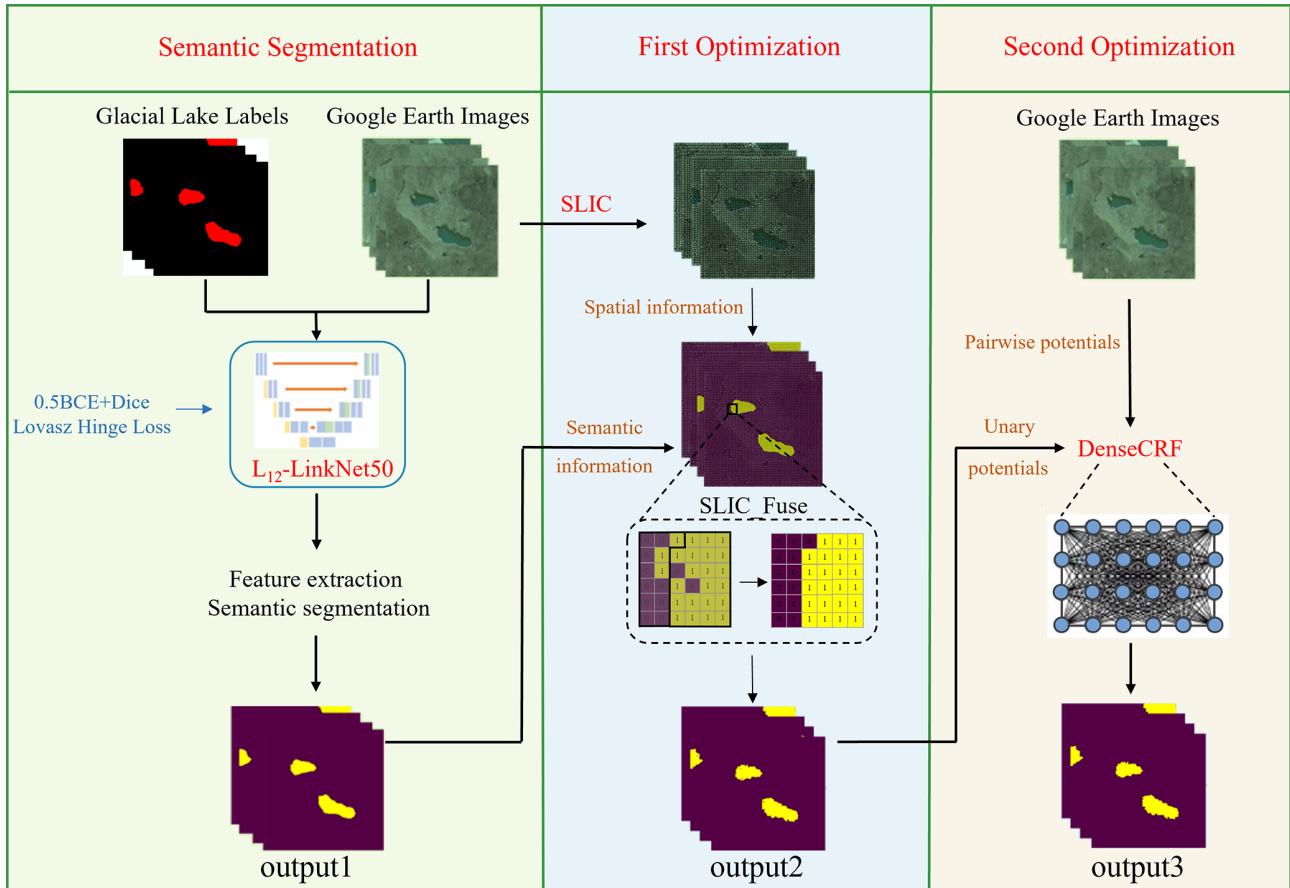


Figure 3. Structure diagram of glacial lake extraction strategy in this study. All images © Google Earth 2020.

Table 3. The structure of ResNet of different network depths.

Stage	Output feature map size (width × height)	ResNet18	ResNet34	ResNet50	ResNet101	ResNet152
C_1	128×128	A convolutional layer with the kernel of 7×7 , the stride of 2 A max-pooling layer with the kernel of 3×3 , the stride of 2				
C_2	64×64	$\begin{bmatrix} 3 \times 3.64 \\ 3 \times 3.64 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3.64 \\ 3 \times 3.64 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1.64 \\ 3 \times 3.64 \\ 1 \times 1.256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1.64 \\ 3 \times 3.64 \\ 1 \times 1.256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1.64 \\ 3 \times 3.64 \\ 1 \times 1.256 \end{bmatrix} \times 3$
C_3	32×32	$\begin{bmatrix} 3 \times 3.128 \\ 3 \times 3.128 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1.128 \\ 3 \times 3.128 \\ 1 \times 1.512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1.128 \\ 3 \times 3.128 \\ 1 \times 1.512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1.128 \\ 3 \times 3.128 \\ 1 \times 1.512 \end{bmatrix} \times 8$
C_4	16×16	$\begin{bmatrix} 3 \times 3.256 \\ 3 \times 3.256 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3.256 \\ 3 \times 3.256 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1.256 \\ 3 \times 3.256 \\ 1 \times 1.1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1.256 \\ 3 \times 3.256 \\ 1 \times 1.1024 \end{bmatrix} \times 23$	$\begin{bmatrix} 1 \times 1.256 \\ 3 \times 3.256 \\ 1 \times 1.1024 \end{bmatrix} \times 36$
C_5	8×8	$\begin{bmatrix} 3 \times 3.512 \\ 3 \times 3.512 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3.512 \\ 3 \times 3.512 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1.512 \\ 3 \times 3.512 \\ 1 \times 1.2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1.512 \\ 3 \times 3.512 \\ 1 \times 1.2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1.512 \\ 3 \times 3.512 \\ 1 \times 1.2048 \end{bmatrix} \times 3$
	1×1	Average pooling layer, 1000-dimensional fully connected layer, the softmax function				

The size of the input image is $256 \times 256 \times 3$. In the matrix multiplication expressions (five columns on the right), take ResNet18 as an example, where 3×3 indicates the convolution kernel size, 64 indicates the number of channels of the output image, and 2 indicates two residual blocks.

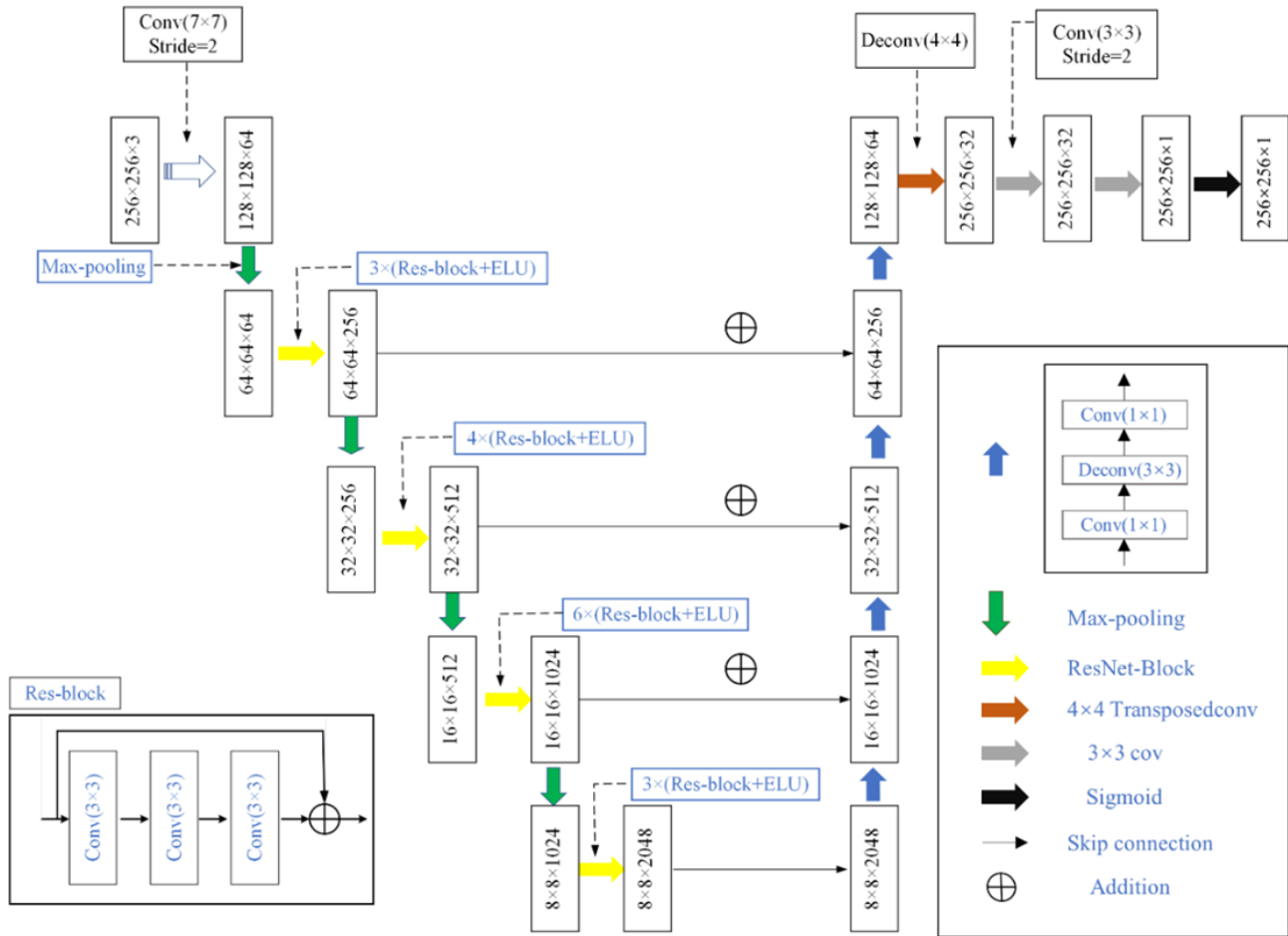


Figure 4. Schematic diagram of the LinkNet50 network structure used in this study. In the decoder, “conv” (1 × 1) is responsible for reducing the number of channels (×1/4), and “deconv” (3 × 3) only changes the size of the feature map (×2). After the decoder, “transposedconv” (deconv (4 × 4)) reduces the number of channels (×1/2) and expands the size of the feature map (×2).

3.1.2 Loss function

In the dataset in this study, the glacial lake area is small and the background area is large (unbalanced samples), so the target’s features cannot be fully learned during the model training process. Therefore, to solve this problem, dice loss, which can help to reduce the impact of unbalanced positive and negative samples in binary classification, was used in this study. Dice loss essentially measures the overlap of two samples, and the calculation formula is

$$d = 1 - 2|X \cap Y| / (|X| + |Y|), \tag{1}$$

where $|X|$ and $|Y|$ indicates the number of pixels of sample X and sample Y , respectively, and $|X \cap Y|$ indicates the intersection of X and Y . For the common part that is repeatedly calculated, the coefficient of the $|X \cap Y|$ is 2.

However, dice loss will affect backpropagation, making the loss change unstable during model training. To increase the stability of the training process, BCE loss was introduced in this study. BCE loss belongs to the cross-entropy loss

function, which is used to evaluate the difference between the probability distribution obtained by the training model and the natural distribution. In binary classification, the model predicted the probability of each category as p and $1 - p$, respectively, and the loss function is

$$L = \frac{1}{N} \sum_i L_i = \frac{1}{N} \sum_i - [y_i \cdot \log(p_i) + (1 - y_i) \cdot \log(1 - p_i)], \tag{2}$$

where y_i is the label of sample i , 1 is for positive, and 0 is for negative. p_i is the probability that sample i is predicted to be positive. After testing, we finally adopted Loss1 = 0.5BCE + Dice as the loss function of the training model, which not only solved the problem of unbalanced positive and negative samples but also increased the stability of the network training process.

In addition, the Lovasz hinge loss is a convex Lovasz extension of submodular losses, which could optimize the IoU (intersection over union) loss of the network in the condi-

tion of unbalanced sample distribution (Berman et al., 2018). It is worth noting that the LinkNet50 with the Loss1 as the loss function is called L_1 -LinkNet50 in this study. After L_1 -LinkNet50, the Lovasz hinge loss (Loss2) was further used to fine-tune the deep semantic segmentation network in this study, which is referred to L_{12} -LinkNet50.

3.2 High-precision edge optimization algorithm

3.2.1 Simple linear iterative clustering (SLIC)

SLIC is a superpixel segmentation algorithm proposed by Achanta et al. (2012) with the advantages of a simple calculation process, high computation speed, and good edge matching. First, it converts the image from RGB to CIELAB, in which the five-dimensional vector $V [l, a, b, x, y]$ consists of a (l, a, b) color value and (x, y) coordinates of the corresponding pixel. Then based on the idea of K means, k superpixels are initialized in an image, and the distance between them is set as S . The core part is to iteratively calculate the centers of these superpixels by a clustering method. The distance for five-dimensional vectors (D') includes the distance of the CIELAB color space (d_c) and the geometric space (d_s). The following formulas are used to calculate the distance:

$$d_c = \sqrt{(l_j - l_i)^2 + (a_j - a_i)^2 + (b_j - b_i)^2}, \quad (3)$$

$$d_s = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2}, \quad (4)$$

$$D' = \sqrt{(d_c/m)^2 + (d_s/s)^2}, \quad (5)$$

where m indicates the maximum possible distance in the CIELAB color space and s indicates the maximum possible value in the geometric space.

For each superpixel center, the range of pixel searching is $2S \times 2S$. If the distance from a pixel to the superpixel center i is less than the distance from it to the superpixel center to which it previously belonged, then this pixel is assigned to the superpixel i . In this iterative optimization algorithm, optimization persists until pixel distances to the new and previous superpixel centers stabilize. After iterating out the superpixel segmentation blocks of the image, the semantic segmentation results and SLIC segmentation are fused based on a rule. First, count the number of pixels with different semantic segmentation labels in a superpixel. Then the semantic label with more pixels is added to all pixels of this superpixel segmentation block.

3.2.2 DenseCRF

DenseCRF overcomes the limitation that CRF can only be performed in a local area and cannot connect full-text information (L. Zhang et al., 2018). The global context information of the whole image is organically combined, and all the pixels in the entire image are connected with the current pixel. DenseCRF is composed of unary potentials and pairwise potentials. Unary potentials come from the output of the

front-end semantic segmentation network, which refers to the potential of predicting the pixel point (i) as a semantic label (x_i) through the semantic segmentation network. Pairwise potentials describe the relationship of each pixel to all other pixels in the image, mainly providing position and spectral information through the original input image (Berman et al., 2018). Therefore, it not only makes predictions for a single pixel but also calculates the probability of different classes appearing simultaneously.

3.3 Accuracy assessment indicators

Scientific selection of evaluation indicators is the key to testing the accuracy of glacial lake extraction results. Four indicators – recall, precision, F1 score (Yacouby and Axman, 2020), and IoU (Rahman and Wang, 2016) – were selected as the indicators for the accuracy evaluation of glacial lake extraction results. And all of them are generated based on the true positive (TP), true negative (TN), false positive (FP), and false negative (FN).

4 Experiments and results

Based on high-resolution Google Earth images (level 18, with an image resolution of 0.52 m) in the study area, we used L_1 -LinkNet50 for the preliminary extraction in the rough extraction stage. The commonly used semantic segmentation models (U-Net, LinkNet) and traditional machine learning methods (SVM and RF) were chosen for comparison. Moreover, among the improved semantic segmentation methods for glacial lakes, EfficientNet U-Net (Qayyum et al., 2020) was proposed based on high-resolution images (3–4 m) and was also chosen for comparison. Our experiment was based on Python 3.6 and the open-source deep learning framework PyTorch. In the SVM classifier, the penalty coefficient (C) was set to 12, the kernel was set to the radial basis function (rbf), and the gamma in the kernel function was set to 0.187. In the RF classifier, the number of decision trees was set to 150, and the number of features is set to 2. The training was performed on an NVIDIA GeForce RTX 2080 Ti, using cuDNN10.0 for acceleration. The batch size was set to 4. For the semantic segmentation network, the optimization method adopted the Adaptive Moment Estimation (Adam), and the initial learning rate was set to 1×10^{-4} . Moreover, the learning rate update strategy of polynomial decay was adopted to prevent the network from sinking into local optimal solutions later in model training, in which the momentum and weight decay were set at 0.9 and 1×10^{-4} , respectively. This study trained all networks for 45 epochs in this stage.

4.1 Comparative analysis of rough extraction results

Recall, precision, F1 score, and IoU were calculated to evaluate the extraction results against the ground truth obtained

Table 4. Meanings and calculation formulas of four evaluation indicators.

Evaluating indicators	Formulas	Meanings
Recall	$TP/(TP + FN)$	The proportion of the number of correctly identified samples to the number of all positive samples.
Precision	$TP/(TP + FP)$	The proportion of the number of correctly identified samples to the number of predicted positive samples.
F1 score	$2 \cdot \text{precision} \cdot \text{recall}/(\text{precision} + \text{recall})$	The harmonic average of recall rate and accuracy rate.
IoU	$TP/(TP + FN + FP)$	The ratio of intersection and union of the real value and predicted value.

by manual digitization. As can be seen in Fig. 5, the edges of the glacial lakes extracted by SVM and RF are rough. There are difficulties in the complete extraction of complex glacial lakes and small glacial lakes, which affects the recall. In deep learning models, the U-Net network is greatly affected by snow (Fig. 5e), and the probability of being wrongly classified as glacial lakes is high, which decreases the precision and F1 score. EfficientNet U-Net obtains the highest precision, and the predicted water masks have few false positives, which is consistent with the conclusions of Qayyum et al. (2020). This method reduced false detections of glacial lakes in snow-covered areas compared to U-Net (Fig. 5e). However, there are still problems for glacial lakes with similar spectral information to the background and for glacial lakes that are shaded by mountain shadows. LinkNet can identify more glacial lakes than EfficientNet U-Net. But some false detections are prone to occur in the shaded area (Fig. 5c), and the precision is reduced. Finally, after introducing the deep residual network, L_1 -LinkNet50 improved the extraction of glacial lakes with small areas and glacial lakes shaded by mountain shadows. Although the precision is slightly lower than that of EfficientNet U-Net, the final F1 score of L_1 -LinkNet50 reaches 87.77 % and the recall is 3.46 % higher than that of EfficientNet U-Net (Table 5). Therefore, it can be found that L_1 -LinkNet50 has the most vital comprehensive ability for glacial lake extraction in these models.

4.2 Comparative analysis of optimization results

L_1 -LinkNet50 showed the best effect to extract glacial lakes in Sect. 4.1. This study improved L_1 -LinkNet50 and carried out post-processing on the glacial lake segmentation results. L_1 -LinkNet50 was trained for another 25 epochs for L_{12} -LinkNet50, using the Lovasz hinge loss (Loss2) as the loss function. Then the two-level optimization strategy (SLIC and DenseCRF) was used to optimize semantic segmentation results. For superpixel segmentation, the number of superpixels, compactness, and iteration times were set as 2800, 60,

Table 5. Quantitative evaluation for glacial lake extraction.

Method	Recall	Precision	F1 score	IoU
SVM	73.61 %	88.28 %	80.28 %	67.05 %
RF	74.38 %	89.32 %	81.17 %	68.31 %
U-Net	79.88 %	80.50 %	80.19 %	70.33 %
EfficientNet U-Net	81.04 %	92.08 %	85.72 %	77.70 %
LinkNet	83.65 %	88.44 %	85.97 %	75.02 %
L_1 -LinkNet50	84.50 %	91.31 %	87.77 %	78.21 %

The bold font represents the highest score.

and 10, respectively. Then the semantic segmentation result by L_{12} -LinkNet50 or the fusion result by SLIC and L_{12} -LinkNet50 was input into DenseCRF as the unary potential. In this process, the mean-field approximation method was used for inference to minimize the potential function.

In Table 6, after L_1 -LinkNet50 was trained with the Lovasz hinge loss, the IoU reached 83.77 % in the study area, which was 5.56 % higher than that of L_1 -LinkNet50. It alleviated the problems of adhesions (multiple glacial lakes nearby detected as one) (Fig. 6 (3) and (4)) and missed detections (Fig. 6 (2)). In addition, it is difficult for the semantic segmentation network to recover all the lost spatial information when upsampling, resulting in imprecise segmentation edges. The glacial lakes after superpixel segmentation optimization are closer to the natural boundary, especially the small glacial lakes. Post-processed results by DenseCRF had smoother edges, and the precision increased by 0.86 %. Moreover, after using two-level optimization (SLIC–DenseCRF), missed detections of glacial lakes were effectively reduced. Compared to L_{12} -LinkNet50, the IoU and F1 score increased by 6.92 % and 3.96 %, respectively. The comparison of results based on different optimization algorithms proves that the post-processing based on SLIC–DenseCRF for deep learning semantic segmentation results can improve the accuracy of glacial lake extraction.

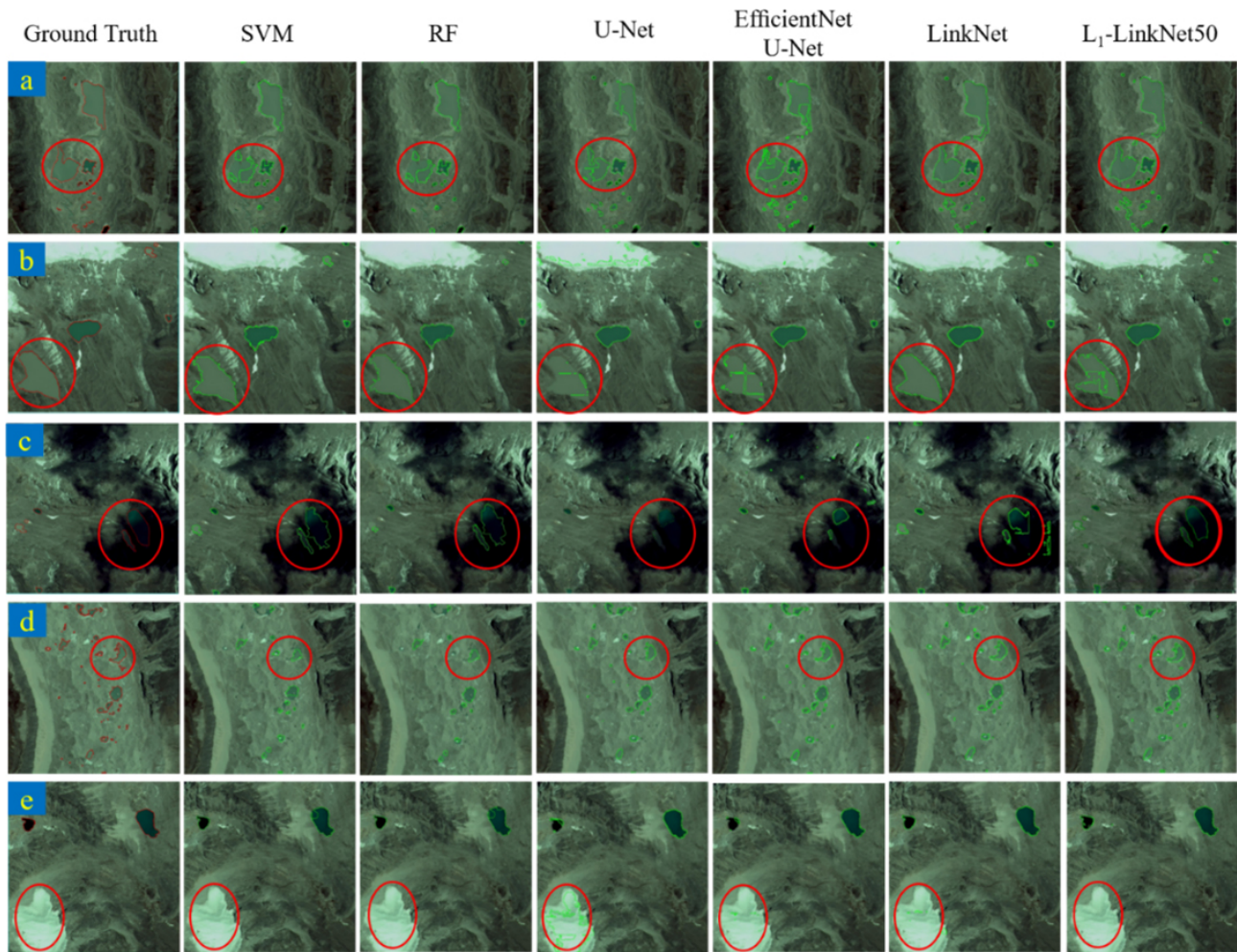


Figure 5. Performance comparison of different models for the glacial lake extraction. All images © Google Earth 2020. Five regions of the same size (2048×2048) were chosen based on Google Earth images (0.52 m). The red shapes are the boundary of the ground truth glacial lakes, and the green shapes are the boundary of the predicting results, including areas with glacial lakes of complex outlines (a), the inconsistent color of water bodies (b), mountain shadows (c), and areas with multiple small glacial lakes (d) and ice and snow (e).

Table 6. Evaluation indicators for glacial lake identification results under different optimization conditions.

Method	Recall	Precision	F1 score	IoU
L_1 -LinkNet50	84.50 %	91.31 %	87.77 %	78.21 %
L_{12} -LinkNet50	88.26 %	92.85 %	90.50 %	83.77 %
L_{12} -LinkNet50-SLIC	92.66 %	89.77 %	91.19 %	85.11 %
L_{12} -LinkNet50-DenseCRF	89.38 %	93.71 %	91.49 %	84.23 %
L_{12} -LinkNet50-SLIC-DenseCRF	96.52 %	92.49 %	94.46 %	90.69 %

The bold font represents the highest score.

4.3 Glacial lake extraction in Qomolangma National Nature Reserve

After evaluating the ability of the glacial lake extraction method proposed in this study, we applied L_{12} -LinkNet50-SLIC-DenseCRF to the extraction of glacial lakes in the Qo-

molangma National Nature Reserve (QNNR). This reserve has a total area of $33\,819\text{ km}^2$, including the core area, the buffer area, and the experimental area. Because it is difficult to obtain the sub-meter-level Google Earth image of the entire QNNR, Google Earth images (level 16) in 2020 within the 10 km buffer zone from the end of the glacier in QNNR

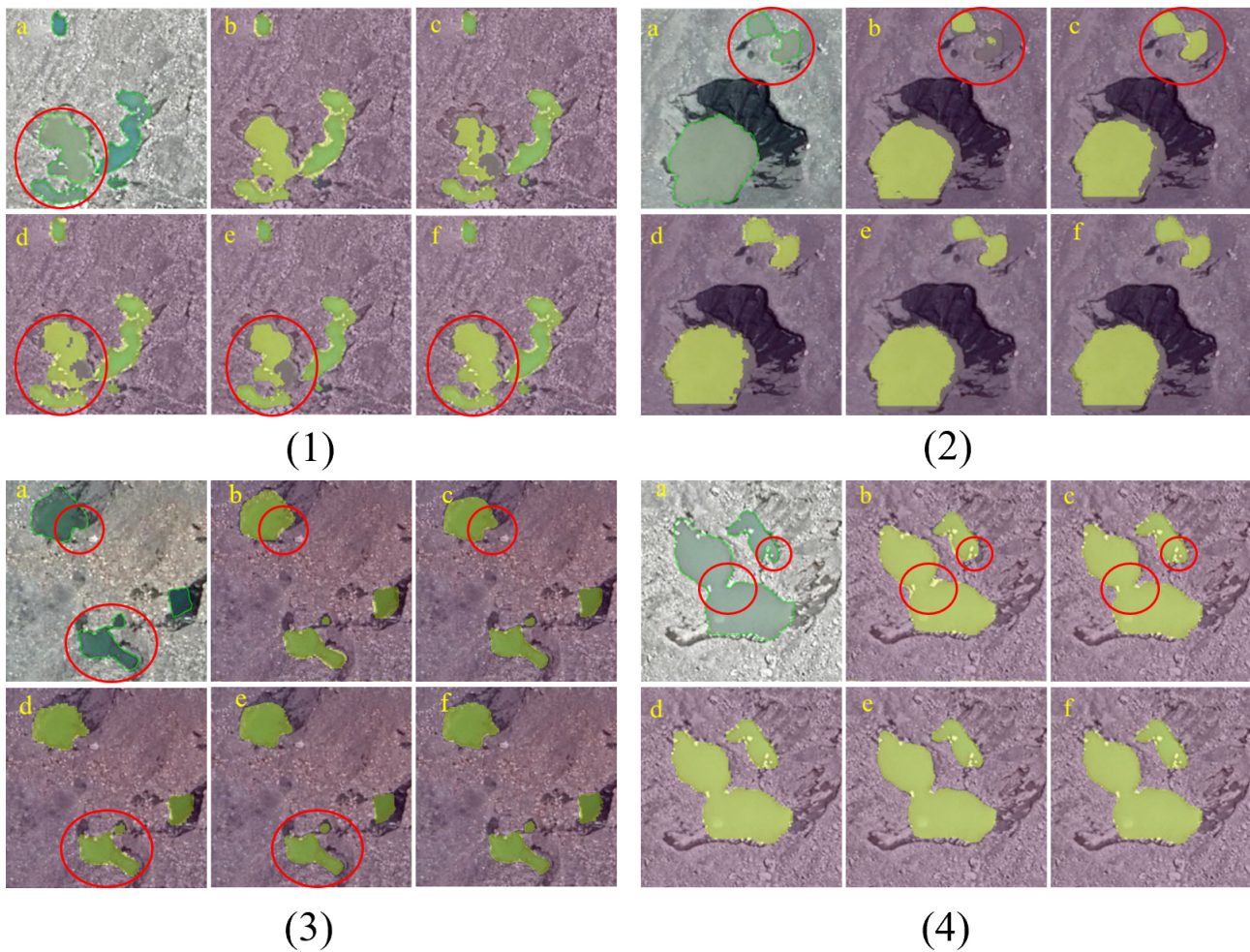


Figure 6. Comparison of glacial lake identification results based on Google Earth images (0.52 m) under different optimization conditions. All images © Google Earth 2020. Ground truth (a), L_1 -LinkNet50 (b), L_{12} -LinkNet50 (c), L_{12} -LinkNet50-SLIC (d), L_{12} -LinkNet50-DenseCRF (e), and L_{12} -LinkNet50-SLIC-DenseCRF (f). The green shapes are the boundary of the ground truth glacial lakes, and the yellow masks are the predicting results.

were used as the data source, with an image resolution of 2.11 m. For the glacial lake extraction in QNNR, the evaluation results show that the precision, F1 score, recall, and IoU are 85.55 %, 82.49 %, 79.65 %, and 70.20 %, respectively. The IoU of the actual application is lower than the rectangular study area, but the precision has reached more than 85 %. The final result is shown in Fig. 7. In the QNNR, glacier lakes are mainly distributed in the altitude range of 4000 to 6000 m. The area and number of glacial lakes are approximately normally distributed, and both peak at 5000–5500 m, which is consistent with the research of Yang et al. (2019) and Zhang et al. (2021). The area and number of glacial lakes at the peak account for 66.58 % and 39.70 % of all glacial lakes, respectively.

Compared with the existing glacial lake inventory and three land cover datasets in Sect. 2.2.2, the glacial lakes extracted based on the proposed method are closer to the

ground truth in terms of the number and area (Fig. 8a). The area of the largest glacial lake (5.943 km²) of the reserve extracted in this study is consistent with the other four and is closest to the ground truth. For glacial lakes with an area greater than 0.01 km², the distribution of the number of glacial lakes is similar for all datasets. However, for small glacial lakes, due to the advantages of remote sensing image sources and methods, the accuracy of the extraction results of glacial lakes in this study is significantly better than the other four existing datasets. Moreover, we checked the results of the glacial lake extraction in the QNNR and found that the smallest glacial lake that can be fully and correctly extracted has an area of 160 m².

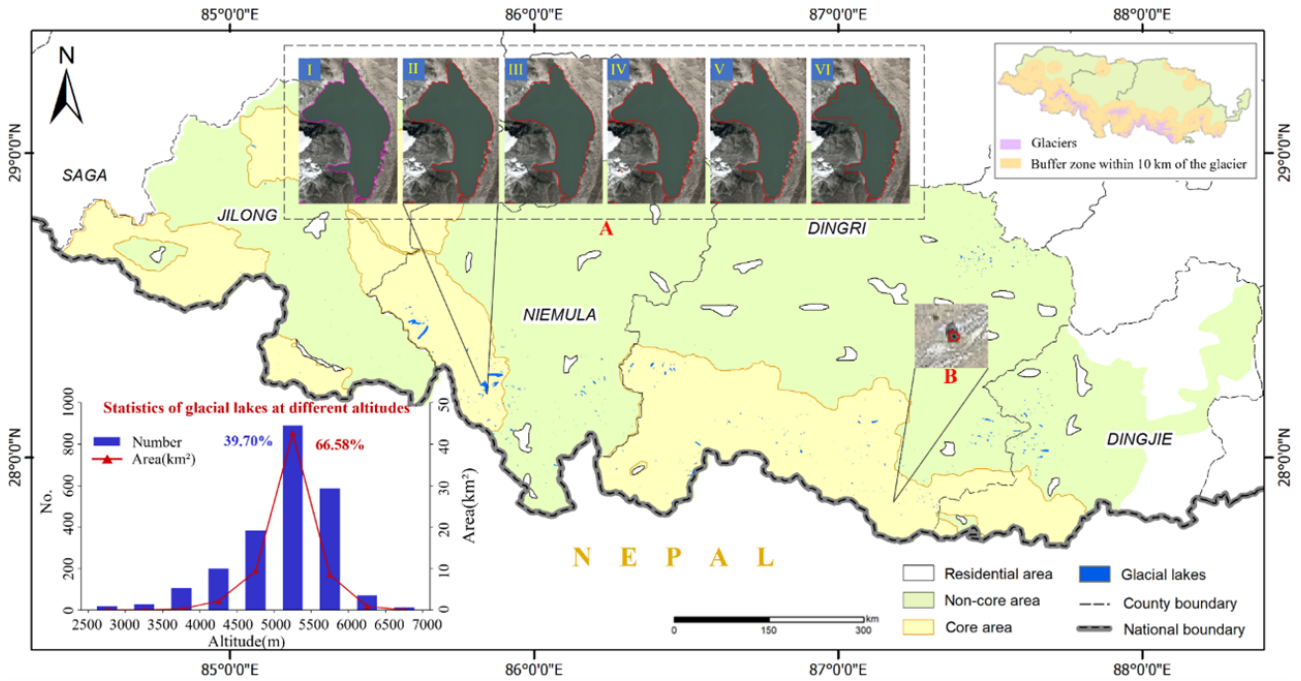


Figure 7. Glacial lake extraction result in the QNNR based on Google Earth images in 2020 (2.11 m). All images © Google Earth 2020. A and B are the largest and smallest glacial lakes extracted in this region, respectively. The purple shapes are the boundary of the ground truth glacial lakes, and the red shapes are the boundary of the predicting results. Ground truth (I), L_{12} -LinkNet50–SLIC–DenseCRF 2.11 m (II), inventory data of glacial lake 30 m (III), ESA WorldCover 10 m (IV), Esri Land Cover 10 m (V), and FROM-GLC10 10 m (VI).

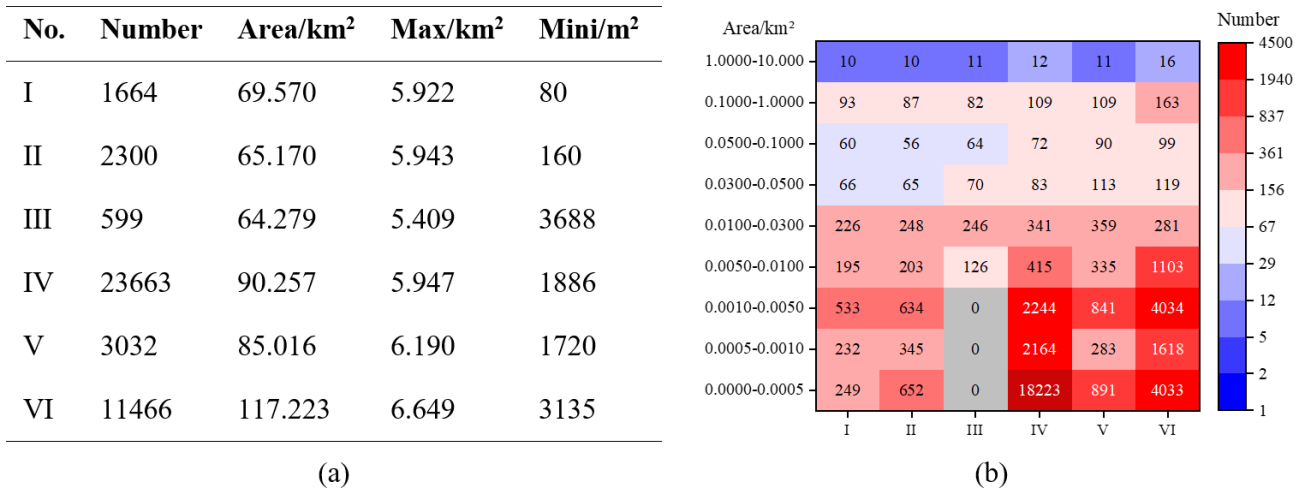


Figure 8. Comparison of glacial lakes between this study and the other four datasets. Basic information (a) and statistical information on the number distribution of glacial lakes in different areas (b). Ground truth (I), L_{12} -LinkNet50–SLIC–DenseCRF 2.11 m (II), inventory data of glacial lake 30 m (III), ESA WorldCover 10 m (IV), Esri Land Cover 10 m (V), and FROM-GLC10 10 m (VI).

5 Discussion

The lack of reliable glacial lake samples is one of the difficulties in the development of glacial lake extraction research based on deep learning networks (J. Wang et al., 2022). Qayyum et al. (2020) also stated that because of insufficient types of glacial lakes in the training dataset, some

muddy brown glacial lakes could not be identified. The high-resolution samples built in this study help to improve the evaluation indicators. Moreover, L_{12} -LinkNet50 uses a deep residual structure (ResNet50) as the backbone of the network, which enhances the ability to extract glacial lakes under different background conditions and is also better at the boundary of the small glacial lake (Fig. 9b). Therefore, com-

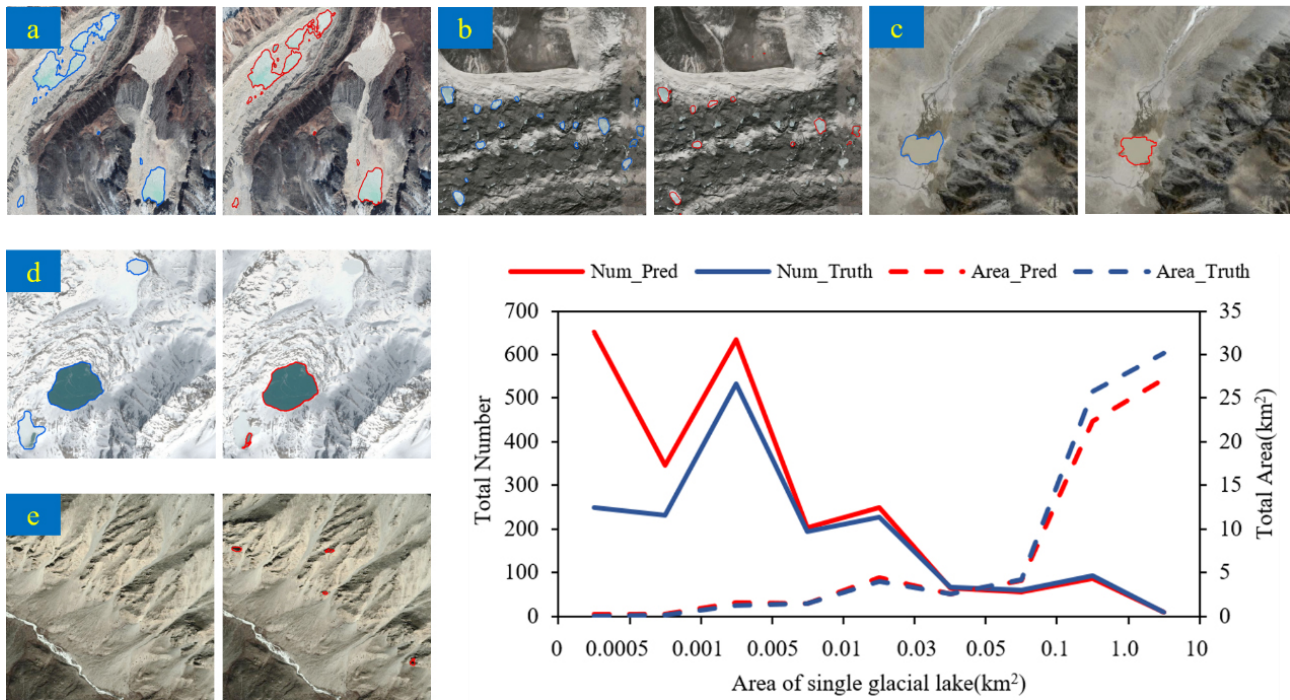


Figure 9. The display of glacial lake extraction results in different situations in the QNNR, as well as the statistics of area and number. All images © Google Earth 2020. The blue shapes are the boundary of the ground truth glacial lakes, and the red shapes are the boundary of the predicting results.

pared with other methods as shown in Sect. 4.1, the evaluation indicators of glacial lakes extracted by L_{12} -LinkNet50 are improved. The proposed method in this study can effectively reduce missed detections of some glacial lakes that show similar spectral features with soil (Fig. 9c) and shadows.

For post-processing, the parameter values used in the SLIC algorithm in this study, including the number of superpixel blocks (2800) and the compactness (60), were obtained through multiple experiments by the control variable method based on sub-meter-level images. When the image resolution differs greatly, the amount of semantic information in a single superpixel will change. Thus, these parameter values are not applicable to images with a spatial resolution of 10 m such as band 2, 3, and 4 of Sentinel images. In addition, frozen lakes generally start from the edge with shallow water bodies and more small rocks, resulting in more noise on the edge of glacial lakes. DenseCRF connects the local and global information to set up pairwise potentials on all pairs of pixels, providing more detailed labeling and reducing the small-area noise generated by the image segmentation of high-resolution images. As a result, the optimized glacial lakes have smoother edges (Fig. 9a) and fewer false spots on the lake surface.

For the glacial lake extraction results of the QNNR, the curve in Fig. 9 shows that the area distribution of small-area glacial lakes is consistent with the results of manual digitiza-

tion. However, although the proposed method is effective in identifying glacial lakes with similar spectral information to shadows, it is prone to misjudge in small areas of shadows (Fig. 9e). Because the area of these shadows is too small, little spectral and texture information on the background can be extracted, so it is difficult to distinguish by the method in this study. The number distribution of large-scale glacial lakes is consistent with the results of manual digitization, but more large-area glacial lakes have not been fully identified. The spectral information of glacial lakes completely covered by snow and some glacial lakes that have been frozen for a long time is very similar to that of snow. At present, the proposed method in this study cannot fully identify those glacial lakes (Fig. 9d). In the Google images captured over the QNNR area, large areas of the land were covered with snow. This is also the reason why glacial lake evaluation indicators of the QNNR are lower than those of the small study area in Sect. 4.2.

6 Conclusions

Aiming at the demand for high-accuracy outline extraction of glacial lakes in the high-Asia region, this study built a dataset for glacial lakes based on the global meter-level to sub-meter-level Google Earth images and then proposed the glacial lake extraction method of the L_{12} -LinkNet50 se-

semantic segmentation network with two-level optimization of SLIC–DenseCRF.

Based on the dataset containing glacial lakes of multiple types, the ability to identify glacial lakes of different types and colors is improved in this study. The Lovasz hinge loss and 0.5 BCE + dice are combined to improve the loss function of the deep semantic segmentation network and suppress the impact of the unbalanced positive and negative samples in the dataset. It has the advantage of small glacial lake detection and effectively reduces the missed detection of glacial lakes that have similar spectral features to the bare soil or shadows. The F1 score in the study area reaches more than 90%. At the same time, it is applied to a wider range of QNNR. Due to the misjudgment of the small glacial lake in the shadow, the F1 score is reduced, but it also reaches 82.49%. By post-processing for the semantic segmentation results, the edges of glacial lakes are more consistent with the actual situation, and the noise spots on the lake surface are also reduced.

Although the proposed method has achieved good extraction results on the new dataset, there are still shortcomings in the recognition of snow-covered glacial lakes and terrain shadows with small areas. For future research, multi-source remote sensing images can be used to reduce the impact of snow cover and shadows.

Data availability. All raw data can be provided by the corresponding authors upon request.

Author contributions. YC and MP designed the experiments; XB, RP, RL, and PD prepared experimental data; XB and MP developed the code; YC, RP, and XB performed the data analysis; YC, XB, and RP wrote a draft of the manuscript; YC, XB, RP, and MP reviewed and edited the manuscript; and all authors contributed to correcting and editing the final version.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors.

Acknowledgements. This research was funded by the National Natural Science Foundation of China (grant no. 41771451) and the Sichuan Province Youth Science and Technology Innovation Team (grant no. 2020JDTD0003). We thank Peng Gong et al., ESA, and ESRI for providing the land cover products (FROM-GLC10, ESA

WorldCover, and Esri Land Cover). We also thank the TPDC for the second glacier inventory dataset of China and the inventory data of the glacial lake in west China.

Financial support. This research has been supported by the National Natural Science Foundation of China (grant no. 41771451) and the Sichuan Province Youth Science and Technology Innovation Team (grant no. 2020JDTD0003).

Review statement. This paper was edited by Homa Kheyrollah Pour and reviewed by Connor Shiggins and one anonymous referee.

References

- Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., and Süsstrunk, S.: SLIC superpixels compared to state-of-the-art superpixel methods, *IEEE T. Pattern Anal.*, 34, 2274–2282, <https://doi.org/10.1109/TPAMI.2012.120>, 2012.
- Begam, S., Sen, D., and Dey, S.: Moraine dam breach and glacial lake outburst flood generation by physical and numerical models, *J. Hydrol.*, 563, 694–710, <https://doi.org/10.1016/j.jhydrol.2018.06.038>, 2018.
- Berman, M., Triki, A. R., and Blaschko, M. B.: The lovasz-softmax loss: A tractable surrogate for the optimization of the intersection-over-union measure in neural networks, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 18–22 June 2018, Salt Lake City, United States, 4413–4421, <https://doi.org/10.1109/CVPR.2018.00464>, 2018.
- Bourgeois, J., Cisternas, M. E., Braucher, R., Bourlès, D., and Frutos, J.: Geomorphic records along the general Carrera (Chile)–Buenos Aires (Argentina) glacial lake (46–48 S), climate inferences, and glacial rebound for the past 7–9 ka, *J. Geol.*, 124, 27–53, <https://doi.org/10.1086/684252>, 2016.
- Chaurasia, A. and Culurciello, E.: Linknet: Exploiting encoder representations for efficient semantic segmentation, *IEEE Visual Communications and Image Processing VCIP*, 1–4, <https://doi.org/10.1109/VCIP.2017.8305148>, 2017.
- Chen, F.: Comparing Methods for Segmenting Supra-Glacial Lakes and Surface Features in the Mount Everest Region of the Himalayas Using Chinese GaoFen-3 SAR Images, *Remote Sens.*, 13, 2429, <https://doi.org/10.3390/rs13132429>, 2021.
- Chen, F., Zhang, M., Guo, H., Allen, S., Kargel, J. S., Haritashya, U. K., and Watson, C. S.: Annual 30 m dataset for glacial lakes in High Mountain Asia from 2008 to 2017, *Earth Syst. Sci. Data*, 13, 741–766, <https://doi.org/10.5194/essd-13-741-2021>, 2021.
- Chen, Y., Fan, R., Yang, X. C., Wang, J. X., and Latif, A.: Extraction of urban water bodies from high-resolution remote-sensing imagery using deep learning, *Water*, 10, 585, <https://doi.org/10.3390/w10050585>, 2018.
- Cordeiro, M. C., Martinez, J. M., and Peña-Luque, S.: Automatic water detection from multidimensional hierarchical clustering for Sentinel-2 images and a comparison with Level 2A processors, *Remote Sens. Environ.*, 253, 112209, <https://doi.org/10.1016/j.rse.2020.112209>, 2021.
- Gong, P., Liu, H., Zhang, M. M., Li, C. C., and Wang, J.: Stable classification with limited sample: transferring a 30-m res-

- olution sample collection collected in 2015 to mapping 10-m resolution global land cover in 2017, *Sci. Bull.*, 64, 370–3734, <https://doi.org/10.1016/j.scib.2019.03.002>, 2019.
- He, K. M., Zhang, X. Y., Ren, S. Q., and Sun, J.: Deep residual learning for image recognition, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 26 June–1 July 2016, Las Vegas, United States, 770–778, <https://doi.org/10.1109/CVPR.2016.90>, 2016.
- He, Y., Yao, S., Yang, W., Yan, H. W., Zhang, L. F., Wen, Z. Q., and Liu, T.: An extraction method for glacial lakes based on Landsat-8 imagery using an improved U-Net network, *IEEE J. Sel. Top. Appl. Earth Obs.*, 14, 6544–6558, <https://doi.org/10.1109/JSTARS.2021.3085397>, 2021.
- Huggel, C., Käab, A., Haerberli, W., Teyssere, P., and Paul, F.: Remote sensing based assessment of hazards from glacier lake outbursts: a case study in the Swiss Alps, *Can. Geotech. J.*, 39, 316–330, <https://doi.org/10.1139/t01-099>, 2002.
- Jain, S. K., Sinha, R. K., Chaudhary, A., and Shukla, S.: Expansion of a glacial lake, Tsho Chubda, Chamkhar Chu Basin, Hindukush Himalaya, Bhutan, *Can. Geotech. J.*, 75, 1451–1464, <https://doi.org/10.1007/s11069-014-1377-z>, 2015.
- Ke, L., Ding, X., Zhang, L. E. I., Hu, J. U. N., Shum, C. K., and Lu, Z.: Compiling a new glacier inventory for southeastern Qinghai–Tibet Plateau from Landsat and PALSAR data, *J. Glaciol.*, 62, 579–592, <https://doi.org/10.1017/jog.2016.58>, 2016.
- Li, D., Shangquan, D. H., and Huang, W. D.: Study on the area change of lakes Merzbacher in the Tianshan mountains during 1998–2017, *J. Glaciol. Geocryol.*, 42, 1126–1134, <https://doi.org/10.7522/j.issn.1000-0240.2019.0308>, 2020.
- Li, S. and Liu, X.: Multi-type road extraction and analysis of high-resolution images with D-LinkNet50, in: 2022 3rd International Conference on Geology, Mapping and Remote Sensing (ICGMRS), 244–248, <https://doi.org/10.1109/ICGMRS55602.2022.9849390>, 2022.
- Li, Z. Y., Wang, R., Zhang, W., Hu, F. M., and Meng, L.: Multiscale features supported deeplabv3+ optimization scheme for accurate water semantic segmentation, *IEEE Access*, 7, 155787–155804, <https://doi.org/10.1109/ACCESS.2019.2949635>, 2019.
- Liu, S. Y., Guo, W. Q., and Xu, J. L.: The second glacier inventory dataset of China (version 1.0) (2006–2011), National Tibetan Plateau/Third Pole Data Center [data set], <https://doi.org/10.3972/glacier.001.2013.db>, 2012.
- Nie, Y., Sheng, Y. W., Liu, Q., Liu, L. S., Liu, S. Y., Zhang, Y. L., and Song, C. Q.: A regional-scale assessment of Himalayan glacial lake changes using satellite observations from 1990 to 2015, *Remote Sens. Environ.*, 189, 1–13, <https://doi.org/10.1016/j.rse.2016.11.008>, 2017.
- Nie, Y., Deng, Q., Pritchard, H. D., Carrivick, J. L., Ahmed, F., Huggel, C., Liu, L. J., Wang, W., Lesi, M., Wang, J., Zhang, H., Zhang, B., Lü, Q., and Zhang Y.: Glacial lake outburst floods threaten Asia's infrastructure, *Sci. Bull.*, 68, 1361–1365, <https://doi.org/10.1016/j.scib.2023.05.035>, 2023.
- Pandey, P., Ali, S. N., and Champati Ray, P. K.: Glacier-glacial lake interactions and glacial lake development in the central Himalaya, India (1994–2017), *J. Earth Sci.*, 32, 1563–1574, <https://doi.org/10.1007/s12583-020-1056-9>, 2021.
- Qayyum, N., Ghuffar, S., Ahmad, H. M., Yousaf, A., and Shahid, I.: Glacial lakes mapping using multi satellite PlanetScope imagery and deep learning, *ISPRS Int. J. Geoinf.*, 9, 560, <https://doi.org/10.3390/ijgi9100560>, 2020.
- Qi, W., Zhang, Y., Gao, J., Yang, X., Liu, L., and Khanal, N. R.: Climate change on the southern slope of Mt. Qomolangma (Everest) Region in Nepal since 1971, *J. Geogr. Sci.*, 23, 595–611, <https://doi.org/10.1007/s11442-013-1031-9>, 2013.
- Rahman, M. A. and Wang, Y.: Optimizing intersection-over-union in deep neural networks for image segmentation, *International symposium on visual computing*, 234–244, https://doi.org/10.1007/978-3-319-50835-1_22, 2016.
- Rick, B., McGrath, D., Armstrong, W., and McCoy, S. W.: Dam type and lake location characterize ice-marginal lake area change in Alaska and NW Canada between 1984 and 2019, *The Cryosphere*, 16, 297–314, <https://doi.org/10.5194/tc-16-297-2022>, 2022.
- Sathananthavathi, V. and Indumathi, G.: Encoder enhanced atrous (EEA) unet architecture for retinal blood vessel segmentation, *Cogn. Syst. Res.*, 67, 84–95, <https://doi.org/10.1016/j.cogsys.2021.01.003>, 2021.
- Song, C., Sheng, Y., Ke, L., Nie, Y., and Wang, J.: Glacial lake evolution in the southeastern Tibetan Plateau and the cause of rapid expansion of proglacial lakes linked to glacial-hydrogeomorphic processes, *J. Hydrol.*, 540, 504–514, <https://doi.org/10.1016/j.jhydrol.2016.06.054>, 2016.
- Song, J., Gao, S., Zhu, Y., and Ma, C.: A survey of remote sensing image classification based on CNNs, *Big Earth Data*, 3, 232–254, <https://doi.org/10.1080/20964471.2019.1657720>, 2019.
- Talal, M., Panthakkan, A., Mukhtar, H., Mansoor, W., Almansoori, S., and Al Ahmad, H.: Detection of water-bodies using semantic segmentation, *International Conference on Signal Processing and Information Security*, 7–8 November 2020, United Arab Emirates, Dubai, <https://doi.org/10.1109/CSPIS.2018.8642743>, 2018.
- Taylor, C., Robinson, T. R., Dunning, S., Carr, J. R., and Wsetoby, M.: Glacial lake outburst floods threaten millions globally, *Nat. Commun.*, 14, 487, <https://doi.org/10.1038/s41467-023-36033-x>, 2023.
- Veh, G., Korup, O., Roessner, S., and Walz, A.: Detecting Himalayan glacial lake outburst floods from Landsat time series, *Remote Sens. Environ.*, 207, 84–97, <https://doi.org/10.1016/j.rse.2017.12.025>, 2018.
- Wang, J., Chen, F., Zhang, M., and Yu, B.: NAU-Net: A New Deep Learning Framework in Glacial Lake Detection, *IEEE Geosci. Remote Sensing Lett.*, 19, 1–5, <https://doi.org/10.1109/LGRS.2022.3165045>, 2022.
- Wang, J. X., Chen, F., Zhang, M. M., and Yu, B.: ACFNet: A Feature Fusion Network for Glacial Lake Extraction Based on Optical and Synthetic Aperture Radar Images, *Remote Sens.*, 13, 5091, <https://doi.org/10.3390/rs13245091>, 2021.
- Wang, R., Meng, Y., Zhang, W., Li, Z., Hu, F., and Meng, L.: Remote sensing semantic segregation for water information extraction: Optimization of samples via training error performance, *IEEE Access*, 7, 13383–13395, <https://doi.org/10.1109/ACCESS.2019.2894099>, 2019.
- Wang, S. D., Peppia, M. V., Xiao, W., Maharjan, S. B., Joshi, S. P., and Mills, J. P.: A second-order attention network for glacial lake segmentation from remotely sensed imagery, *ISPRS J. Photogramm. Remote Sens.*, 189, 289–301, <https://doi.org/10.1016/j.isprs.2022.05.007>, 2022.

- Wang, X.: Inventory data of glacial lake in west China (2015), National Tibetan Plateau/Third Pole Environment Data Center [data set], <https://doi.org/10.11922/sciencedb.615>, 2018.
- Wang, Z., Gao, X., Zhang, Y., and Zhao, G.: MSLWENet: A novel deep learning network for lake water body extraction of Google remote sensing images, *Remote Sens.*, 12, 4140, <https://doi.org/10.3390/rs12244140>, 2020.
- Yacouby, R. and Axman, D.: Probabilistic extension of precision, recall, and F1 score for more thorough evaluation of classification models, in: Proceedings of the first workshop on evaluation and comparison of NLP systems, Online, 20 November 2020, 79–91, <https://doi.org/10.18653/v1/2020.eval4nlp-1.9>, 2020.
- Yang, C. D., Wang, X., Wei, J. F., Liu, Q. H., Lu, A. X., Zhang, Y., and Tang, Z. G.: Chinese glacial lake inventory based on 3S technology method, *J. Geogr. Sci.*, 74, 544–556, <https://doi.org/10.11821/dlxb201903011>, 2019.
- Yao, X. J., Liu, S. Y., Han, L., and Sun, M. P.: Definition and classification systems of glacial lake for inventory and hazards study, *J. Geogr. Sci.*, 72, 1173–1184, <https://doi.org/10.11821/dlxb201707004>, 2017.
- Zhang, L., Li, H., Shen, P. Y., Zhu, G. M., Song, J., Shah, S. A. A., Bennamoun, M., and Zhang, L.: Improving Semantic Image Segmentation with a Probabilistic Superpixel-Based Dense Conditional Random Field, *IEEE Access*, 6, 15297–15310, <https://doi.org/10.1109/ACCESS.2018.2814568>, 2018.
- Zhang, M. M., Chen, F., and Tian, B. S.: An automated method for glacial lake mapping in High Mountain Asia using Landsat 8 imagery, *J. Mt. Sci.*, 15, 13–24, <https://doi.org/10.1007/s11629-017-4518-5>, 2018.
- Zhang, M. M., Chen, F., Zhao, H., Wang J. X., and Wang, N.: Recent Changes of Glacial Lakes in the High Mountain Asia and Its Potential Controlling Factors Analysis, *Remote Sens.*, 13, 3537, <https://doi.org/10.3390/rs13183757>, 2021.
- Zhang, M. M., Chen, F., Guo, H. D., Yi, L., Zeng, J. Y., and Li, B.: Glacial Lake Area Changes in High Mountain Asia during 1990–2020 Using Satellite Remote Sensing, *Research*, 2022, 9821275, <https://doi.org/10.34133/2022/9821275>, 2022.
- Zhang, T. G., Wang, W. C., Gao, T. G., An, B. S., and Yao, T. D.: An integrative method for identifying potentially dangerous glacial lakes in the Himalayas, *Sci. Total Environ.*, 806, 150442, <https://doi.org/10.1016/j.scitotenv.2021.150442>, 2022.
- Zhao, H., Chen, F., and Zhang, M. M.: A systematic extraction approach for mapping glacial lakes in high mountain regions of Asia, *IEEE J. Sel. Top. Appl. Earth Obs.*, 11, 2788–2799, <https://doi.org/10.1109/JSTARS.2018.2846551>, 2018.
- Zheng, G., Mergili, M., Emmer, A., Allen, S., Bao, A., Guo, H., and Stoffel, M.: The 2020 glacial lake outburst flood at Jinwuco, Tibet: causes, impacts, and implications for hazard and risk assessment, *The Cryosphere*, 15, 3159–3180, <https://doi.org/10.5194/tc-15-3159-2021>, 2021.
- Zhong, Y., Liu, Q., Sapkota, L., Luo, Y., Wang, H., Liao, H., and Wu, Y.: Rapid glacier Shrinkage and Glacial Lake Expansion of a China-Nepal Transboundary Catchment in the Central Himalayas, between 1964 and 2020, *Remote Sens.*, 13, 3614, <https://doi.org/10.3390/rs13183614>, 2021.