GLACIER IDENTIFICATION FROM LANDSAT8 OLI IMAGERY USING DEEP U-NET

Q. He [1], Z. Zhang [2], G. Ma [1, *], J. Wu [1]

[1] State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, 430079, Wuhan, Hubei, China – hqj_mail@163.com, mgr@whu.edu.cn, 1250230255@qq.com
[2] Xining Comprehensive Survey Center for Natural Resources, China Geological Survey, 810000, Xining, Qinghai, China – zhangzhijun-0001@163.com

Commission III, WG III/9

KEY WORDS: Glacier identification, Landsat8 OLI, U-Net, NDSI, Semantic segmentation, Remote sensing

ABSTRACT:

Glacier is one of the clearest signal of climate change, and its changes have important effects on regional climate and water resources. Glacier identification is the basic of glacial changes research. Traditional remote sensing glacier identification methods usually perform simple bands calculation based on the spectral characteristics of glacier. The identification results are greatly affected by threshold segmentation. In addition, there is a misclassification of water body and glacier. As a simple and efficient semantic segmentation network, U-Net has been widely used in many fields of image processing. This paper performs an improved semantic segmentation network Deep U-Net for glacier identification using Landsat 8 OLI image as the data source, and compares it with the traditional NDSI glacier identification method. The identification results are validated by the glacier label data produced by visual interpretation. The results indicate that the proposed method achieves an identification accuracy of 97.27%, which is higher than the NDSI glacier identification method. It can effectively exclude the interference of water bodies on glacier identification, and has a higher degree of automation.

1. INTRODUCTION

Glaciers are important freshwater resources on the land surface and are extremely sensitive to climate changes. Glacier changes have important impacts on climate changes, ecology and the environment, and water resources. In recent years, the issue of global warming has caused great concern worldwide, and brought about significant glacial ablation. According to research, the contribution of global glacial ablation to sea level rise is 29 ± 13% (Gardner et al., 2013; Ye et al., 2016). The continuous rise of sea level could lead to the submersion of some island nations and a large number of people will be homeless. As an important source of drinking water, glacial ablation also implies a severe water crisis (Gao et al., 2019; Pritchard, 2019). Therefore, the identification and monitoring of glaciers is of great importance in the study of global climate change.

The glacier areas usually have complex topography and are difficult to reach. Thus, field monitoring is costly and it is difficult to perform large-scale and long-term monitoring. The development of remote sensing technology has provided accurate and timely data sources for glacier research, making timely and large-scale glacier monitoring possible. Glacier identification based on remote sensing images is mainly divided into visual interpretation and computer-aided interpretation. The accuracy of visual interpretation is high, but the professional knowledge of the interpreter is required, along with heavy workload and time consumption. Computer-aided interpretation can quickly obtain wide-range glacier information, which has become a research focus in recent years. Common computer-aided glacier identification methods include band

ratio method and Normalized Difference Snow Index method (Salomonson, Appel, 2004). Based on the principle that glaciers have high reflection in the visible light band and low reflection in the near-infrared band, suitable bands are selected for band calculation to extract glacier area. These methods are easy to operate, but usually need a threshold, which is greatly influenced the identification result. Due to the differences of glaciers' types, images spatial resolution, images quality, and study areas, the threshold is difficult to adapt to different scenes. Moreover, these methods are based on the analysis of one or several features of the images, such as spectral features, texture features, or geometric features. How to use these features comprehensively to analyze the images to improve the accuracy of glacier identification is a pending issue. Therefore, the end-to-end method needs to be used to solve the above problems, so that glacier features can be directly learned from the images. The deep learning that develop rapidly these years has provided a way for end-to-end glacier identification.

Pixel-by-pixel glacier identification can be considered as a semantic segmentation problem for two classes. Early representative research on semantic segmentation using deep learning, proposed the Fully Convolutional Networks (FCN) (Long et al., 2015). FCN is mainly implemented through three technologies, convolutions, up-sampling and skip-connection. In the FCN, the fully connected networks are replaced by the convolutional networks, so that the architecture can accept images of any size and output segmentation result of the same size as the original image to achieve pixel-by-pixel classification. Up-sampling is necessary to convert the feature map to original size, which is implemented as transposed

---

* Corresponding author

convolution. The skip-connect layer take advantage of feature maps of different pooling layers for up-sampling to optimize the output. Many classical semantic segmentation methods based on deep learning are proposed based on FCN (Garcia-Garcia et al., 2017; Noh, 2015). DeepLab series (Chen et al., 2014; Chen et al., 2016; Chen et al., 2017; Chen et al., 2018) introduce the atrous convolution, which allows us to enlarge the field of view of filters to incorporate larger context. However, the atrous convolution makes the networks much more complex. With a large amount of weights to be computed, the DeepLab series networks are difficult to train, they need lots of training samples and will take a long time to fit. U-Net is a simple and efficient semantic segmentation network, which is first used in medical image processing. U-Net has the similar layers (convolutions, up-sampling) to FCN, while it adopts a completely different feature fusion method from FCN. U-Net concatenate feature maps in channel dimension to form a "thicker" feature map, so that it can make use of semantic information from multi-scale. Furthermore, due to the simple structure, U-net is very adaptable to small sample set. It is appropriate to glacier identification that hard to obtain label data.

In this study, Landsat8 OLI images will be used to propose a glacier identification method based on an improved semantic segmentation network Deep U-Net to perform glacier identification automatically, and evaluate the result. Then, compare the result with traditional remote sensing NDSI glacier identification result. The technical flowchart of this paper is shown in Figure 1.
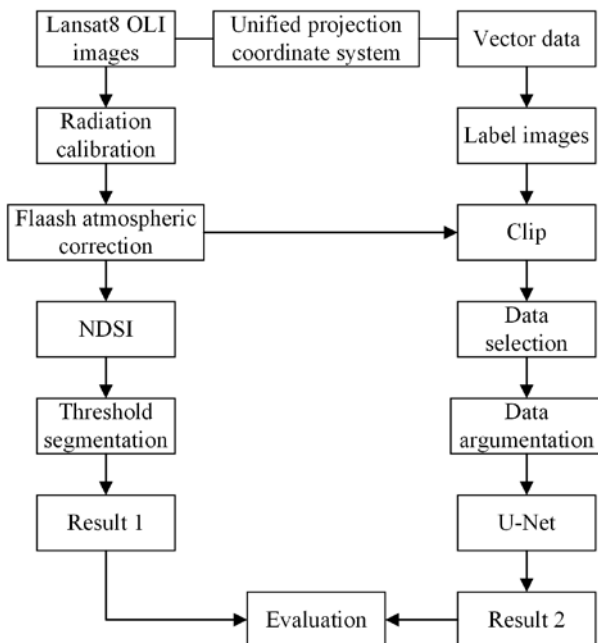

Figure 1. Flowchart of this paper

2. STUDY AREA AND DATA

The study area (Figure 2) is located in the central part of Eurasia. It belongs to a warm temperate continental arid climate zone with long light duration. Influenced by topography, landform and atmospheric circulation, the region has a complex and diverse climate, with large annual and daily changes in temperature, and less precipitation (Kan et al., 2016). Affected by the topography and regional precipitation, the source of each river system is mostly located in glaciers and mountain snow
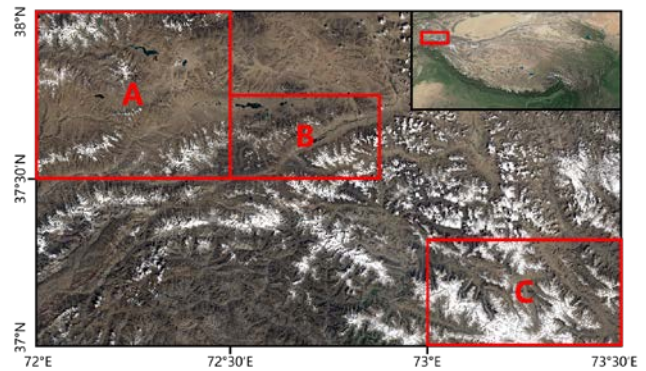

Figure 2. Study area (A, C are for training, B is for testing)

ID	Path/Row	Date
1	151/34	2018/09/19
2	151/35	2018/09/19
3	150/34	2018/09/12
4	150/35	2018/09/12
5	149/34	2018/08/20
6	149/35	2018/08/04
7	148/35	2018/08/29

Table 1. The related parameters of Landsat8 OLI images

Band	Bandwidth / μ m
1-COASTAL/AEROSOL	0.43-0.45
2-Blue	0.45-0.51
3-Green	0.53-0.59
4-Red	0.64-0.67
5-NIR	0.85-0.88
6-SWIR1	1.57-1.65
7-SWIR2	2.11-2.29

Table 2. Spectral information table

zones. The rivers usually have great seasonal contrast, floods in summer, drought in autumn and winter, and have extreme water shortage in spring.

The Landsat8 OLI images, and the vector glacier label data of the study area are used in this study.

The landsat8 OLI images are from the U.S. Geological Survey website (https://earthexplorer.usgs.gov/). The OLI (Operational Land Imager) includes 9 bands. Compared with the previous ETM + sensor, the bandwidth has been refined for six of the heritage bands. Two new bands are added. For example, the bandwidth of OLI Band 5 is adjusted to 0.845-0.885 μm to exclude the influence of water vapor absorption at 0.825 μm. The Band 1 (blue band, 0.433–0.453 μm) is mainly used for coastal zone observation, and the Band 9 (short-wave infrared band, 1.360–1.390 μm) is used for cirrus cloud detection (Roy et al., 2014). In this study, 7 OLI images are selected. The details are shown in the Table 1. After the radiation calibration, flaash atmospheric correction, and image mosaic through ENVI Classic 5.1, the typical areas A, B, and C (Figure 2) with few clouds are selected for research. After preprocessing, the
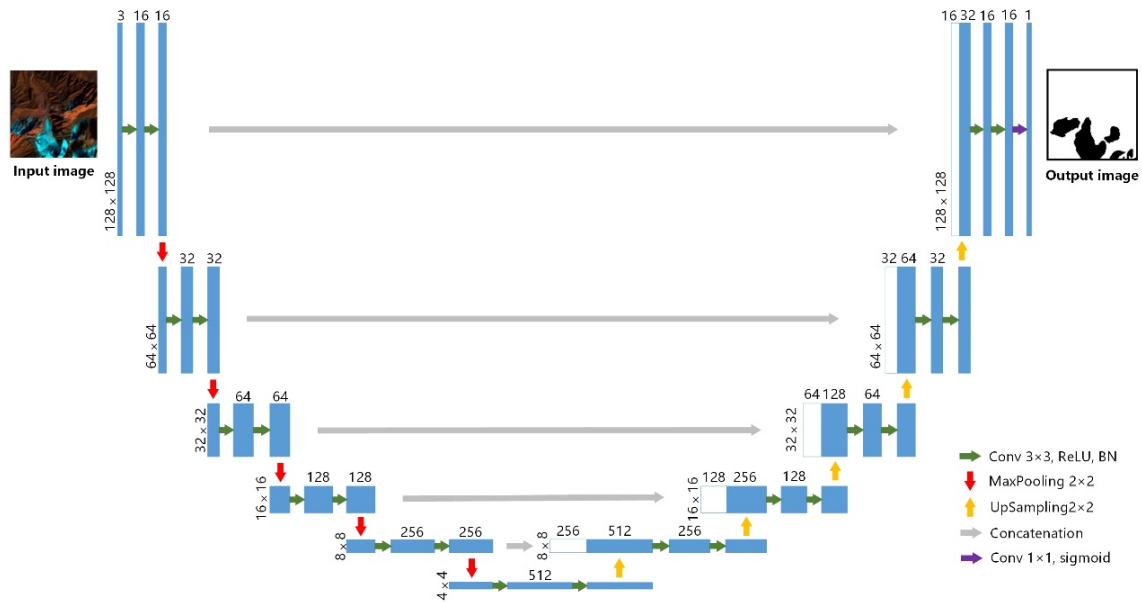
Figure 3. Architecture of the proposed Deep U-Net

Landsat8 OLI images contains 7 bands. The information of bands is shown in Table 2.

The vector glacier label data of the study area are produced by visual interpretation on the OLI images by professionals. The label data are preprocessed in arcgis 10.2. They are unified to the same coordinate system as the OLI image, and converted into raster data. The pixel size should be same as the corresponding OLI image pixel size. The glacier raster data will be used as ground true for follow-up research.

3. METHODOLOGY

3.1 Normalized Difference Snow Index

Based on the characteristics that glacier has different reflection in visible light band and short-wave infrared band, the Normalized Difference Snow Index (NDSI) (Hall et al., 1995; Salomonson and Appel, 2004) uses these two bands for normalization, and highlights the snow-covered parts of the image. The NDSI is defined as follows:

$$NDSI = \frac{Green - SWIR1}{Green + SWIR1} \qquad (1)$$

Where *Green* denotes OLI green band 3, *SWIR1* denotes OLI short-wave infrared band 6. The result of NDSI is between -1 and 1. A threshold segmentation is performed to change the result into a binary map of glacier.

3.2 Glacier Identification From Landsat8 OLI Imagery Using Deep U-Net

3.2.1 Deep U-Net: In this study, an improved fully convolutional network, U-Net (Ronneberger, 2015; Feng, 2019) was used to perform glacier identification. As a simple and efficient semantic segmentation network, U-Net has received widespread attention in medical image processing. The network has three characteristics, fully convolution, up-sampling and skip-connection. As shown in Figure3, the architecture consists of a contracting path (left side in Figure 3) and a symmetric

expanding path (right side). The contracting path repeated apply two 3*3 convolutions (stride 1, padding 1), each followed by rectified linear unit (ReLU), batch normalization (BN), and a 2×2 max pooling layer (stride 2) for downsampling. At each group of convolutions, we double the number of feature channels. The expansive path consists of 5 groups of operations for upsampling, which are a 2×2 stride 2 convolution ("upconvolution"), a concatenation with the correspondingly cropped feature map from the contracting path, and two 3x3 convolutions which halve the number of feature channels, each followed by ReLU and BN. At the last layer of the network, a 1×1 convolution followed by sigmoid are used to map the output feature map into the required number of channels.

The number of weight parameters in the convolution layer has nothing to do with the size of the feature map. Therefore, a fully convolutional network can accept input images of any size for training. At the same time, the GPU can speed up the convolution operation, which in turn reduce computing time and improve computing efficiency. U-net combines the characteristics of Encoder-Decoder structure and skip-connect network, and is more elegant. The down-sampling branch obtains the high-level semantic features of the image (Zeiler, 2013), and the up-sampling branch and the skip-connection concatenate feature maps in channel dimension to form a "thicker" feature map, to make full use of spatial information and multi-scale semantic information to output a more accurate segmentation result. The U-Net network is relatively simple with fewer parameters. It can adapt to small datasets (Ronneberger, 2015; Feng, 2019). It is appropriate to glacier identification that hard to obtain label data.

3.2.2 Data Processing: We observe the spectral characteristics of the glacier, and finally select the band combination 6 (R), 5 (G), 2 (B) for experiments. On this fake color synthesis image (Figure 4), the glacier information is prominent and well distinguished from the clouds. The Landsat8 images and the label images of the A and C areas are divided into patches with 128 × 128 pixels. The study area has fewer glacier pixels than non-glacier pixels. For deep learning algorithms, in order to ensure the stability of the training model,

the number of positive samples and negative samples should be as close as possible. Based on this, images with fewer glacier pixels are removed from the dataset, and the dataset is randomly divided into training and validation sets in proportion. Due to the small size of the dataset, a data argumentation (flip, rotate at different angles) is performed on the training samples. Finally, there are 4866 pairs of training samples and 80 pairs of validation samples (Figure 5).

3.2.3 Implementation: The full implementation is based on the Keras (Chollet et al., 2015) library with TensorFlow as its backend. All experiments were conducted on a computer with Inter(R) Xeon(R) E5-2687W CPU at 3.00 GHz, 16.0-GB RAM and NVIDIA GRID RTX8000-8Q (8 GB). With the "Binary-Entropy" in Keras as loss function. We used the "Adadelta" optimization algorithm (Zeiler, 2012) and the "sigmoid" activation function. The batch size of the training and validation datasets is set to 8 images. Within 50 epochs, the model with the highest validation accuracy is saved, and the model is used to identify the glacier from the test image. The training process takes about 8 hours.

3.3 Evaluation Metrics

The evaluation criteria are calculated based on a confusion matrix shown as Table 3.

(a)   (b)

Figure 4. (a) 432-band true color synthesis.
(b) 652-band fake color synthesis

Images	Landsat8 images	Label images
1		
2		
3		

Figure 5. Examples of the training samples

Ground true	prediction	
	glacier	Non glacier
Glacier	TP	FN
Non glacier	FP	TN

Table 3. Confusion matrix

Where TP is true positive that denotes glacier pixels in both result and label image. TN is true negative that non-glacier pixels are identified as glacier. FP is false positive when non-glacier pixels are erroneously identified as glacier. While FN is false negative which is calculated for pixels are both non-glacier. The evaluation criteria are defined as below:

$$\text{The false alarm rate: } FA = \frac{FP}{TP + FP} \quad (2)$$

$$\text{The missed alarm rate: } MA = \frac{FN}{TP + FN} \quad (3)$$

$$\text{The overall accuracy: } OA = \frac{TP + TN}{TP + FN + TN + FP} \quad (4)$$

$$\text{The Kappa coefficient: } KC = \frac{OA - PRE}{1 - PRE} \quad (5)$$

$$PRE = \frac{(TP + FP) \times (TP + FN) + (FN + TN) \times (FP + TN)}{(TP + FN + TN + FP)^2} \quad (6)$$
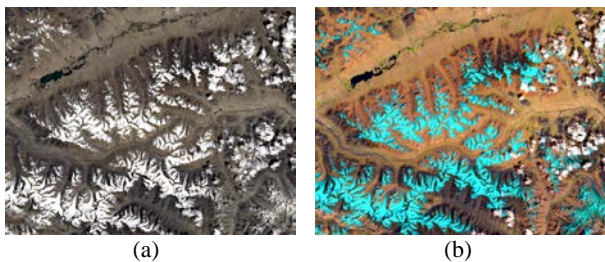

Figure 6. Histogram of the NDSI grayscale image

Methods		MA (%)	FA (%)	OA (%)	KC (%)
NDSI Threshold	0	13.97	22.21	95.73	79.29
	0.1	**16.06**	**20.36**	**95.84**	**79.39**
	0.2	17.79	19.40	95.84	79.05
	0.3	19.56	18.57	95.80	78.57
	0.4	21.43	17.92	95.72	77.89
	0.5	23.43	17.42	95.61	77.01
	0.6	25.75	16.99	95.46	75.86
	0.7	28.66	16.68	95.24	74.23
Deep U-Net		**19.49**	**6.02**	**97.27**	**85.22**

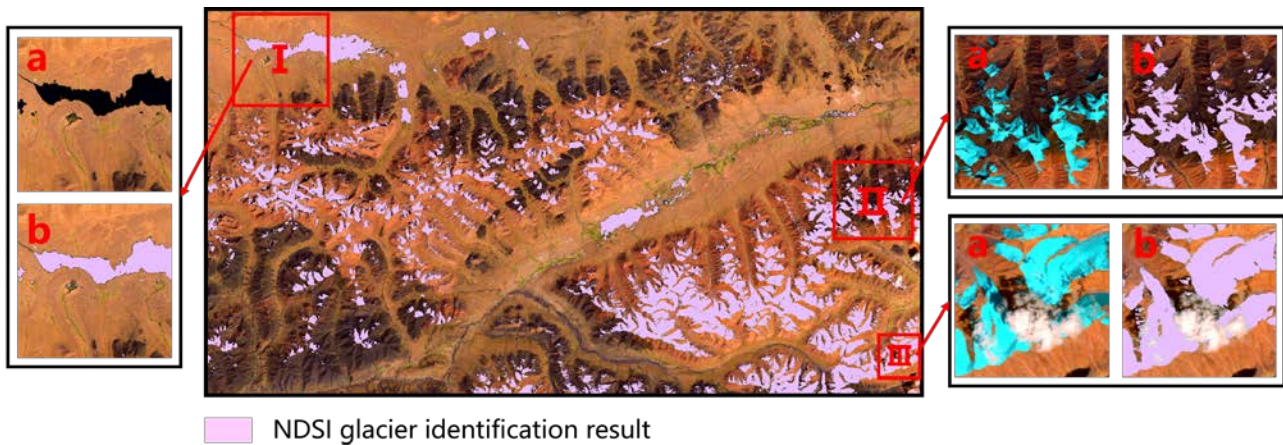Table 4. Accuracy of glacier identification

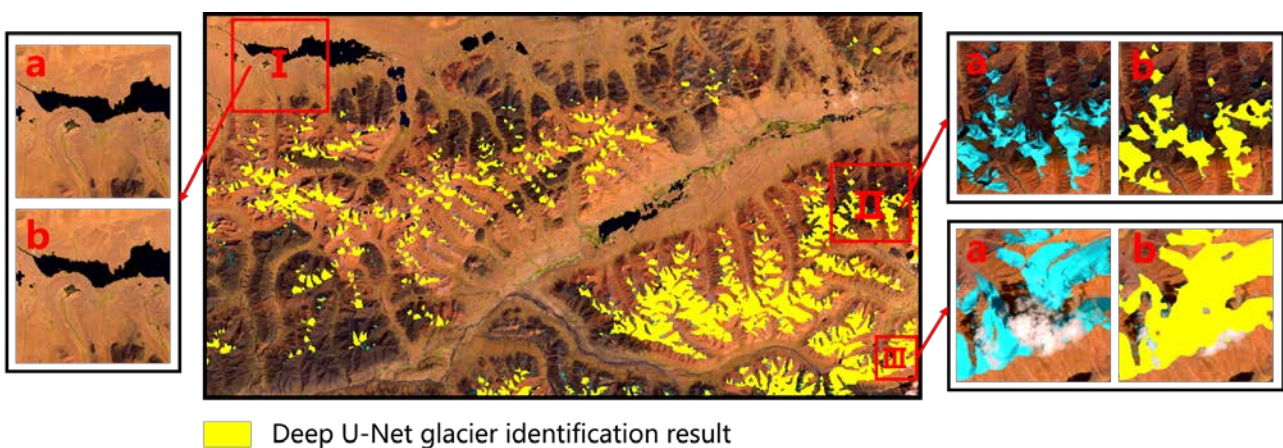Figure 7. NDSI glacier identification result. (a): original image. (b): identification result

Figure 8. Deep U-Net glacier identification result. (a): original image. (b): identification result

4. RESULTS AND ANALYSIS

After a histogram analysis on the NDSI grayscale image (Figure 6), and we can see the obvious "double peak". Thresholds are taken at intervals of 0.1 between the two peaks (data value $\in [0, 0.7]$) to binarize the NDSI images, and then compute the accuracy of segmentation results using these thresholds. The results are shown in Table 4.

It can be seen that the missed alarm rate is proportional to the thresholds, however the other three are negatively correlated to the thresholds. In summary, the glacier identification reaches the best result when the threshold is 0.1. Take the segmentation map of threshold 0.1 as the NDSI glacier identification result, shown in Figure 7.

Glacier identification result using Deep U-Net is shown in the Figure 8. The result is "clean", with few broken objects, and highly matches with the ground true image. The NDSI glacier identification result (Figure 7) has many misclassifications and omissions. Comparing area Ⅰ, NDSI method misclassifies some water bodies as glaciers, while the Deep U-Net method performs very well. As for the results of area Ⅱ, it can be seen that both methods have a small amount of missing glacier pixels and are difficult to identify the shadow area. There are some clouds in area Ⅲ, the Deep U-Net method fails to distinguish them.

The accuracy results (Table 4) show that the glacier identification from Landsat8 OLI imagery using Deep U-Net can get a low false alarm rate of 6.02%, but there are some cases of missing detection and some details are lost. This method can reach a high overall accuracy of 97.27%, which is higher than the traditional NDSI glacier identification results. The identification results are basically reliable.

5. DISCUSSION

Generally speaking, the automatic glacier identification method based on Deep U-Net network has higher accuracy than the NDSI glacier identification method, but the situation of missing detection needs to be improved, especially for some small targets, the identification result is worse than the NDSI method. Further, the Deep U-Net model can be trained with the NDSI image as a channel of the image to verify whether the combination of the NDSI method and the Deep U-Net method can improve the accuracy of glacier identification, and reduce the missed alarm rate of small targets.

The automatic glacier identification method based on Deep U-Net network can reach a high accuracy in a short time. It can be used for large-scale glacier identification, which is helpful to the study of the global glaciers distribution. With this method, it is also possible to monitor long-term glacier changes, study the rule of regional glacier changes, and provide a basis for climate and environmental change research.

6. CONCLUSION

For the task of glacier identification, we observe the spectral characteristics of the glacier based on the Landsat8 OLI data, derive the dominant bands for glacier identification. An automatic glacier identification method based on Deep U-Net network is proposed, and compared to existing glacier identification methods.

On the combined image of Landsat8 OLI bands 6 (R), 5 (G), 2 (B), the glacier information is prominent, which is conducive to the glacier identification. The Normalized Difference Snow Index method can generally identify glacial areas on Landsat8 OLI images, but it will misclassify some water bodies into glaciers, and it cannot effectively distinguish shadows and clouds. Glacier identification method based on Deep U-Net network can well exclude water bodies and shadow areas, but the misclassifications of shadows and clouds are still existed. It has been proved that this method is efficient and more automatic, and has higher accuracy than NDSI method in glacier identification in our study area.

ACKNOWLEDGEMENTS

REFERENCES

Chen, L.C., Papandreou G., Kokkinos I., Murphy, K., Yuille, A.L., 2014. Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs. *Computer Science*, 4, 357-361.

Chen, L.C., Papandreou G., Kokkinos I., Murphy, K., Yuille, A.L., 2016. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 99, 834-848.

Chen, L.C., Papandreou, G., Schroff, F., Adam, H., 2017. Rethinking Atrous Convolution for Semantic Image Segmentation. arXiv:1706.05587v3 [cs.CV]

Chen, L.C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H., 2018. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. arXiv:1802.02611v3 [cs.CV].

Chollet, F., et al., 2015. Keras. GitHub repository. https://github.com/fchollet/keras.

Feng, W., Sui, H., Huang, W., Xu, C., An, K., 2019. Water Body Extraction From Very High-Resolution Remote Sensing Imagery Using Deep U-Net and a Superpixel-Based Conditional Random Field Model. *IEEE GEOSCIENCE AND REMOTE SENSING LETTERS,* 16(4), 618-622.

Gao, J., Yao, T., Masson-Delmotte, V., Steen-Larsen, H.C., Wang W., 2019. Collapsing glaciers threaten Asia's water supplies. *Nature*, 565, 19-21.

Garcia-Garcia, A., Orts-Escolano, S., Oprea, S.O., Villena-Martinez, V., Garcia-Rodriguez, J., 2017. A Review on Deep Learning Techniques Applied to Semantic Segmentation. arXiv:1704.06857v1 [cs.CV].

Gardner, A.S., Moholdt, G., Cogley, J.G, et al., 2013. A reconciled estimate of glacier contributions to sea level rise: 2003 to 2009. *Science*, 340(6134), 852-857.

Hall, D., Riggs, G., Salomonson, V., 1995. Development of Methods for Mapping Global Snow Cover Using Moderate Resolution Imaging Spectroradiometer Data. *Remote Sensing of Environment,* 54(2), 127-140.

Kan, X., Zhang, Y., Cao, T., Wang, J., Tian, W., 2016. Snow Cover Recognition for Qinhai-Tibetan Plateau Using Deep Learning and Multispectral Remote Sensing. *Acta Geodaetica et Cartographica Sinica,* 45(10), 1210-1221.

Li, Z., Zhou, J., Tian, B., Xie, C., 2008. The Glacier Identification using SAR Interfermetric and Polarimetric Information in Qinghai-Tibetan Plateau. IEEE International Geoscience & Remote Sensing Symposium, IGARSS 2008, IV, 1054-1056.

Long, J., Shelhamer, E., Darrell, T., 2015. Fully Convolutional Networks for Semantic Segmentation. CVPR 2015, Computer Vision Foundation, 3431-3440.

Noh, H., Hong, S., Han, B., 2015. Learning Deconvolution Network for Semantic Segmentation. Computer Vision (ICCV), 2015 IEEE International Conference.

Pritchard, H.D., 2019. Asia's shrinking glaciers protect large populations from drought stress. *Nature*, 569, 649–654.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer International Publishing, 2015.

Roy, D.P., Wulder, M.A., Loveland, T.R., et al., 2014. Landsat-8: Science and product vision for terrestrial global change research. *Remote Sensing of Environment*, 145, 154-172.

Salomonson, V.V., Appel, I., 2004. Estimating Fractional Snow Cover from MODIS Using the Normalized Difference Snow Index. *Remote Sensing of Environment*, 89, 351-360.

Ye, Q, Cheng, W, Zhao, Y, Song, J., Zhao, R., 2016. A review on the research of glacier changes on the Tibetan Plateau by remote sensing technologies. *Journal of Geo-information Science,* 18(7), 920-930.

Zeiler, M.D., 2012. ADADELTA: AN ADAPTIVE LEARNING RATE METHOD. *Computer Science*.

Zeiler, M.D., Fergus, R., 2013. Visualizing and Understanding Convolutional Networks. Springer International Publishing Switzerland, ECCV 2014, Part I, LNCS 8689, 818–833.