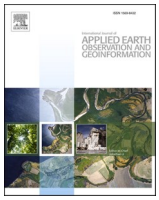
International Journal of Applied Earth Observations and Geoinformation 112 (2022) 102928

Contents lists available at ScienceDirect

International Journal of Applied Earth Observations and Geoinformation

journal homepage: www.elsevier.com/locate/jag



Deep learning empowers the Google Earth Engine for automated water extraction in the Lake Baikal Basin

Kai Li [a,b], Juanle Wang [a,c,*], Wenjing Cheng [d,e], Yi Wang [f], Yezhi Zhou [a,b], Ochir Altansukh [g]

[a] State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China
[b] College of Geoscience and Surveying Engineering, China University of Mining & Technology (Beijing), Beijing 100083, China
[c] Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing 210023, China
[d] Chinese Academy of Meteorological Sciences, Beijing 100081, China
[e] Chinese Academy of Meteorological Sciences Training Center, Beijing 100081, China
[f] National Science and Technology Infrastructure Center, Beijing 100862, China
[g] Department of Environment and Forest Engineering, National University of Mongolia, Ulaanbaatar 210646, Mongolia

ARTICLE INFO

Keywords:
Water extraction
Deep learning
Noise correction
Quality assessment bands
Google Earth Engine
Lake Baikal

ABSTRACT

Studying the spatial and temporal water distribution in the Lake Baikal Basin, which hosts the freshwater lake with the most water storage in the world, is essential to understand the water resources and environment of the basin and its impact and influence in terms of climate change and disaster prevention and mitigation. The basin spans two countries, Russia and Mongolia, which, along with its vastness, makes it challenging to accurately automate the acquisition of large-scale and long-term series data. The Google Earth Engine (GEE) is capable of processing large amounts of remote sensing imagery but does not support the computation and application of deep learning models. This study uses a combination of local deep learning training and GEE cloud-based big data intelligent computing to empower GEE with deep learning computing power, enabling it to rapidly automate the deployment of deep learning models. Visible light, near infrared (NIR), modified normalized difference water index (MNDWI), short-wave infrared 1 (SWIR1), linear enhancement band (LEB), and digital elevation model (DEM), which are more sensitive to water bodies, were selected as input features, along with the optimized input features of the existing pixel-based convolutional neural network (CNN) model. This method corrects the initial water labels from the Landsat quality assessment bands to reduce the time cost of manually drawing the labels and improving the classification accuracy of the water bodies. On average, only 1–2 h are required to generate the results for each water body product for each period in Lake Baikal Basin. The extraction of water bodies from the Lake Baikal Basin was achieved for nine yearly periods between 2013 and 2021. The validation accuracy was 92.9 %, 92.7 %, and 92.4 % for the three years 2013, 2017 and 2021, respectively. The results showed that the mean area of water bodies in the basin was 37,500 km^2 and that the area of water bodies in the basin fluctuated without significant change from 2013 to 2021. This study provides methodological support for the continuous monitoring and assessment of water body dynamics at more catchment scales and other large scenarios.

1. Introduction

Lake Baikal in Northeast Asia is the deepest and most abundant freshwater lake globally, with enough water to sustain five billion people for half of a century. It was selected as a UNESCO World Natural Heritage Site in 1996 (Zheng, 2002). As early as the last century, scientists have been studying Lake Baikal in the fields of climate change (Bolgrien et al., 1995, Mamaev, 1987), geology (Zonenshain and Savostin, 1981, Soloviev et al., 1989), and water environment (Bezrukov et al., 1990, Oshchepkov and Shlyakhova, 1987), etc. The United Nations Sustainable Development Goals 6 (SDG 6) refers to the provision and sustainable management of water and sanitation for all, with sub-

* Corresponding author at: State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China.
E-mail address: wangjl@igsnrr.ac.cn (J. Wang).

https://doi.org/10.1016/j.jag.2022.102928
Received 19 February 2022; Received in revised form 20 June 2022; Accepted 15 July 2022
Available online 28 July 2022

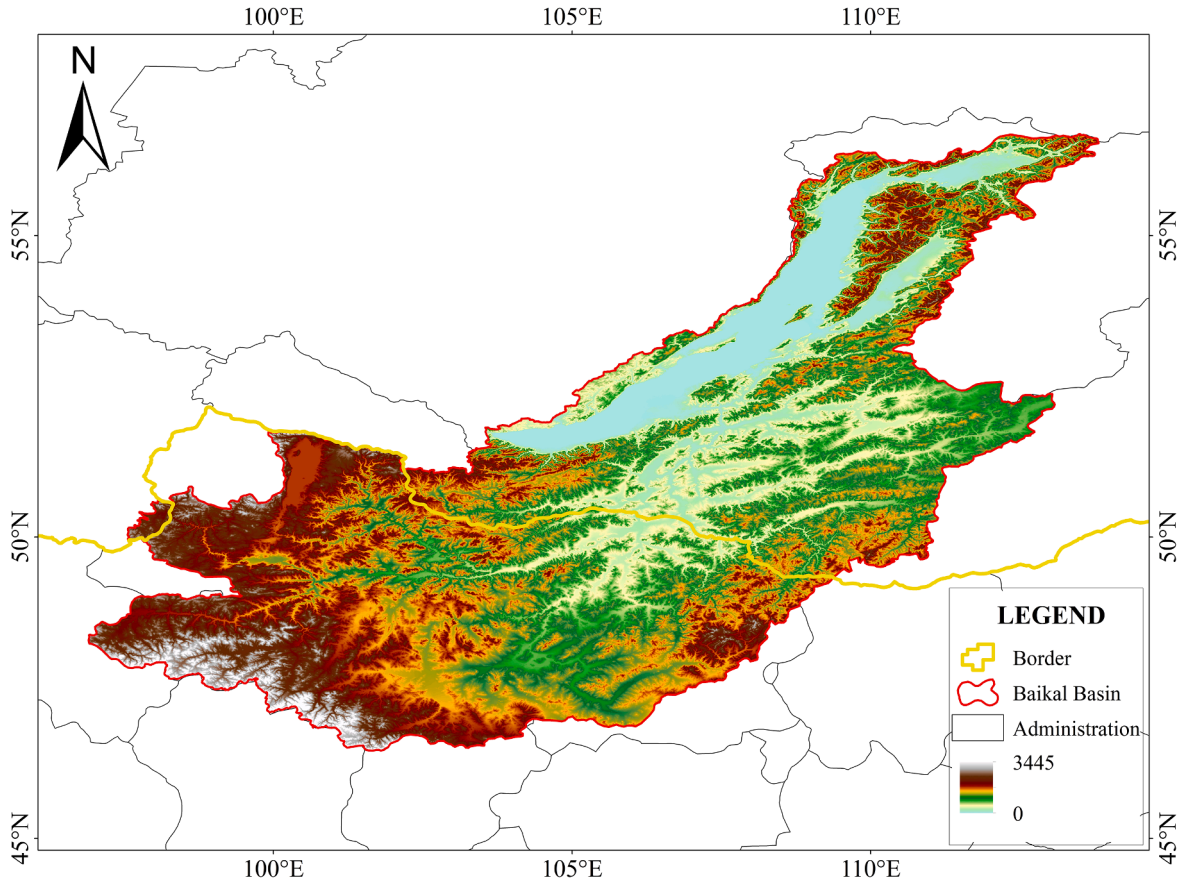


Fig. 1. Study area.

goal 6.6 discussing the protection and restoration of water-related ecosystems by 2020, including mountains, forests, wetlands, rivers, underground aquifers, and lakes. Understanding the spatiotemporal variability of surface water bodies in the Lake Baikal Basin is of great scientific importance for sustainable water resources and environmental development in Northeast Asia and the surrounding regions. The development of earth-orbiting satellite technology and observations has made the annual monitoring and assessment of water bodies on a large scale possible. However, the vast area of Lake Baikal and the fact that it straddles two countries, Russia and Mongolia, makes obtaining information on its dynamic water bodies difficult, as they are influenced by different spatiotemporal conditions, scenarios, and feature types. This also makes the automated extraction of large-scale, long-time series water bodies at the basin scale very challenging.

In 1996, McFeeters proposed the normalized difference water index (NDWI), which uses the green and near infrared (NIR) bands of remote sensing imagery to normalize the water index, to suppress soil and vegetation features for water body characterization and surface water areas estimation (McFEETERS and S., 1996). Xu constructed a new index suitable for extracting water information in built-up areas, the modified normalized difference water index (MNDWI), by replacing the NIR band in the NDWI index with short-wave infrared (SWIR) (Xu, 2006). Feyisa et al. constructed an automated water extraction index (AWEI) to improve the classification accuracy of commonly misclassified shaded and dark surface features, with better overall accuracy than MNDWI and maximum likelihood classification methods (Glf et al., 2014). However, these indexes for classifying water suffer from the difficulty of determining thresholds for different scenarios; specifically, single thresholds or single indices often make it difficult to achieve more accurate water classification results in fine rivers and shaded mountain areas.

Although many global landcover products have been developed (Gong et al., 2013, Gong et al., 2019, Zhang et al., 2020), an in-depth discussion on water misclassification is lacking. Studies dedicated to the extraction of water bodies have mostly been focused on the local scale in small towns, mountains, and the like. For example, Yang et al. extracted water bodies in urban areas of Beijing and Yantai city from Sentinel imagery using image sharpening NDWI (Xiucheng et al., 2017), and Wang Fan et al. carried out an improved water extraction technique in the predominantly mountainous Yibin region (Wang, 2021). These studies mainly address the extraction of water bodies in specific environments with clouds, cloud shadows, and building shadow coverage. Even as we enter the era of remote sensing artificial intelligence (AI) and a plethora of algorithms continue to emerge (Long et al., 2015, Badrinarayanan et al., 2017, Ronneberger et al., 2015), water segmentation methods remain dominated by small-scale, regional studies due to over-reliance on labeled samples (Li et al., 2019, Weng et al., 2020, Li et al., 2021a). Therefore, the main problems faced by multi-feature, multi-temporal and multi-scene water extraction are the difficulty of sample annotation and the difficulty of deploying a large-range and long-term deep learning model.

Google Earth Engine (GEE) is a remote sensing cloud computing platform that provides significant image data while allowing for image processing, statistics, and analysis. Its advantage is that it allows fast and efficient research and mapping of large areas. Deep learning from remote sensing imagery has become a popular research topic. However, the GEE platform struggles to support deep learning computation in the cloud, making it difficult to deploy substantial cloud-based work.

To address these issues, this study collaboratively invokes the local deep learning environment and the GEE remote sensing data and computing platform, uses local deep learning training, and gives the trained models to the GEE platform for deep learning computing power, so that GEE can rapidly automate the deployment of deep learning

Table 1
Raw data information.

Source	Band	Min	Max	Resolution	Wavelength
Landsat 8	B1	1	65,455	30 m	0.435–0.451 μm
	B2	1	65,455	30 m	0.452–0.512 μm
	B3	1	65,455	30 m	0.533–0.590 μm
	B4	1	65,455	30 m	0.636–0.673 μm
	B5	1	65,455	30 m	0.851–0.879 μm
	B6	1	65,455	30 m	1.566–1.651 μm
	B7	1	65,455	30 m	2.107–2.294 μm
	QA	0	65,535	30 m	/
NASADEM	elevation	−512	8,768	30 m	/

models. This study produced a set of machine-learning samples for water classification in the Baikal Basin, a set of surface water products for 2013–2021, a set of pixel-based CNN-based water extraction models, and a set of deep learning model weight resolution empowerment tools.

2. Study area and materials

2.1. Study area overview

The northern region of Lake Baikal Basin is in Russia, and the southern region is in Mongolia, with each country accounting for approximately half of the 570,000 km^2 total basin area (see Fig. 1). Lake Baikal is the oldest, deepest (1,637 m), and largest freshwater lake in the world, storing approximately 20 % of the global unfrozen freshwater while hosting more than 1,500 endemic species. The study area has a low inland elevation and a high external elevation, especially in the southwest. As a low-altitude area throughout the region, Lake Baikal is fed by the Selenge River in Mongolia and a total of 336 other rivers of all sizes and flows, including the Angara River, a tributary of the Yenisei River.

2.2. Materials

The primary data sources for this study are listed in Table 1. The data were selected from Landsat 8 Operational Land Imager (OLI) imagery and NASA's digital elevation model (DEM) data, both at a resolution of 30 m. The B1–B7 bands and the DEM data were chosen to build water-sensitive features as input features for the water segmentation model. Min and Max in the Table means the minimum and maximum value of the band. The quality assessment (QA) band (USGS) was used for cloud removal and water label construction. Overall, we synthesized 65 images in the research area in one year and 585 images in nine years.

3. Methodology

Based on the GEE cloud computing platform and local deep learning computing environment, this study developed a set of water extraction techniques for remote sensing images that were suitable for combining big data acquisition and AI computing. The technical route of this process is shown in Fig. 2, which is divided into three major parts: high-quality remote sensing data acquisition, deep learning model construction and noise correction, and deep learning model online deployment. We aimed to combine local deep learning training with online cloud computing and rapidly automate the acquisition of training model weights and implement large-scale batch deployment in GEE. Session 1 uses the GEE platform to select data and provide the necessary Landsat images, initial labels, and DEM data for Session 2, which then performs feature selection on the images, noise correction on the initial labels, and combines the feature bands with the water labels to train the pixel-based CNN model. Session 3 is the parsing of the above water segmentation model and the deployment of the model via the gee-Python interface to enable model transcription and allow the model parameters to be deployed to the cloud for quick large-scale water extraction.

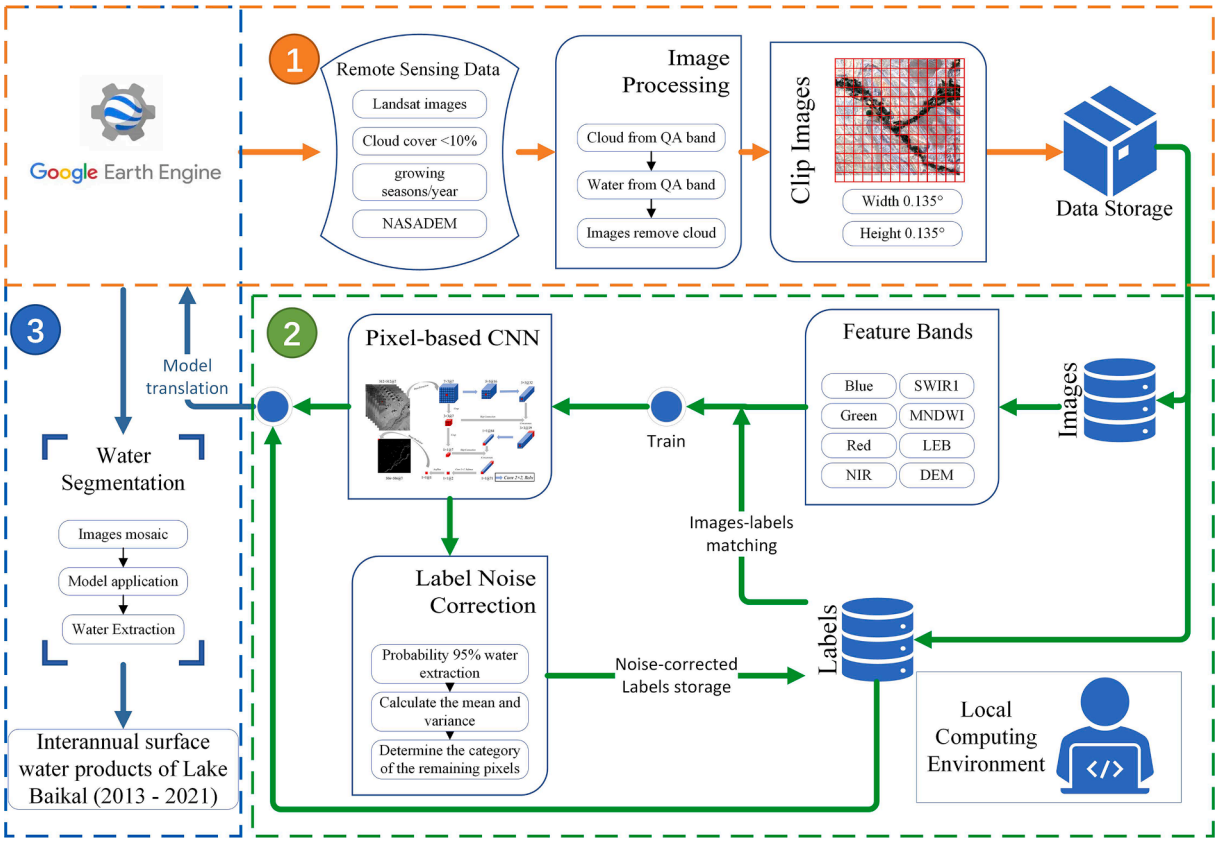


Fig. 2. Technical route.

Table 2
QA band information.

Bits	Meanings	Bits	Meanings
Bit 0	Fill	Bits 10–11	Cloud Shadow
Bit 1	Dilated Cloud	0	None
Bit 2	Cirrus (high confidence)	1	Low
Bit 3	Cloud	2	Medium
Bit 4	Cloud Shadow	3	High
Bit 5	Snow	Bits 12–13	Snow/Ice
Bit 6	Cloud or Dilated Cloud	0	None
0	set	1	Low
1	not set	2	Medium
Bit 7	Water	3	High
Bits 8–9	Cloud Confidence	Bits 14–15	Cirrus
0	None	0	None
1	Low	1	Low
2	Medium	2	Medium
3	High	3	High

3.1. Remote sensing data acquisition

In the high-quality image data acquisition session, the QA bands for Landsat 8 imagery were primarily evaluated on a per-bit basis. The QA bands are the individual pixel quality bands assessed by NASA through the FMASK (Function of MASK) algorithm (Qiu et al., 2019), first applied to Landsat 8 satellite imagery in 2014 and later expanded to Landsat 5 and 7 satellites. The meaning of the binary bits of the QA band of the Landsat 8 Collection 2 Level 2 image is shown in Table 2, from which the QA band covers information such as clouds, snow, cloud shadows, and water.

The pixels recorded in bits 1–4 of the Landsat 8 QA band are cloud-covered, often losing much of their ground information. Bitwise operations masked out the cloud and cloud shadow-covered areas of the same area image, and the masked image was superimposed to obtain a cloud-free image of the area, reducing the interference of clouds and cloud shadows in the water extraction process. As shown in Fig. 3, the image visualization effect and data availability improved after the cloud removal and stacking processes. The 7th bit in the QA band is also indicated in Table 2 as water, allowing the extraction of water information in that band by bit, which provides the initial label for water label construction.

3.2. Feature bands construction

The effectiveness of deep learning models is often determined by the input training datasets, and thus the choice of remote sensing image features governs the merit of image classification. Training water-sensitive indexes and bands improve the accuracy of water classification effectively. Therefore, the more water-sensitive bands, blue, green, red, NIR, MNDWI, short-wave infrared 1 (SWIR1), linear enhancement band (LEB), and DEM were chosen for the water segmentation process. The involvement of DEM can better remove the misclassification of water caused by mountain shadows. LEB is obtained by performing a linearly enhanced convolution operation on the MNDWI index. Fig. 4 shows the four convolutional kernels of the linear enhancement band, which linearly enhances the MNDWI data in the vertical, horizontal, and left–right 45-degree directions. After convolution, the maximum value of each pixel is calculated to obtain the LEB in the eight neighborhoods around that pixel. Fig. 5 shows a comparison between MNDWI and LEB, where the first column is the false color synthesis images, the second column is the MNDWI index images, and the third column are LEBs. The LEBs enhance the features of linear features and suppresses the expression of homogeneous features.

3.3. Deep learning model building and label noise correction

The pixel-based CNN model proposed by Li et al. (Li et al., 2021b) was used in this study. Visible light, NIR, MNDWI, SWIR1, LEB, and DEM were chosen as the input features of the model to enhance the separability of water and improve classification accuracy. The pixel-based CNN model is a water extraction model that considers the pixel texture features and spectral feature information. The model was trained

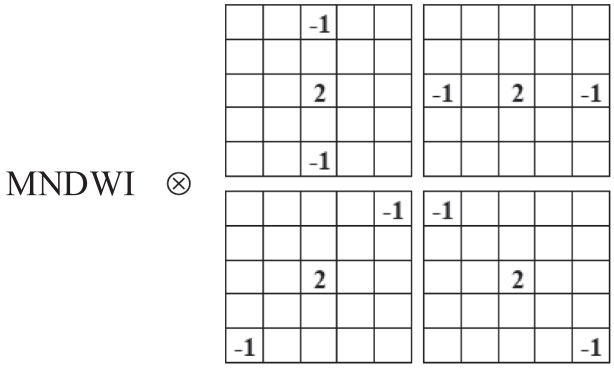


Fig. 4. Kernels of linear enhancement.

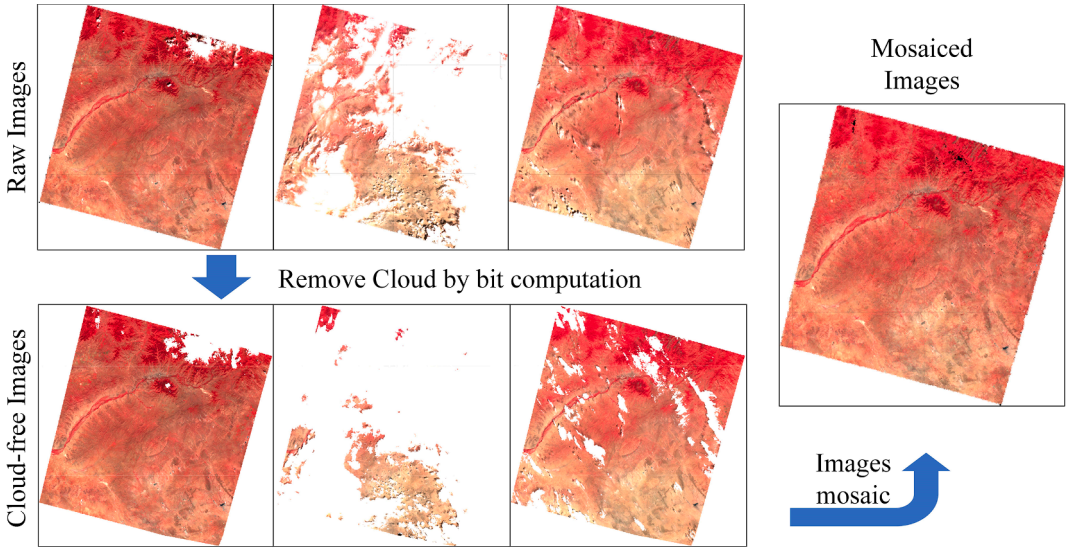


Fig. 3. Basic de-clouding process.

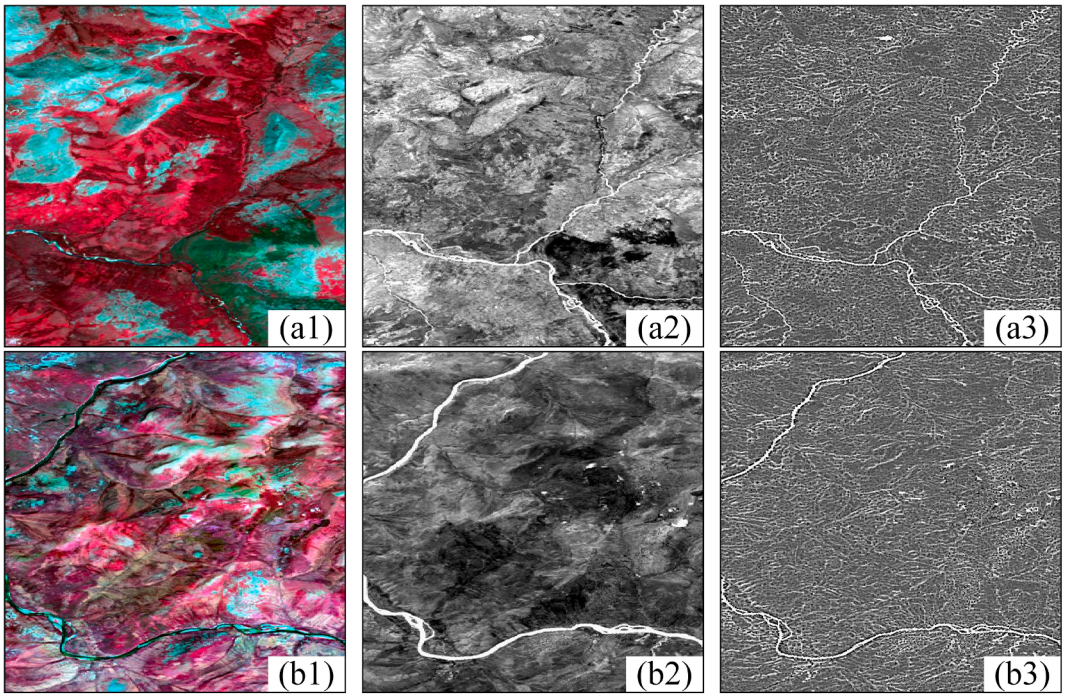


Fig. 5. Linear enhancement.

iteratively by traversing the entire image. The 7 × 7 neighborhood of the feature band pixel was first taken as the input. After two 3 × 3 convolutions, it was spliced with a 3 × 3 region in the middle of the input features. Another convolution and splicing were performed, and, finally, a 1 × 1 convolution and softmax function were used to determine the attribution probability for each category of the central pixel. The cross-entropy loss was calculated by combining the attribution probability with the labeled pixel and was reduced by iterative optimization of the

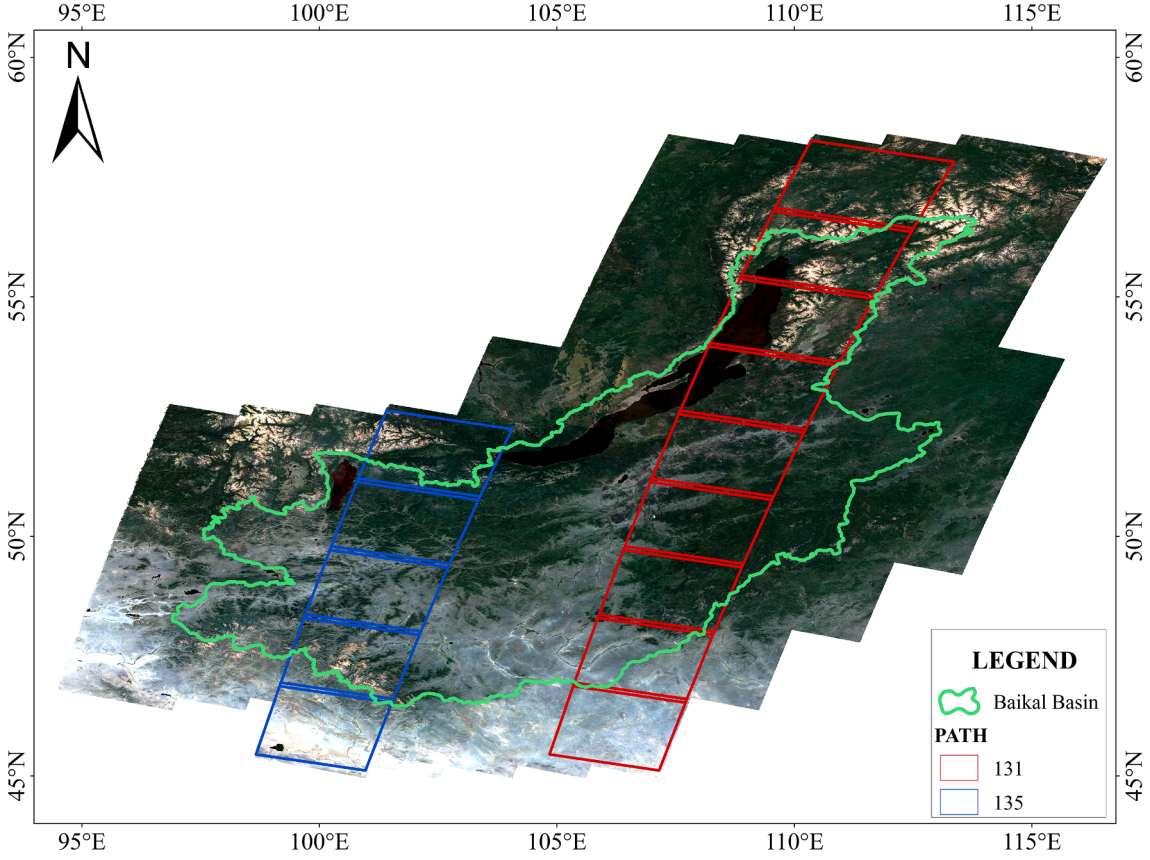


Fig. 6. Stacked image after removal of clouds.

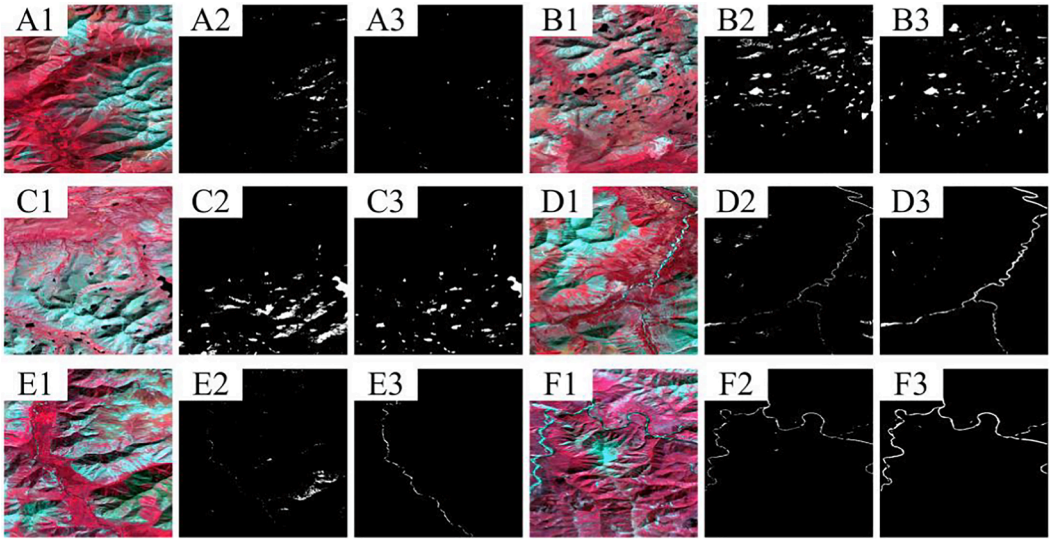


Fig. 7. Comparison of labels before and after noise correction.

parameters through gradient descent.

Once the model was constructed, it was applied for tag noise correction. Water information in the QA band is implemented by the FMASK algorithm using local dynamic thresholding, which suffers from poor water body continuity and is affected by clouds and cloud shadows. Therefore, noise deduction is required for the QA band. In this study, a normal distribution was fitted to the characteristic wavebands under the assumption that the reflectance distribution of the local area water body is normal. The pixel-based CNN model was used to train the initial labels 30 times, making the model more sensitive to the water of the initial labels. If the probability of a pixel being water is higher than 95 %, the pixel will be classified as water. The mean (μ) and variance (σ) of the characteristic bands were then calculated for these determined water labels. The remaining images in the region to be determined were considered to be water bodies if they fall within the interval (μ-3σ, μ + 3σ), and vice versa. This cycle was repeated three times.

The noise-corrected water labels and feature images were obtained using this method. After initializing the pixel-based CNN model and training the corrected water labels, a water segmentation model for the Lake Baikal Basin was obtained.

3.4. Online deployment of models

The water extraction model for the Lake Baikal Basin was obtained using the above method, but the application of the model to a large range of remote sensing images is also a major challenge. The traditional method requires manually downloading the images and then applying the model to each. When faced with large-scale and long-time series applications, the problems of difficult image acquisition, low computer performance, and slow image mosaic are evident. This study automates the parsing of the trained model, assigns the local model architecture and parameter weights to the GEE, and integrates the corresponding computational module of the GEE to achieve online deployment of the model.

4. Results

4.1. Remote sensing image data and initial labels acquisition

This study used a deep learning model to train remote sensing images of the Lake Baikal Basin and achieve the water extraction results; therefore, image data and labels were needed as the basis for model training. One year of full-coverage images of Lake Baikal required 65 scenes of OLI images, with paths from 128 to 138 and rows spanning from 20 to 28.

Fig. 6 shows the results of image stacking across the study area. The main month of restriction was the northern growing season (June to August), and the data were averaged by image element after filtering. As shown in Fig. 6, controlling the growing season effectively reduces interference at the seams. In particular, the seam and boundary information can have a significant impact on the classification of water bodies during convolution operations. The data are mean or median processed by image element location to mosaic the 65 images into one image. Relying on the cloud computing power of the GEE cloud computing platform makes it possible to input a single image of a large area of the Lake Baikal Basin. The data prediction process reduces the redundant operations of cyclic iterations and removes the duplication of calculations in areas where different images overlap, thereby increasing the efficiency of the product output.

The input training images are superimposed synthetic de-clouded images for the 2020 growing season of paths 131 and 135. Fig. 6 shows the geospatial locations of columns 131 and 135, which are in the main area of the Lake Baikal Basin and cover part of Lake Baikal and the small tributaries of the Selenge River. These images were north–south oriented and covered a wide multi-surface area, which helped the model learn information in different scenarios during the training process, including large lakes, small rivers, and other special surface properties. Ultimately, the model learned information from different scenarios during training, including large lakes, small rivers, and other surface-specific attributes.

For the images within the sample strips, we reprojected the data to the EPSG:4326 (WGS84) coordinate system to match the projected coordinates of the NASADEM data. QA bands were used to obtain the de-clouded and initial label images. Then, a 502 × 502 pixel image set containing the seven OLI bands, elevation data, and the corresponding initial water label data set were obtained by cropping at 0.135°. Finally, the initial water label dataset was noise-corrected to obtain a standard water label library.

A comparison of labels before and after noise correction is shown in Fig. 7. Among them, the first column is the standard false color synthesis image of remote sensing image, the second column is the water body information extracted from QA band, and the third column is the water body label after noise correction. A-C shows the extraction of water bodies in mountainous areas, and it can be seen that the QA band tends to misclassify mountain shadows as water. This situation is improved after noise correction, which can shield the interference caused by

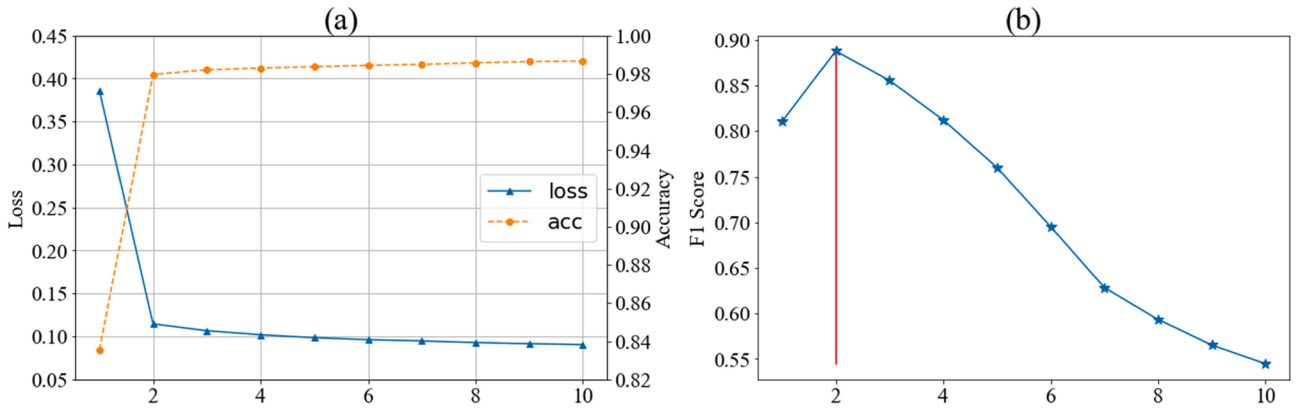


Fig. 8. Training iteration curves.

Table 3
Model parameters.

Layer (type)	Output shape	Param	Connected to	GEE module
input_1 (InputLayer)	[(None, 7, 7, 8)]	0		ee.Image
conv2d (Conv2D)	(None, 5, 5, 16)	1,168	input_1[0][0]	ee.Kernel.convolve
conv2d_1 (Conv2D)	(None, 3, 3, 32)	4,640	conv2d[0][0]	ee.Kernel.convolve
tf_op_layer_strided_slice	(None, 3, 3, 7)	0	input_1[0][0]	ee.Image.select
concatenate (Concatenate)	(None, 3, 3, 39)	0	conv2d_1[0][0] tf_op_layer_strided_slice[0][0]	ee.Image.cat
conv2d_2 (Conv2D)	(None, 1, 1, 64)	23,104	concatenate[0][0]	ee.Kernel.convolve
tf_op_layer_strided_slice_1	(None, 1, 1, 7)	0	input_1[0][0]	ee.Image.select
concatenate_1 (Concatenate)	(None, 1, 1, 71)	0	conv2d_2[0][0] tf_op_layer_strided_slice_1[0][0]	ee.Image.cat
conv2d_3 (Conv2D)	(None, 1, 1, 128)	9,216	concatenate_1[0][0]	ee.Kernel.convolve
conv2d_4 (Conv2D)	(None, 1, 1, 2)	258	conv2d_3[0][0]	ee.Kernel.convolve
Total params: 38,514 Trainable params: 38,514 Non-trainable params: 0				

mountain shadows during the extraction of water bodies. D-F are the labels of fine rivers, and we can see that the QA band is less effective for the extraction of fine rivers, and there is the problem of river breakage or even omission. After noise correction, the pixels of the underestimated water bodies are extracted, and the continuity of the fine water bodies was further improved. After that, 369 labels corresponding to each of the data image sets were obtained and converted into a pixel-based CNN trainable training data volume of 4.7 million.

4.2. Model training and prediction

The pixel-based CNN model was initialized and 369 images were selected to construct seven input images with blue, green, red, NIR, MNDWI, SWIR1, LEB, and DEM features. The format was transformed with the 369 labels to construct 4.7 million deep learning training samples that were $7 \times 7 \times 8$ (all bands are 8) for the input and $1 \times 1 \times 2$ (one-hot encoding of the labeled data) for the output. Of these, 2.1 million were water samples and 2.6 million were non-water samples, with a balanced sample size of 1:1.2. This loop was iterated 10 times. Fig. 8 shows the training iteration curve, where (a) shows the accuracy and loss of each image in 10 iterations and (b) shows the F1 score of the model for 10 iterations. The accuracy steadily improved while the loss gradually decreased, with the 10th training accuracy exceeding 98.66 %, indicating that the above seven features helped the model to train and extract water bodies, and are better features for water segmentation. F1 score in Fig. 8 (b) is a statistical measure of the accuracy of a dichotomous model, which takes into account both the precision and recall of the classification model, and a higher F1 score indicates a higher quality model. By analyzing the iteration curves, it was found that accuracy and loss did not change significantly after the second iteration, while F1 score was the highest. Therefore, the model obtained from the second training iteration was chosen as the classification model for the extraction of water bodies in this study.

4.3. Water segmentation results and accuracy

4.3.1. Online deployment of models

The structure of the pixel-based CNN model is shown in Table 3, with input, conv2d, concatenate, and slice corresponding to the ee.Image, ee.Kernel.convolve, ee.Image.cat, and ee.Image.select modules in GEE, respectively. A model conversion function was built in Python to convert CNN-based deep learning architectures to GEE format, and the code has been open-sourced in GitHub.

4.3.2. Results of interannual water extraction in Lake Baikal

Fig. 9 shows the annual water extraction results from 2013 to 2021, the 2021 water segmentation results, and a demonstration of the extraction results of some rivers and water bodies. It can be seen that the AI and remote sensing big data fusion method can effectively extract large-area lakes such as Lake Baikal and Lake Khovsgol. Similarly, the fine Selenge River was clearly displayed.

Figs. 10 and 11 show the interannual distribution of water bodies in some tributaries of the Selenge River and the Selenge River delta, respectively. The Selenge River, the most important source of water for Lake Baikal, flows into Lake Baikal after the Selenge River delta. It can be seen from the figure that the curved and fine Selenge River was extracted, and the continuity of the extracted rivers is strong. Fine rivers in the delta as well as small area lakes were also extracted.

We counted the area of water bodies within the Lake Baikal basin. It was compared with the water level changes. The water level change

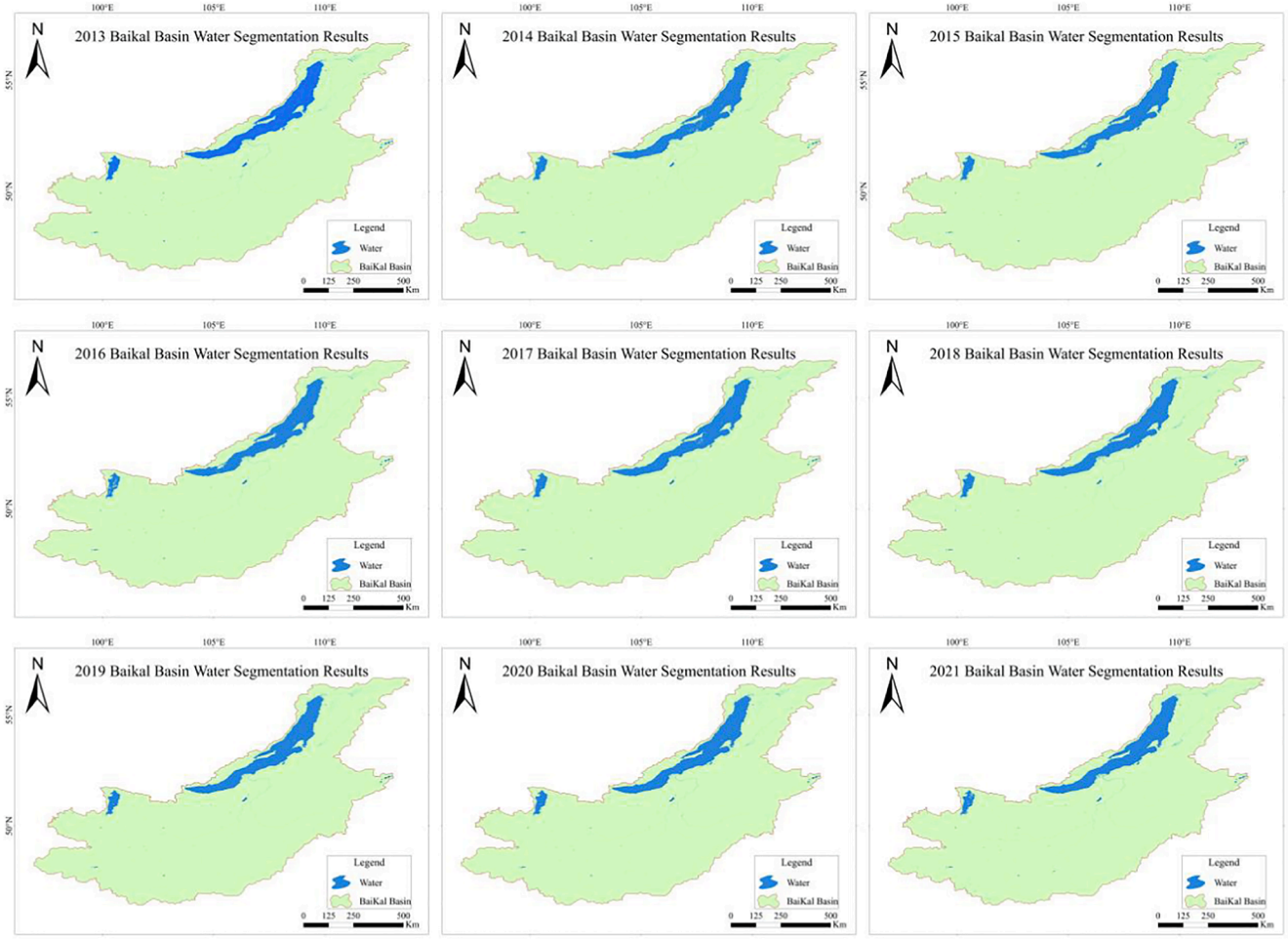


Fig. 9. Interannual spatial and temporal variability of water bodies during 2013–2021 in the Lake Baikal Basin.

information was taken from the annual report of the Ministry of Natural Resources and Environment of the Russian Federation on the state of Lake Baikal and measures for its protection, which contains information on the water level of Lake Baikal from 1980 to 2019. From 2013 to 2021, the water area of the Lake Baikal basin varies from 37,201 to 37,993 km^2, with an average value of 37,532 km^2 or 6.6 % of the total basin area. Overlaying this with the annual report on the state of Lake Baikal and protection measures (Fig. 12), it is clear that the water level of Lake Baikal decreased between 2013 and 2015 and remained low from 2015 to 2017, with a sudden increase in water level in 2018 before it dropped again in 2019.

4.3.3. Accuracy evaluation

The validation points were manually selected to validate the accuracy of water products extracted from the Pixel-based CNN model, QA band, and MNDWI using Otsu thresholds, respectively. Validation points were selected on Google Earth Pro for the three years 2013, 2017, and 2021. A total of 4,798 validation points were selected, including 2,262 validation points for water bodies, 1,847 validation points for others and 689 validation points for shadows. The confusion matrix is listed in Table 4. The three-year accuracy for 2013, 2017, and 2021, of the AI remote sensing big data fusion method was 92.9 %, 92.7 %, and 92.4 %, respectively, and the Kappa coefficients were 0.86, 0.85, and 0.85, respectively. The accuracy of the QA band and NDWI with Otsu thresholds method were lower than that of the noise-corrected deep learning method. As can be seen from Table 4, the water body products obtained from the Pixel-based CNN model for the above three years are able to reduce the interference of shadows in comparison with the water body information in the QA band and MNDWI threshold method, and the misclassification of water bodies into shadows is improved. It can be seen that MNDWI threshold method, compared with other methods, is more likely to misclassify shadows as water. Therefore, using the threshold method to do water classification needs to overcome the problem of difficulty in defining thresholds between water and shadows. The averages of Acc, Recall, MIoU, FWR, FWR, Kappa for 3 years were counted, and it was found that the Pixel-based CNN method performed better under these indicators. In summary, the method in this paper gets better performance in water body extraction and shadow pixels avoidance.

Among them, the calculation formula of these indicators are as follow:

$$p_o = Acc = \frac{TP + TN}{TP + TP + FP + FN} \tag{1}$$

$$\text{Recall} = \frac{TP}{TP + FN} \times 100\% \tag{2}$$

$$MIoU = \frac{1}{k+1} \sum_{i=0}^{k} \frac{TP}{FN + TP + FP} \times 100\% \tag{3}$$

$$TWR = \frac{TP}{FP + TP} \times 100\% \tag{4}$$

$$FWR = \frac{FP}{FP + TP} \times 100\% \tag{5}$$

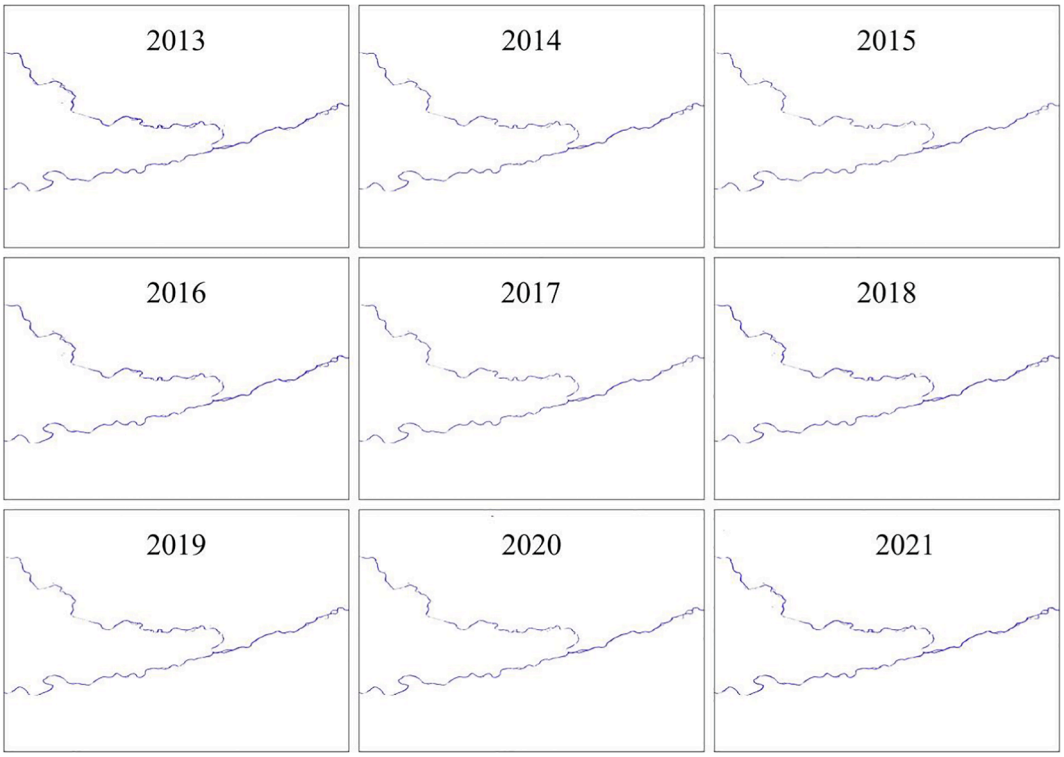


Fig. 10. Interannual water bodies in the Selenge River tributaries from 2013 to 2021.

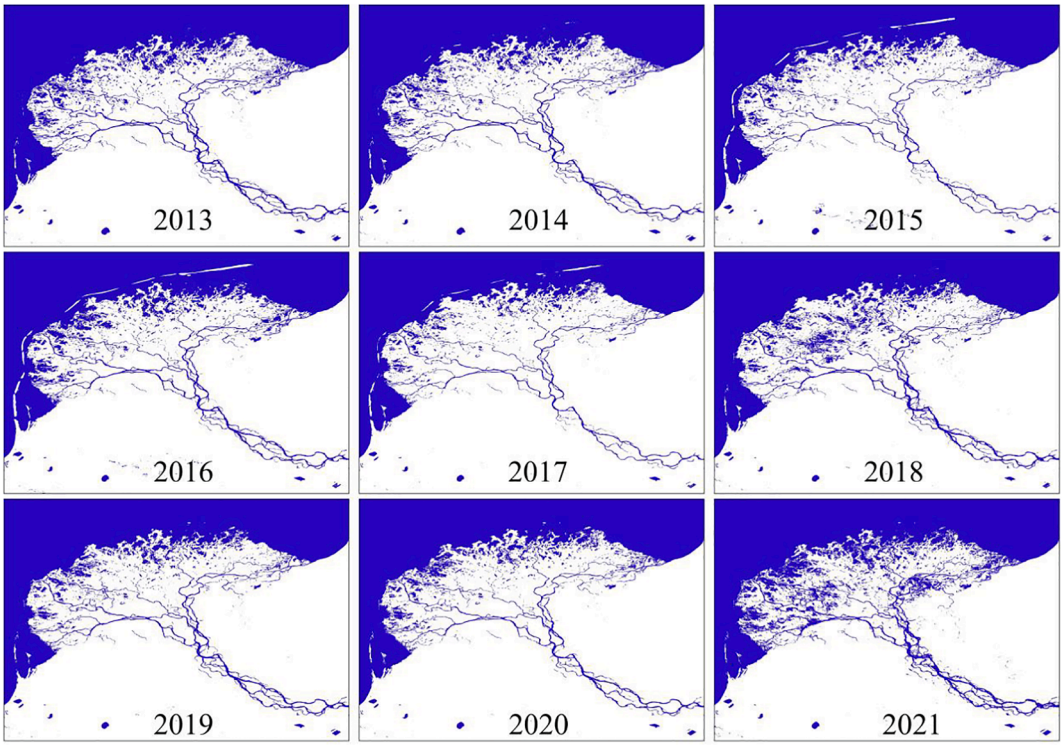


Fig. 11. Interannual water bodies in the Selenge River Delta from 2013 to 2021.

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

$$p_e = \frac{(TP + TN) \times (TP + FP) + (FP + FN) \times (TN + FN)}{N^2} \tag{7}$$

True Positive (TP) is the number of pixels whose label and prediction are both “water.” True Negative (TN) is the pixel whose label and prediction are both “ non-water.” False Positive (FP) represents the number of pixels with the label ”water“ and predicted as ” non-water,“ and False Negative (FN) represents the number of pixels with the label ” non-water

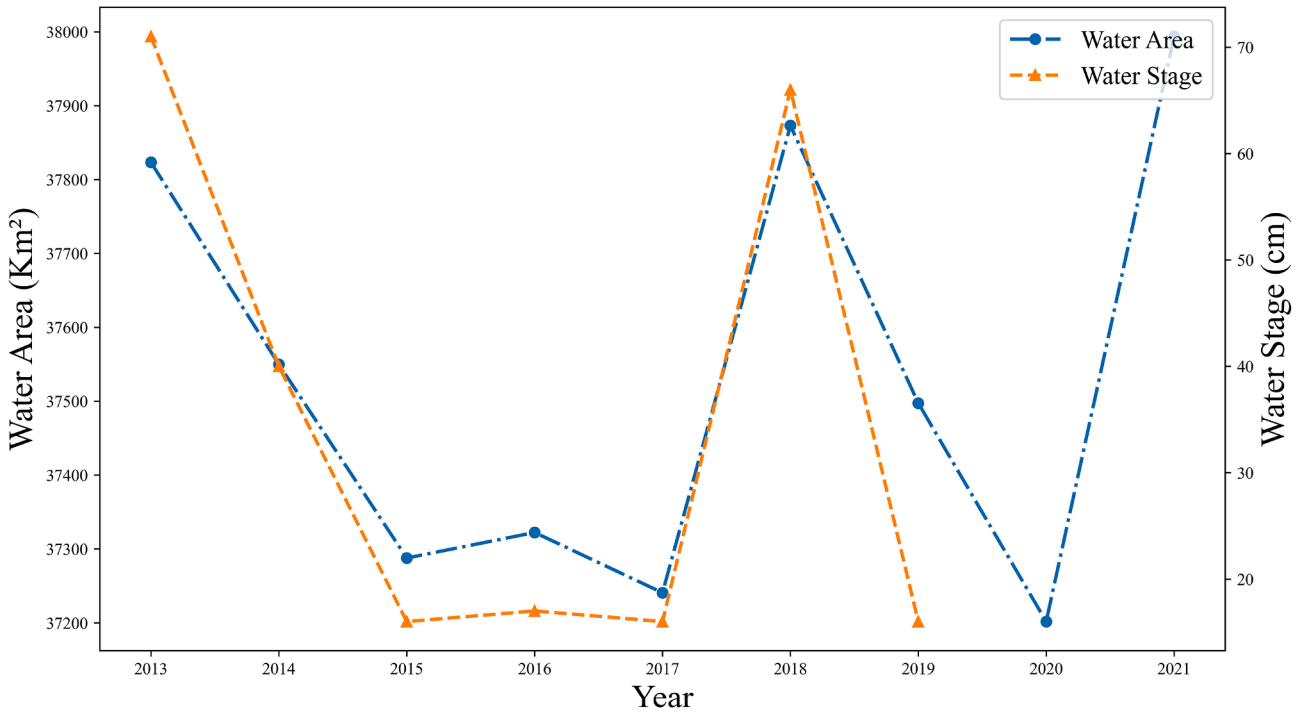


Fig. 12. Interannual variations in the basin area and water levels.

Table 4
Confusion matrix for water classification.

Methos	Year	2013				2017			
	Ground truth / Prediction	Water	Non-water		SUM	Water	Non-water		SUM
			Shadow	Others			Shadow	Others	
Pixel-based CNN	Water	637	11	11	659	671	13	9	693
	Non-water	89	213	611	913	94	222	594	910
	SUM	726	846		1572	765	838		1603
QA band	Water	617	27	9	653	675	32	13	720
	Non-water	109	197	613	919	90	203	590	883
	SUM	726	846		1572	765	838		1603
MNDWI	Water	632	37	18	687	663	42	13	718
	Non-water	94	187	604	885	102	193	590	885
	SUM	726	846		1572	765	838		1603

Methos	Year	2021			
	Ground truth / Prediction	Water	Non-water		SUM
			Shadow	Others	
Pixel-based CNN	Water	679	13	17	709
	Non-water	92	217	605	914
	SUM	771	852		1623
QA band	Water	664	29	29	722
	Non-water	107	201	593	901
	SUM	771	852		1623
MNDWI	Water	667	43	27	737
	Non-water	104	187	595	886
	SUM	771	852		1623

	QA	MNDWI	CNN
Acc	90.73%	90%	**92.73%**
Recall	86.45%	86.74%	**87.84%**
MIoU	81.47%	80.36%	**85.06%**
TWR	93.4%	91.61%	**96.42%**
FWR	6.6%	8.39%	**3.58%**
Kappa	0.81	0.8	**0.85**

Table 5
Comparison of different band weights.

Bands	Blue	Green	Red	NIR	SWIR1	MNDWI	LEB	DEM
Weight	11.34	11.12	11.60	12.31	11.63	12.95	13.27	11.55
Ratio / %	11.84	11.61	12.11	12.85	12.15	13.52	13.85	12.06

Table 6
Efficiency comparison between local computation environment and GEE cloud computing environment.

	Local computation environment	**Efficiency**	**GEE cloud computing environment**	**Efficiency**
Image screening and acquiring	Cloud volume filtering, month filtering, data download, with cloud area filling	It takes around 10 days to complete one year of 65 Cloud-free Imaging and requires around 100 GB of local disk	QA band cloud removing, month screening	Seconds
Pre-processing	Reflectance calculation, image cropping to regular size	1–2 Days	Reflectance calculation, free-cloud images mosaic	Seconds
Deep Learning Prediction	Cropped image prediction	3–4 Days	Deep learning weights for one-period mosaic image prediction	1–2 h
Post-processing	Splicing, seamless mosaic	1–2 Days	/	/
Total model deployment time	**15–18 Days**		**1–2 h**	

“ and predicted as ”water.“.

5. Discussion

This study proposes an automated water segmentation method for AI big data fusion, using a pixel-based CNN model as the water segmentation model, and constructs a local deep learning environment and a model structure parsing procedure, combining the parsed weights and each layer of the model’s network in coordination with the method in the GEE cloud computing platform. Automated label noise correction and rapid deployment of training weights to predict water bodies were achieved. Using the Lake Baikal Basin as an example, the surface water distribution for 2013–2021 was predicted and extracted at the basin scale using sample strips as training data.

Table 5 provides and compares the first convolution parameters of the incoming band weights. Eight bands have similar weight information. The highest weight is in the LEB at 13.85 % and the lowest weight in the green band at 11.61 %. The top three in terms of weight are the LEB, MNDWI and NIR. This indicates that these three bands play an important role in the water segmentation process in the Lake Baikal Basin, confirming the superiority of MNDWI as a water segmentation index.

In summary, deep learning helps to efficiently discover feature information, which facilitates the exploration of ways to construct features for water bodies or other information. Knowledge mining is achieved through the backward propagation of data-driven models. For example, when analyzing the weights, redundant information can be removed from the neural network, where the proportion of weights is much smaller than other feature bands, or meaningful features can be uncovered in the process of extracting feature information from the neural network and exploring the mechanism by which these features operate.

The deep learning-enabled automated extraction method for water bodies in the GEE Lake Baikal Basin data reduced the data acquisition difficulties, removed the redundant image screening, downloading, and slicing operations, and allowed for the batch acquisition of cloud-free images, greatly enhancing the rapid deployment of the model in the cloud. Table 6 demonstrates the efficiency of the local deep learning approach compared to the deep learning-enabled GEE online cloud computing approach. The deep learning-enabled GEE deep learning approach in the cloud significantly reduced time and labor costs in image data acquisition, data pre-processing, deep learning prediction, and image post-processing. It only took approximately 1–2 h to obtain the distribution of surface water bodies in the Lake Baikal Basin using data from one year.

Water bodies at the watershed scale can be efficiently obtained using this method. However, improvements or in-depth studies can be made in the following areas in the future. (1) Optimization of labels. Noise correction in QA band can reduce the interference of noise such as mountain shadow, however, manual participation in labeling water bodies or filtering labels can get higher accuracy water body labels currently. (2) Optimization of mosaic. In this study, the image mosaic process was restricted in months to June to August each year, and the median of cloud-free images was used to obtain images of the growing season of the study area for the entire region for 9 years. Restricting the month and median processing can alleviate the problem of uneven mosaic caused by inconsistent radiation intensity of features at different times in large scenarios, but there is still the problem of color difference in some cloudy areas. This problem can be solved by using a better declouding method or a color correction algorithm. (3) A better combination of characteristic bands. The characteristic bands in this paper are visible light, NIR, SWIR1, MNDWI, LEB, and DEM. Exploring bands with higher sensitivity to water bodies can ease the training difficulty of the model. In addition, the combination with synthetic aperture radar images can explore the real-time water extraction under multi-cloud coverage. (4) Method application. The method can provide a reference method for large scale remote sensing applications, not only in water body extraction but also for more land cover features mapping. However, it needs to be tailored for different tasks in terms of models, datasets, etc.

6. Conclusion

This study proposes an automated water extraction method with AI big data fusion and extracts and analyzes interannual surface water in Lake Baikal Basin from 2013 to 2021. First of all, it bridges the gap in traditional water classification, which is difficult to deploy empowered by deep learning on a large scale. By combining local training with real-time prediction, deep learning models are automatically deployed on the GEE cloud computing platform. It automatically assigns weights from the local model to GEE and transforms the neural network layers in the pixel-based CNN into GEE modules. Additionally, the efficiency of water body product acquisition was greatly improved; the overall accuracy reached more than 92 % as well. Secondly, the QA band of the combined images removed the interference of clouds and cloud shadows, and accurate water labels were quickly obtained using the water information from the QA band. The overfitting phenomenon of a single scene was reduced by the iterative training of large samples and multiple scenes. Thirdly, in-terannual surface water body products were obtained for

2013–2021. Our results show that the mean area of water bodies in the basin is 37,500 km^2, representing 6.6 % of the total basin area. The water body area of the Lake Baikal Basin has fluctuated over the years but has not changed significantly since 2013.

This study helps enhance the advancement of large-scale water-body-centered remote sensing studies. It provides examples of basin scenarios for the application of dynamic monitoring of surface water bodies globally and in large regions. This depth further enhances the applicability of the GEE platform, extending it from the traditional machine-learning domain to the deep learning domain. Models and methods have been uploaded to GitHub (https://github.com/CaryLee17/water_gee).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This research was funded by the National Natural Science Foundation of China (grant number 32161143025), the Strategic Priority Research Program (Class A) of the Chinese Academy of Sciences (grant number XDA20030200), Mongolian Foundation for Science and Technology, and National University of Mongolia (grant number P2020-3779).

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jag.2022.102928.

References

Badrinarayanan, V., Kendall, A., Cipolla, R., 2017. SegNet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. 1.

Bezrukov, L.B., Budnev, N.M., Galperin, M.D., Dzhilkibaev, Z.A.M., Lanin, O.Y., Tarashansky, B.A., 1990. Measurement of light-absorption coefficient in Lake Baikal water media. Okeanologiya 30, 1022–1026.

Bolgrien, D.W., Granin, N.G., Levin, L., 1995. Surface temperature dynamics of Lake Baikal observed from AVHRR lmages. Photogramm. Eng. Remote Sens. 61.

Glf, A., Hm, A., Rf, B., Srp, B., 2014. Automated water extraction index: A new technique for surface water mapping using Landsat imagery. Remote Sens. Environ. 140, 23–35.

Gong, P., Wang, J., Yu, L., Zhao, Y., Zhao, Y., Liang, L., Niu, Z., Huang, X., Fu, H., Liu, S., Li, C., Li, X., Fu, W., Liu, C., Xu, Y., Wang, X., Cheng, Q., Hu, L., Yao, W., Zhang, H., Zhu, P., Zhao, Z., Zhang, H., Zheng, Y., Ji, L., Zhang, Y., Chen, H., Yan, A., Guo, J., Yu, L., Wang, L., Liu, X., Shi, T., Zhu, M., Chen, Y., Yang, G., Tang, P., Xu, B., Giri, C., Clinton, N., Zhu, Z., Chen, J., Chen, J., 2013. Finer resolution observation and monitoring of global land cover: first mapping results with Landsat TM and ETM+ data. Int. J. Remote Sens. 34, 2607–2654.

Gong, P., Liu, H., Zhang, M., Li, C., Wang, J., Huang, H., Clinton, N., Ji, L., Li, W., Bai, Y., Chen, B., Xu, B., Zhu, Z., Yuan, C., Ping Suen, H., Guo, J., Xu, N., Li, W., Zhao, Y., Yang, J., Yu, C., Wang, X.i., Fu, H., Yu, L.e., Dronova, I., Hui, F., Cheng, X., Shi, X., Xiao, F., Liu, Q., Song, L., 2019. Stable classification with limited sample: transferring a 30-m resolution sample set collected in 2015 to mapping 10-m resolution global land cover in 2017. Sci. Bull. 64 (6), 370–373.

Li, J., Wang, C., Xu, L., Wu, F., Zhang, H., Zhang, B., 2021a. Multitemporal water extraction of Dongting Lake and Poyang Lake based on an automatic water extraction and dynamic monitoring framework. Remote Sens. 13 (5), 865.

Li, K., Wang, J., Yao, J., 2021b. Effectiveness of machine learning methods for water segmentation with ROI as the label: A case study of the Tuul River in Mongolia. Int. J. Appl. Earth Obs. Geoinf. 103, 102497.

Li, L., Yan, Z., Shen, Q., Cheng, G., Gao, L., Zhang, B., 2019. Water body extraction from very high spatial resolution remote sensing data based on fully convolutional networks. Remote Sens. 11 (10), 1162.

Long, J., Shelhamer, E., Darrell, T., 2015. Fully Convolutional Networks for Semantic Segmentation. IEEE Trans. Pattern Anal. Mach. Intell. 39, 640–651.

Mamaev, O.I., 1987. The thermal regime of Lake Baikal. DOKLADY AKADEMII NAUK SSSR, 292, 1477–1482.

Mcfeeters, S.K., 1996. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. Int. J. Remote Sens. 17, 1425–1432.

Oshchepkov, S.L., Shlyakhova, L.A., 1987. Determination of structure and hydrosol concentration by optical measurements. IZVESTIYA AKADEMII NAUK SSSR FIZIKA ATMOSFERY I OKEANA 23, 75–83.

Qiu, S., Zhu, Z., He, B., 2019. Fmask 4.0: Improved cloud and cloud shadow detection in Landsats 4–8 and Sentinel-2 imagery. Remote Sens. Environ. 231, 111205.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. Cham. Springer International Publishing, 234-241.

Soloviev, S.L., Kovachev, S.A., Misharina, L.A., Ufimtsev, G.F., 1989. SEISMIC ACTIVITY OF TRANSVERSE RUPTURES IN THE OLKHON-SVYATONOSSKY ZONE OF THE LAKE BAIKAL. DOKLADY AKADEMII NAUK SSSR 309, 61–64.

USGS USGS. Landsat Collection 2 Quality Assessment Bands.

Wang, F., Li, C., Ma, T., Liu, M., Zhang, Z., 2021. A modified method for water body information rapid extracting from remote sensing image. Yangtze River, 52, 223-228.

Weng, L., Xu, Y., Xia, M., Zhang, Y., Xu, Y., 2020. Water areas segmentation from remote sensing images using a separable residual SegNet network. Int. J. Geo-Inform. 9, 256.

Xu, H., 2006. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. Int. J. Remote Sens. 27, 3025–3033.

Yang, X., Zhao, S., Qin, X., Zhao, N., Liang, L., 2017. Mapping of urban surface water bodies from sentinel-2 MSI imagery at 10 m resolution via NDWI-based image sharpening. Remote Sens. 9 (6), 596.

Zhang, X., Liu, L., Chen, X., Gao, Y., Mi, J., 2020. GLC_FCS30: Global land-cover product with fine classification system at 30 m using time-series Landsat imagery.

Zheng, Z.L., 2002. Probe into the natural and cultural development of the Lake Baykal. Siberian Studies.

Zonenshain, L.P., Savostin, L.A., 1981. Geodynamics of the Baikal rift-zone and plate-tectonics of Asia. Tectonophysics 76, 1–45.