Grain Size Distribution of Bedload Transport in a Glaciated Catchment
Glacial Lake Detection from GaoFen-2 Multispectral Imagery Using an Integrated Nonlocal Active Contour Approach: A Case Study of the Altai Mountains, Northern Xinjiang Province

Meimei Zhang 1, Fang Chen 1,2,3,* and Bangsen Tian 1

1 Key Laboratory of Digital Earth Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, No. 9 Dengzhuang South Road, Beijing 100094, China; zhangmm@radi.ac.cn (M.Z.); tianbs@radi.ac.cn (B.T.)
2 University of Chinese Academy of Sciences, Beijing 100049, China
3 Hainan Key Laboratory of Earth Observation, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Sanya 572029, China
* Correspondence: chenfang_group@radi.ac.cn; Tel.: +86-010-8217-8105

Received: 15 March 2018; Accepted: 6 April 2018; Published: 10 April 2018

Abstract: Due to recent global climate change, glacial lake outburst floods (GLOFs) have become a serious problem in many high mountain areas. Accurately and rapidly mapping glacial lakes is the basis of other glacial lake studies that are associated with water resources management, flood hazard assessment, and climate change. Most glacial lake detection studies have mainly used medium to coarse resolution images, whose application is limited to large lakes. Because small glacial lakes are abundant and because changes in these lakes are small and occur around the lake shores, fine-resolution satellite imagery is required for adequate assessments. In addition, the existing detection methods are mainly based on simply applying a threshold on various normalized difference water indices (NDWIs); this cannot give appropriate results for glacial lakes that have a wide range of turbidity, mineral, and chlorophyll content. In the present study, we propose a region-dependent framework to overcome the spectral heterogeneity of glacial lake areas using a nonlocal active contour model that is integrated with the NDWI. As the first trial, the glacial lakes were detected using high-resolution GaoFen-2 multispectral imagery in the test site of Altai Mountains (northern Xinjiang Province). The validation of the results was carried out using the manually digitized lake boundaries. The average probabilities of false positives $P_{FP}$ and false negatives $P_{FN}$ were found to be 0.0106 and 0.0039, respectively. After taking into consideration the spectral features of the water and making slight NDWI threshold adjustments, this method can also be used for lake detection in any glaciated environment elsewhere in the world.

Keywords: nonlocal active contour; glacial lakes; Altai Mountains; high resolution; GaoFen-2

1. Introduction

Glacial lakes are a sensitive indicator of global climate changes. With the increase in air temperature and changing precipitation, mountain glaciers have been widely retreating, thinning, and even disappearing in recent decades; this facilitates the formation and rapid development of glacial lakes [1–4]. The continued increase in size of glacial lakes may result in the sudden breaching of lake dams and can produce catastrophic glacial lake outburst floods (GLOFs), causing severe damage to downstream settlements as well as other structures and the natural environment worldwide [5,6]. In particular, some GLOFs in transboundary basins exert significant impacts on bordering countries.
One striking example is the Cirenmaco GLOF, which occurred in 1981 and destroyed the China-Nepal Friendship Bridge and the Sun-Kosi Hydropower Station, and which also blocked roads and killed hundreds of people [7,8]. Therefore, timely monitoring and mapping of glacial lakes is critically important for the analysis of the spatiotemporal development of the lakes, and further determine the spatial variation in shoreline position for each lake. This will aid the better understanding of the processes driving lake basin expansion, as well as the prediction of lake outburst events.

Field measurements constitute an efficient way of detecting and mapping glacial lakes. However, due to the harsh natural environment, rapid expansion, and the vast numbers of lakes that are found in glaciated regions, measurements of this type are time- and labor-intensive, meaning that extensive and continuous field observations are not possible [9]. Satellite remote sensing offers the most feasible technique for monitoring glacial lakes and the related changes in mountainous areas. Optical sensors, such as the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER), SPOT-5, as well as Landsat data and Sentinel-2a/b, have been widely used to derive the spatial distribution of glacial lakes over large areas and long temporal scales [10–14]. In spite of this, the coarse spatial resolution of these sensors is a major limitation for the identification of the small glacial lakes that are common in regions of human activity; subsequent tracking of the tiny changes that occur along the lakeshores is also not possible while using this type of data [15,16]. Because glacial lake can cause or trigger natural hazards in several days and its small size, in this study, satellite with a short revisit cycle of less than ten days and a very high spatial resolution of about 1 m is necessary to monitor glacial lake changes.

So far, glacial lakes have been mainly mapped and monitored using manual interpretation and digitization based on optical data [8,11,17–20]. A few studies have used semi-automatic methods for classifying glacial lakes: these methods include single-band thresholding [21], band ratioing [13], normalized difference water indices (NDWIs) [22], the use of modified NDWIs [23], or other logic-based methods [24]. These methodologies, which make use of different bandwidths of the electromagnetic spectrum and identify an appropriate threshold yielding the most accurate lake extraction using a particular index, have proven to be simple as well as effective. These techniques are pixel-based and work well under conditions of homogeneous contrast between the water bodies and their adjacent backgrounds. In practice, however, these backgrounds are generally heterogeneous, both spatially and temporally, due to variations in the aquatic environment and imaging conditions, especially for glacial lakes that are characterized by differences in the amount of suspended solids, chlorophyll concentration, water depth, and snow/ice, or the presence of aquatic vegetation and mudflats [25,26]. All of these factors can complicate the spectral responses of lake water and pose special challenges for frequent and large-area lake mapping. Moreover, with increasing the spatial resolution of the satellite data, the heterogeneous reflectance between the lakes and the background becomes particularly prominent. In recent years, the active contour model has been shown to have many advantages for shoreline extraction because it can produce a smooth, continuous outline, and, at the same time, reduce the amount of manual post-processing, including reshaping and merging of the shoreline segments [9,27–29]. Active contour models based on the nonlocal principle, which makes use of the local homogeneity and includes the benefits of a region-based approach, have superior performance with respect to specific noise and spatial heterogeneity [30–32]. It is thus of great interest to investigate the use of nonlocal principles for high-resolution image segmentation combined with active contour models.

The present study was undertaken at a test site in the Altai Mountains in the north of Xinjiang Province and had two main objectives. The first was to introduce a nonlocal active contour-based glacial lake detector that can adaptively extract lake outlines from high-resolution imagery. Local NDWI information was incorporated into the proposed detector to compensate for the poor spectral contrast and the heterogeneous reflectance in the imagery. The second objective was to test the robustness of the new method under different environmental conditions and to evaluate its accuracy in comparison with existing classification techniques. The investigations that were performed in this study are
crucial in the context of hazard risk assessment and hydrological modeling. Furthermore, the derived mapping results with an average error of less than 0.5 pixels could lead to a better understanding of the characteristics of water surfaces in undulating mountainous terrain.

2. Study Area

The present study was carried out in parts of the Altai Mountains, as indicated by the red polygon in Figure 1. The Altai Mountains run from northwest to southeast and extend over four countries, namely China, Kazakhstan, Russia, and Mongolia. The elevation of this region varies from 1000 m to 4374 m above sea level (a.s.l) and decreases gradually towards the southeast [33,34].

Since the 1980s, the climate in the northwest of China has changed from warm and dry to warm and wet [1,35–37]. The hydrological cycle in the Altai area has accelerated and the frequency of extreme hydrologic events has increased [38–40]. Among the plateaus in western China, the glacier area in this region has shown the strongest recession, giving rise to a dramatic increase in the number of glacial lake systems [41–43]. Glacial lakes are usually located in low-lying areas near glaciers. They are formed by the glacier melt water and heavy rainfall. These lakes are relatively small and densely distributed over the region.

![Figure 1. Location map of the study area. The red rectangle indicates the coverage of the GaoFen-2 (GF-2) images.](image)

3. Methodology

3.1. Data and Workflow of the Proposed Method

Launched in 19 August 2014, the GaoFen-2 (GF-2) remote sensing satellite is a Chinese high-resolution optical satellite that is equipped with two panchromatic/multispectral (PAN/MSS) cameras with a combined swath of 45 km [17]. It is capable of collecting images with a ground sampling distance of 1 m in the PAN band and 4 m in the MSS bands, and a revisit cycle of five days [44,45]. The imagery contains four multi-spectral bands: 450–520 nm (blue); 520–590 nm (green); 630–690 nm (red); 770–890 nm (near-infrared); and, a panchromatic band (450–900 nm). Thus, GF-2 has the potential for use in land-cover change monitoring, high-accuracy mapping surveys, disaster assessments, and other scientific research. Here, we first attempt to use GF-2 imagery to map numerous small glacial lakes in high mountain regions.
The experimental data that were used are listed in Table 1. The six data cover two neighbouring tracks, each of which has three successive images. For automated detection and mapping of glacial lakes in the Altai Mountain region, imagery from late August or September is most suitable as this is when the lake areas are largest and the seasonal snow cover is at a minimum. In this study, all of the images were acquired synchronously by the PMS1 and PMS2 cameras on 8 September 2016. The GF-2 images were almost free of cloud (four images had cloud cover of less than 1.0%; two images had 8.0% cloud cover) and seasonal snow cover. All of the six GF-2 scenes were geometrically rectified using 1-arc-second elevation data from the Shuttle Radar Topography Mission (SRTM) that was produced in 2000. The slope image obtained from the 30-m SRTM was resampled to 1 m and then cropped to match the spatial resolution and area of the GF-2 image. By using a glacier inventory, we are able to define the type of delineated glacial lakes. Glacial lakes that were directly connected with glacier termini were classified as lakes in proglacial phase, which can be automatically identified according to the distance of 0 m between the lake and glacier. Other lake types are then classified manually based on the detailed analysis of remote sensing images. The glacier outline data referred to were from the Randolph Glacier Inventory (RGI) Version 5.0, 2015 [37] and downloaded from the Global Land Ice Measurements from Space initiative (GLIMS) website (http://www.glims.org/RGI/).

<table>
<thead>
<tr>
<th>Scene ID</th>
<th>Acquisition Date (yyyy-mm-dd hh:mm:ss)</th>
<th>Sensor</th>
<th>PAN and MSS Spatial Resolution</th>
<th>Cloud Cover (%)</th>
<th>Temporal Resolution (Day)</th>
<th>Swath Width (km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2791517</td>
<td>2016-09-08 13:48:42 PMS1 1 and 4 m</td>
<td>8.0</td>
<td>5</td>
<td>45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2790799</td>
<td>2016-09-08 13:48:42 PMS2 1 and 4 m</td>
<td>8.0</td>
<td>5</td>
<td>45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2791518</td>
<td>2016-09-08 13:48:45 PMS1 1 and 4 m</td>
<td>1.0</td>
<td>5</td>
<td>45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2790800</td>
<td>2016-09-08 13:48:45 PMS2 1 and 4 m</td>
<td>1.0</td>
<td>5</td>
<td>45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2791519</td>
<td>2016-09-08 13:48:48 PMS1 1 and 4 m</td>
<td>0.0</td>
<td>5</td>
<td>45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2790801</td>
<td>2016-09-08 13:48:48 PMS2 1 and 4 m</td>
<td>0.0</td>
<td>5</td>
<td>45</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a Sensor—PMS1/2 represents the first/second panchromatic and multispectral cameras; b PAN and MSS—Panchromatic and multispectral bands.

The proposed method consists of data description, preprocessing of GF-2 PMS images, image enhancement using pan-sharpening, generation of the initial glacial lake area, the core glacial lake segmentation algorithm used to refine the results, quantitative accuracy assessment, and the characteristics of glacial lake distribution in Altai Mountains. The platforms needed for the automated image processing are Interactive Data Language (IDL) and C++ codes. An overview of the workflow for the glacial lake detection is given in Figure 2. More details about the method will be described in the following sub-sections.

Figure 2. Workflow of the proposed method for glacial lake detection.
3.2. Conversion of PMS Digital Number (DN) to Surface Reflectance (SR)

The equations and absolute calibration coefficients that are given on the China Centre for Resources Satellite Data and Application website (http://www.cresda.com/CN/index.shtml) were incorporated into the method in order to obtain the top of atmosphere (TOA) reflectance \( R(\lambda) \). The linear transformation formula used for this is

\[
R(\lambda) = M(\lambda) \times \text{DN} + A(\lambda)
\]

where \( M(\lambda) \) and \( A(\lambda) \) represent the multiplicative and additive rescaling factors for each band with wavelength \( \lambda \), respectively.

After that, the obtained TOA reflectance image was atmospherically corrected using the fast line-of-sight atmospheric analysis of spectral hypercubes (FLAASH) module in ENVI 5.3 to retrieve the corrected surface reflectance (SR). Appropriate atmospheric and aerosol models were selected based on the season-latitude and the area that is covered by the image. In addition, orthorectification was needed to deal with the geometric distortions in the GF-2 images. Using 1-arc-second global elevation data from SRTM, the six GF-2 scenes were rectified on a UTM map projection with a 1 and 4-m-pixel spacing using the RPC orthorectification module in ENVI 5.3. The accuracy of all the orthorectifications in the region was within one pixel.

3.3. Pan-Sharpening of PMS Bands

The availability and use of the panchromatic band for pan-sharpening provided more local detail and thus increased the accuracy of the glacial lake boundary delineation process. Several existing data fusion methods were tried out for pan-sharpening the multispectral data. Recently [38], it has been found that the nearest neighbor diffusion (NNDiffuse) pan-sharpening algorithm works best as apart from enhancing the spatial resolution, it also enhances the salient spatial features while preserving the spectral fidelity.

For our study area, an analysis of the correlation between the original MSS bands and pan-sharpened bands was also conducted. As illustrated in Figure 3, the pan-sharpened band correlates well with the original multispectral band, giving correlation coefficients that are greater than 0.95 for all of the bands. The enhanced view of the glacier and glacial lake is given in Figure 4. These results show that the algorithm is capable of preserving very high-spatial details that are not visible in the multispectral images, while preserving spectral information from the multispectral images.

![Figure 3](image-url)  
*Figure 3.* Analysis of the correlation between the original multispectral (MSS) bands and the pan-sharpened bands. The scale refers to the frequency of pixel.
3.4. Generation of Potential Glacial Lakes

Many indices has been developed for water feature detection and mapping in remote sensing images. However, many spectral features are required in the formation of these indices since they combine information from at least four or five bands using an arithmetic formula. Moreover, the extensive application of a specific index at different sites remains a challenge due to the parameter tuning issue that arises for diverse contexts. The NDWI is among the widely used water indices because of its simplicity and effectiveness. For detecting glacial lakes in mountainous areas, the NDWI can provide good contrast. Therefore, the NDWI was selected as the image feature that is used to identify the location of glacial lakes. The associated equation is

$$NDWI = \frac{\text{band}_2 - \text{band}_4}{\text{band}_2 + \text{band}_4}$$

(2)

where \(\text{band}_2\) and \(\text{band}_4\) are the surface reflectance for the green and near infrared (NIR) bands of the GF-2 multispectral sensors, respectively.

The extent of water bodies could be distinguished from the surrounding land cover in NDWI images (Figure 5a,c). Although clouds and cloud shadows existed in the imagery (the yellow ellipses in Figure 5b), they had little effect on the lake detection because the lakes were rarely covered by clouds or cloud shadows. Besides, the observed clouds is quite thin, which make the separation between these thin clouds/cloud shadows and water bodies obvious in the NDWI image (see the yellow ellipses in Figure 5a). Some areas clearly show a response that is similar to that for the glacial lakes, coming out from the topographic shadows at the back of the hillside (the green ellipses in Figure 5c,d). Glaciers for
which there was not much moisture in the surroundings have a distinct spectral signature (the orange ellipse in Figure 5a).

In order to generate the potential glacial lake areas, a global-level threshold segmentation, which combined the NDWI with other variables, such as the NIR band and topographic factors, was applied to constrain commission errors that are induced by melting ice. A loose initial NDWI threshold of 0 was set to ensure that almost all the lakes were delineated. The DEM-derived surface slope for potential lake areas was required to be less than $10^\circ$ [46,47]. The smallest lakes that were detectable in the GF-2 data had an area of 16 pixels, i.e., $16\, \text{m}^2$. Finally, in order to achieve more accurate segmentation results, a buffer zone for each glacial lake unit was constructed such that the area of the buffer had the same area as the lake.

Figure 5. Normalized difference water indices (NDWI) images of glacial lakes and surrounding land cover for regions (a,c). (b,d) are corresponding false color composite (FCCs) of GF-2 images. Yellow ellipses mark clouds and cloud shadows that do not cover much lake area. The orange ellipse marks glaciers where there was little moisture in the surroundings; these have a distinct spectral signature. Green ellipses mark areas that had a similar response to the glacial lakes; mountain shadows can be seen at the back of the hillsides. The two red rectangles mark typical examples of spectrally heterogeneous glacial lakes that were used for the further investigation described in Section 4.1.

3.5. Glacial Lake Detection Algorithm

The key segmentation model that is proposed in this study consisted of a nonlocal active contour model for glacial lake detection in NDWI feature space. This model makes use of the nonlocal information between patches in the neighborhood of the pixel of interest and is suitable for non-homogeneous regions. The main idea is to minimize the nonlocal energy, which is defined by integrating the interactions between pairs of patches inside and outside the segmented regions using a level set method. The segmentation task can be tackled by solving the following problem:

$$\arg \min_{\theta} \varepsilon(\theta) = E(\theta) + \gamma|\partial \theta|$$  \hspace{1cm} (3)
where $\gamma$ is a weight factor for adjusting the regularization term $|\partial \theta|$ to the length of the boundary; this tends to result in smooth boundaries [39]. The nonlocal energy $E(\theta)$ is computed as:

$$E(\theta) = E(\theta) + E(\theta^c) = \int_{\theta} K(x, y) dx, dy$$

where $\theta^c$ is the region complementary to $\theta$ and $K(x, y)$ is the pairwise interaction kernel that is defined as $K(x, y) = G_\sigma(x, y)d(p_x, p_y)$. Here, $G_\sigma(x, y)$ is a truncated Gaussian function. $\sigma$ is a parameter that controls the decay between $G_\sigma(x, y)$ and the patch homogeneity. $d(p_x, p_y)$ denotes the weighted $L^2$ distance function [40].

A level set function $\phi : [0, 1]^2$ is introduced to represent the segmented lake region, $\theta$ [30]. The parametric active contours with shape gradients are then given by

$$\frac{\partial \phi}{\partial t} = - (\nabla E(\phi) + \gamma \nabla \| \partial \phi \| )$$

and the region update based on the steepest descent is

$$\phi^{(l+1)} = \phi^{(l)} - (\nabla E(\phi^{(l)}) + \gamma \nabla \| \partial \phi^{(l)} \| )$$

where $l$ is the time-step size.

In this study, the nonlocal active model was conducted for each buffer zone image of glacial lake. The glacial lake outlines that were obtained by global segmentation were used for the contour initialization. The convergence condition was based on the difference between two adjacent iterative energy, $E(\theta)$, and the defined maximum iteration number.

After completing the local refinement of the results for all of the glacial lakes, a post-processing step is necessary to ensure the quality of the lake dataset, which consists of removing small inlands and the river segments, adding lake attributes, like the image acquisition date, lake area, perimeter, and centroid, and generating a shoreline vector map. Specifically, some misclassifications of small inlands and rivers can be automatically deleted using the platform of ENVI/IDL by retaining the largest polygon in each buffer zone.

**3.6. Performance Evaluation**

Generally, the proposed method has three main outstanding contributions: (1) it can automatically detect a large amount of small glacial lakes using high-resolution remote sensing images; (2) the incorporating of NNDiffuse pan-sharpening algorithm is useful for enhancing the identification of glacial lakes; (3) the lake boundaries can be precisely extracted by taking into account the spatial heterogeneity between each glacial lake and its surrounding land.

In order to quantitatively validate the results, an accuracy assessment at the pixel level was performed on the glacial lake detection results based on the reference datasets; the referenced shorelines were manually digitized by visual interpretation of the panchromatic image using the multispectral image as an additional decision aid. Four numerical indices were used for quantitative evaluation of the proposed method: the probability of false positives $P_{FP}$, probability of false negatives $P_{FN}$, the F-measure, and the average error (AE). $P_{FP}$ and $P_{FN}$ are direct measurements of the segmentation performance [41,42] and can be calculated as

$$P_{FP} = \frac{num(AD \cap MD)}{num(MD)} \quad \text{and} \quad P_{FN} = \frac{num(\overline{AD} \cap MD)}{num(MD)}$$

where AD and MD refer to the raster classification maps that were converted from the automatically detected and manually digitized lake boundaries, respectively; and, $num(\cdot)$ denotes the number of
elements in the dataset. $P_{FP}$ and $P_{FN}$ also denote the commission and omission errors, respectively. The AE concerns the local segmentation performance [9], which is defined as

$$AE = \frac{1}{h} \sum_{t=0}^{h} t \ast \hat{p}(t) = \frac{N_{AD}}{N_{MD}}$$

(8)

where $t = 0, 1, 2, \ldots, h$ is the buffer layer index and $h$ is the total number of buffer layers. $N_{AD}$ is the number of pixels on the automatically detected shoreline that lie within the $t$-th buffer layer, and $N_{MD}$ is the number of pixels on the manually digitized shoreline. The F-measure is also introduced to balance the producer’s accuracy and user’s accuracy by performing a weighted average computation, according to the following equation:

$$F\text{-measure} = 2 \times \frac{PA \times UA}{PA + UA}$$

(9)

where PA indicates the producer’s accuracy and UA the user’s accuracy.

4. Results and Discussion

4.1. Mapping of All the Glacial Lakes in the Study Area

Using the proposed method, the detected glacial lake boundaries were overlaid on the six false color composite (FCC) GF-2 images; the results are shown in Figure 6. Three subregions that were selected as representative examples for validation are shown in detail. It should be noted that, in order to validate the size-dependent accuracy of the method, the selected lake samples should cover a wide range of lake size. The glacial lakes in the Altai Mountains of northern Xinjiang province have developed on both sides of the Kanas River and are more densely distributed on the eastern side due to abundant rainfall. Three types of glacial lakes can be distinguished in the study area, namely lakes in proglacial phase, lakes in glacier-detached phase, and lakes in non-glacial phase. Lakes in the proglacial phase are connected to glacier termini [48,49], while others are not directly in contact with current glaciers, which can be further categorized into lakes in glacier-detached phase and lakes in non-glacial phase based on the existence of one or more current glaciers in the catchment of the lake [50]. Few lakes in proglacial phase can be observed in the study area because of the rapid melting of glaciers. Lakes in glacier-detached phase are impounded by pronounced rock barriers or are embedded in undulating landscapes. In this area, they are found at some distance from the glaciers (Figure 6a,b). Lakes in the non-glacial phase are impounded by dams that are dominated by coarse blocks, debris, or a combination of both, most commonly formed from the Holocene terminal moraines, landslide deposits, or debris cones (Figure 6c).

Table 2 lists the statistical values of the probability of false positives, the probability of false negatives, average errors, and F-measure for the sample lakes, which are calculated using Equations (7)–(9). As shown in Table 2, the mean values of $P_{FP}$ and $P_{FN}$ are about 1.06% and 0.39%, respectively, with standard deviations of 0.0058 and 0.0041. The average error is less than one pixel with a standard deviation of 0.1530. The mean F-measure is 99.27% with a standard deviation of 0.0046. The evaluation results confirm that the proposed methods are sufficiently robust and perform well in detecting glacial lakes from Chinese GF-2 multispectral imagery.
Figure 6. The lake shorelines extracted within the study area. The left-hand side consists of a mosaic formed from six images. The red contours were detected using the proposed method. The right-hand images (a–c) show the three selected validation areas with manually digitized boundaries shown as yellow contours.

Table 2. Evaluation of the proposed glacial lake detection method for the sampled lakes shown in Figure 6.

<table>
<thead>
<tr>
<th>Evaluation Index</th>
<th>Max</th>
<th>Min</th>
<th>Mean</th>
<th>Median</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_{FP}$</td>
<td>0.0217</td>
<td>0.0014</td>
<td>0.0106</td>
<td>0.0103</td>
<td>0.0058</td>
</tr>
<tr>
<td>$P_{FN}$</td>
<td>0.0258</td>
<td>0.0009</td>
<td>0.0039</td>
<td>0.0024</td>
<td>0.0041</td>
</tr>
<tr>
<td>Average error (pixels)</td>
<td>1.2230</td>
<td>0.0856</td>
<td>0.4280</td>
<td>0.3973</td>
<td>0.1530</td>
</tr>
<tr>
<td>F-measure</td>
<td>0.9762</td>
<td>0.9988</td>
<td>0.9927</td>
<td>0.9936</td>
<td>0.0046</td>
</tr>
</tbody>
</table>

4.2. Distribution Characteristics of Glacial Lakes in Altai Mountains

Glacial lakes are highly variable in location, shape, and size due to complex physical conditions and alpine environments. Here, the spatial distribution of glacial lakes in the Altai Mountains is briefly analyzed by different areal size, elevation gradients, and lake types. A total of 1328 lakes with an area of 13.54 km$^2$ were identified in the study area: 33 lakes (0.98 km$^2$) in the proglacial phase, 1036 lakes (6.99 km$^2$) in the glacier-detached phase, and 259 lakes (5.57 km$^2$) in the non-glacial phase. Most of the glacial lakes are <0.1 km$^2$, with this area class comprising nearly 97.81% of the total number. But, larger lakes (area > 0.1 km$^2$) contribute 46.16% to the total area. Generally, lakes in the proglacial phase and glacier-detached phase are formed in the small depressions that have been eroded by glaciers and the rocks and have a smaller area. Lakes in non-glacial phase, on the other hand, are retained by vast blocks or debris dams and have large surface areas. The largest lake had a surface area of 0.73 km$^2$. The average lake area was 0.02 km$^2$ and the median area was 0.0041 km$^2$. Figure 7 shows the distribution of the glacial lakes of various size. The histogram of the areas of all the glacial lakes is strongly skewed toward smaller areas due to the large number of small lakes (Figure 7a). In order
to show the frequency distribution clearly, the glacial lakes were divided into two groups: those that are larger than $1.0 \times 10^3 \text{ m}^2$ (Figure 7b) and those with an area less than $1.0 \times 10^3 \text{ m}^2$ (Figure 7c). As there is much more uncertainty involved in the extraction of lakes with smaller areas, a statistical error analysis for the small glacial lakes (area $< 0.1 \text{ km}^2$) was carried out; the resulting averages were 1.98% for $P_{FP}$, 0.63% for $P_{FN}$, an AE of 0.6741 pixels, and an F-measure of 98.69%. These results reveal the great potential of the proposed method for monitoring small glacial lakes using high-resolution satellite imagery.

Figure 7. Histogram of glacial lake area for the study area: area distribution for (a) All lakes, (b) Lakes larger than $1.0 \times 10^3 \text{ m}^2$, and (c) Lakes smaller than $1.0 \times 10^3 \text{ m}^2$.

Besides, glacial lakes in the Altai Mountains show distinct altitudinal variations. The histograms of the number and area of glacial lakes shows similar distribution patterns against elevation (Figure 8a). The altitudes of all the glacial lakes mapped in 2016 ranged from 1800 m a.s.l. to 3400 m a.s.l.,
and the majority of them, are situated between 2400 m a.s.l. and 2800 m a.s.l., which accounted for 79.37% and 75.50% of the total number and area, respectively. The average elevation of all the glacial lakes is 2685 m a.s.l. However, the two types of lakes exhibit differences in their altitudinal distributions (Figure 8b), lakes in proglacial phase were not included given its relatively small number and area. Lakes in glacier-detached phase have obvious higher elevation than lakes that are in the non-glacial phase due to the close interaction with glaciers.

![Figure 8. Altitudinal characteristics of glacial lakes in Altai Mountains by (a) Frequency and area, and (b) Different types of lakes.](image)

### 4.3. Discussion

In reality, glacial lake extraction errors are most likely to occur at water boundaries. These areas are frequently flooded or inundated by highly dynamic water bodies and are likely to be mixed with backgrounds, such as aquatic vegetation, mudflats, or other noise to varying degrees. Two typical glacial lakes with different types of heterogeneous surroundings were selected as examples to demonstrate the advantages of the proposed method. The proposed method was also compared with the supervised support vector machine (SVM) [51,52] and unsupervised K-means clustering (KMC) [53]. The SVM uses the radial basis function as the kernel function and four multispectral bands of the GF-2 image for classification. The use of SVM needs a lot of user expertise and proper training samples, it was found to be highly sensitive to noise, while the KMC is often hindered in optically complex images, especially for the complex water conditions [54].

As illustrated in Figure 9, the shorelines that were extracted by the three methods overlapped well with each other except for some inconsistencies in the neighborhood of the mud (see the magenta ellipses in Figure 9a,b). In both cases, it can be seen that the SVM considerably overestimated the lake area. Although there was little difference between the results that were obtained using the KMC and our method, the proposed method was better at capturing the ambiguous lake boundaries. For the lake area that was characterized by clay, wet muddy soil, ripples, and a sediment layer that was filled with water in a relatively low-lying area (field photos in Figure 9a), it can be seen that the water–land interface was dominated by mixed reflectance. Clearly, the spectral heterogeneity depends on the size and type of the loose particles in the shallow water, as well as the fluidity of the water bodies. In addition, aquatic vegetation with high moisture content around the lake, such as alpine meadow and alpine grassland, introduces considerable noise that affects the extraction of the target water bodies (Figure 9b). Apart from these conditions, this method is also very effective for the detection of glacial lakes with turbid and small floating ice surface.

In order to further explore the differences between the three methods, for the areas described above, we conducted analyses of the NDWI and the surface reflectance profiles at the transition zones between land and water. It was found that, in this zone, the NDWI profile had a much sharper
boundary than the surface reflectance profile. This may explain why the SVM failed to achieve the desired accuracy even though more features were used. It was also found that the delineation of water bodies using the proposed method outperformed the KMC in non-homogeneous areas. This is because our method makes use of the local homogeneity principle, which is based on patches around the target pixels, while the KMC applies global statistics to generate clustered classes.

![Figure 9](image.png)

**Figure 9.** Experimental results showing comparison between the proposed method and conventional classification methods for different heterogeneous environments: (a) Glacial lake surrounded by clay, wet muddy soil, and ripples, etc.; (b) Lake surrounded by aquatic vegetation with high moisture content. The figures in the first row show the final shorelines overlaid on the FCC image. The insets show zoomed out views of the area marked by the magenta ellipses. The second row consists of field photographs taken in the same area. The third row shows profiles of the NDWI and surface reflectance. The location of the profiles is indicated by the white dotted arrows in the insets in the first row.

When comparing with recently published literature according to the research by Emmer et al. 2018 [55], the proposed framework has good performance in automatically detecting small glacial lakes from the aspects of data and method used. In terms of data, Landsat and MODIS are the most popular dataset being used for the extraction of glacial lakes [2,26,56–58]. However, because most of glacial lakes are of small size, the utilization of these dataset will possibly lead to large underestimation of results. The SRTM Water Body Data (SWBD), as produced by the US National Geospatial Intelligence Agency in February 2000, also provides lake-water boundaries [59]. But, these data sets only reflect a snapshot of glacial lake status, and consequently, their use is severely limited when compared with the Landsat and MODIS products that can reveal the dynamics of lake-area change. Moreover, wide swath Sentinel-2 Multispectral Image (MSI) sensor imagery can support the buildup and frequent update of globally lake maps to be used for further analysis [60]. In terms of method, visual interpretation is still widely used for glacial lake delineation [61]. Although it can guarantee high quality control, it was largely dependent on the acquisition time of the selected images and the experiences of the interpreters.
Water index values that were derived from NDWI\textsubscript{McFeeters}, NDWI\textsubscript{Xu}, AWEI\textsubscript{non-shadow}, and AWEI\textsubscript{shadow} presented an obvious difference between water and non-water [41]. However, the performance of a water index are unstable and vary greatly depending on atmospheric and topographic conditions, and lake type. SAR images have been gradually used by some researchers to map glacial lakes, which, regardless of weather and cloud, serving as an effective supplement for optical data in high mountains area. But, automatic methods have uncertainties in mapping the lake outlines because of the speckle noise of the radar images and the wind that locally impacts the backscattering intensity [62]. The proposed method exploits the capability of NDWI to locate the approximate glacial lake outline quickly around the evolving contour. The nonlocal active contour model can efficiently identify the shoreline by integrating the nonlocal interactions between pairs of patches inside and outside of the lake. It can also be conveniently applied to the multitemporal and multifeature images.

5. Conclusions

In diverse glacial environments, spectral non-uniformity frequently occurs due to the different land cover types, including aquatic vegetation, silt, shadow, etc., that are present. This makes glacial lake identification using spectral methods difficult. To overcome the problems arising from heterogeneous spectral responses, this paper has presented a novel glacial lake detection method that exploits NDWI features and an active contour model that is based on the nonlocal processing principle. The capability of high-resolution GaoFen-2 multispectral imagery for carrying out lake detection has been explored for the first time. The 1-m spatial resolution panchromatic band proved to be very useful for enhancing the spatial resolution of the multispectral bands using the NNDiffuse pan-sharpening algorithm. The visual enhancement in the image was clearly observable.

Glacial lakes were then delineated by applying the proposed method to the pan-sharpened GaoFen-2 multispectral images, demonstrating that the proposed method can effectively reduce spurious detections in heterogeneous lake regions while producing smoother outlines as compared to other conventional approaches. Validation using the lake shorelines that were derived from visual interpretation showed that the proposed approach can effectively map glacial lakes with a high degree of accuracy. Specifically, small glacial lakes were accurately classified with an AE of 0.6741 pixels and an F-measure of 98.69%.

Future work will focus on improving the robustness of proposed method, particularly by applying it to the extraction of different types of glacial lakes, such as supra-glacial lakes, which experience dramatic spatial-temporal changes and are more sensitive to climate change. The method may also be tested in other glaciated regions and improved by using high resolution of DEM data.

Acknowledgments: This work was supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (XDA19030101), the National Key R&D Program of China (2017YFE0100800) and the National Natural Science Foundation of China (41701481).

Author Contributions: Meimei Zhang, Fang Chen and Bangsen Tian contributed the main idea and designed the experiments; Meimei Zhang performed the experiments and analyzed the remote sensing data; Meimei Zhang wrote the manuscript, which was then improved by the contribution of all the co-authors.

Conflicts of Interest: The authors declare no conflict of interest.

References

23. Xu, H. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. Int. J. Remote Sens. 2006, 27, 3025–3033. [CrossRef]
34. Gribenski, N.; Lukas, S.; Jansson, K.N.; Stroeven, A.P.; Preusser, F. Complex patterns of glacier advances during the late glacial in the Chagan Uzun Valley, Russian Altai. *Quat. Sci. Rev.* 2016, 149, 288–305. [CrossRef]


