

Article

An Automatic Accurate High-Resolution Satellite Image Retrieval Method

Zhiwei Fan ^{1,†}, Wen Zhang ^{1,†}, Dongying Zhang ² and Lingkui Meng ^{1,*} 

¹ School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China; fanzhiwei@whu.edu.cn (Z.F.), wen_zhang@whu.edu.cn (W.Z.)

² School of Water Conservancy & Environment, Zhengzhou University, Zhengzhou 450001, China; zhangdongying@zzu.edu.cn

* Correspondence: lkmeng@whu.edu.cn; Tel.: +86-159-2706-6565

† These authors contributed equally to the work.

Academic Editor: Qi Wang

Received: 9 September 2017; Accepted: 21 October 2017; Published: 26 October 2017

Abstract: With the growing number of high-resolution satellite images, the traditional image retrieval method has become a bottleneck in the massive application of high-resolution satellite images because of the low degree of automation. However, there are few studies on the automation of satellite image retrieval. This paper presents an automatic high-resolution satellite image accurate retrieval method based on effective coverage (EC) information, which is used to replace the artificial screening stage in traditional satellite image retrieval tasks. In this method, first, we use a convolutional neural network to extract the EC of each satellite image; then, we use an effective coverage grid set (ECGS) to represent the ECs of all satellite images in the library; finally, the satellite image accurate retrieval algorithm is proposed to complete the process of screening images. The performance evaluation of the method is implemented in three regions: Wuhan, Yanling, and Tangjiashan Lake. The large number of experiments shows that our proposed method can automatically retrieve high-resolution satellite images and significantly improve efficiency.

Keywords: high-resolution satellite image; automated retrieval method; convolutional neural network; geohash coding; satellite image retrieval; satellite image screening

1. Introduction

The information in satellite images plays an important role in environmental monitoring, disaster forecasting, geological surveying, and other applications. With the steadily expanding demand for remotely sensed images, many satellites have been launched, and thousands of images are acquired every day. Increasingly, many researchers and organizations use satellite images to study the surface evolution process of a Region of Interest (ROI), and the entrance of the remote sensing application is satellite image retrieval.

Conventional satellite image retrieval in remote sensing applications is performed using a compound search composed of spatial search and attribute search. The condition of the spatial search is a defined geographical area commonly referred to as the Region of Interest (ROI), and the condition of attribute search generally refers to the imaging time range, spatial resolution, imaging platform, etc. For example, we need some high-resolution satellite images as first-hand information for a land use survey of Wuhan in 2016; the ROI is Wuhan in the retrieval task, and the conditions of the attribute search are imaging in 2016 and a spatial resolution above 10 m per pixel. Some well-known satellite imagery portals use this form of retrieval [1–3].

However, the image set searched by the composite retrieval method cannot be directly applied in most applications; a screening process is required to finish the retrieval task. The purpose of the

screening is to select images that can be directly used for subsequent processing. The screening is performed according to the following guidelines: (a) The entire ROI should be effectively covered, i.e., there should be at least one image that covers the area, and the content of the corresponding part of the image should be the earth's surface instead of cloud; and (b) The number of images in the collection should be as small as possible because fewer images correspond to lower processing costs.

The screening process of the conventional satellite image retrieval is currently manually performed, which restricts the efficiency of applications using satellite images. We describe the image screening as a process of selecting a suitable satellite image set "B" from the image set "A" queried by the composite retrieval method. We call image set A the "Pre-selected Image Set" (PIS) and image set B the "Selected Image Set" (SIS). To make these concepts more intuitive, an example is shown below.

As shown in Figure 1, there are eight satellite images that cover the ROI, which makes $PIS = \{1, 2, 3, 4, 5, 6, 7, 8\}$, and each image in PIS has a different effective coverage. Figure 2 shows three image sets screened out from PIS: image sets I, II, and III. There are two images in I, and the union of the two images can completely cover the ROI. However, a part of the ROI is covered by cloud; thus, image set I does not satisfy screening criteria (a). Both image sets II and III satisfy screening criteria (a); II has five images, and III has three. Thus, according to screening criteria (b), III is selected as the final SIS.

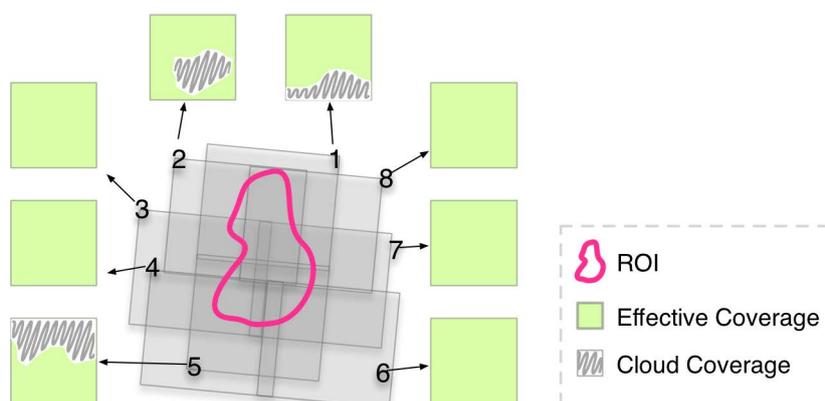


Figure 1. Pre-selected Image Set (PIS) and effective coverage of each satellite image in PIS.

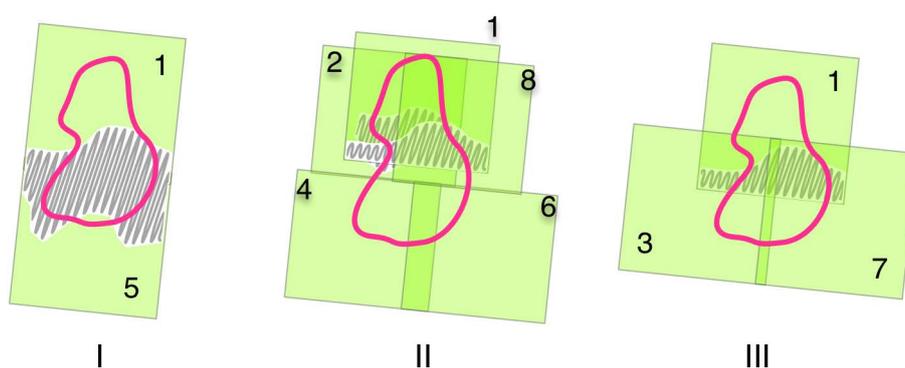


Figure 2. Several image sets screened from PIS.

Although there has been no automatic solution for this problem, some studies are helpful in solving this problem, including cloud detection methods and the cloud distribution expression of remote-sensing images. Cloud detection technology supports the automatic extraction of effective coverage information of satellite images. The reasonable expression of the cloud distribution of satellite images is the prerequisite for realizing automatic screening.

Cloud detection is widely used in remote-sensing applications. In recent years, new cloud detection methods have continued to emerge. Hagolle O, Huc M et al. (2010) developed a multi-temporal cloud detection method for FORMOSAT-2 and LANDSAT images [4]. Zhu Z and Woodcock C. E (2012) proposed an Fmask cloud detection method by combining with Landsat Top of Atmosphere (TOA) reflectance and Brightness Temperature (BT) [5]. Laban N, Nasr A et al. (2012) developed the multi-scale cloud extraction of remote-sensing images using spatial and texture features [6]. Surya S. R and Simon P (2013) used color space transform and Fuzzy C-means clustering to extract cloud in Landsat images [7]. Goodwin N. R, Collett L. J et al. (2013) proposed a fast cloud detection algorithm for the Landsat images using the hierarchical processing and combining the spectral information of the multi-temporal images [8]. Fisher A (2014) developed a cloud detection method specifically for SPOT5 HRG images [9]. Han Y, Kim B et al. (2014) proposed a cloud detection algorithm based on a reference image for high-resolution remote-sensing images [10]. Başeski E and Cenaras Ç (2015) use color and texture information to detect cloud in remote-sensing images [11]. Zhu Z, Wang S et al. (2015) improved the Fmask algorithm for cloud detection of Landsats 4–8 images [12]. An Z and Shi Z (2015) proposed a new automatic supervised approach based on the “scene-learning” scheme for cloud detection on remote-sensing images [13]. Wu T, Hu X et al. (2016) introduced the stereoscopic vision frame to solve the automatic cloud detection problem using DSM and DEM data [14].

In recent years, some studies have discussed the cloud distribution expression of remote-sensing images. Laban N, Nasr A et al. (2012) developed a spatial cloud detection and retrieval system (SCDRS) to retrieve the cloud distribution of remote-sensing images, and the cloud distribution information is expressed using tiling grids in their system [6]. Feng A and Shu S (2014) proposed a model of index of cloud in images based on GeoSOT, which is cloud distribution information organized by a global discrete grid system and saved in a particular file [15].

Obviously, the essential factor that affects the screening process is the effective coverage (EC) of the images in PIS. If the EC of satellite images can be extracted, organized, and rightfully used, the screening task can achieve automation. Based on this idea, an automatic accurate high-resolution satellite image retrieval (AA-HRSIR) method is presented in this study. The AA-HRSIR method consists of three parts: (1) automatic extraction of EC; (2) the effective coverage grid set (ECCGS); and (3) the satellite image accurate retrieval algorithm.

The remainder of the article is organized as follows. In Section 2, details of the AA-HRSIR method are provided. In Section 3, experiments are presented to demonstrate the results of the AA-HRSIR method. In Section 4, the effectiveness of the AA-HRSIR method is discussed. Finally, conclusions are presented in Section 5.

2. Methods

2.1. Effective Coverage Extraction

In a high-resolution satellite image, the effective coverage (EC) is the geographical area that corresponds to the cloudless part of the image. EC can be obtained when cloud in the satellite image is detected.

This paper uses satellite false-color preview images as the input data for cloud detection because false-color preview images are used in the manual screening stage of actual retrieval tasks, which indicates that the preview images carry adequate information for screening, and the false-color preview images have much lower processing costs than the raw data of the satellite images. However, almost all existing cloud detection techniques use the original data of the remote-sensing image as the input and are designed for specific remote-sensing images, so they cannot be used in our study.

We use a convolutional neural network (CNN) to perform cloud detection in satellite images because cloud detection is essentially a problem of single-label image classification, and convolutional neural networks have been proven to solve it with good performance [16]. One key element that

affects the accuracy of the convolution neural network model is the number of samples; in general, more samples correspond to a higher accuracy of the CNN. Fortunately, the number of satellite remote-sensing images is large, which provides sufficient data for the CNN.

The process of cloud detection includes (1) satellite image preprocessing and (2) cloud detection using a convolutional neural network.

2.1.1. Satellite Image Preprocessing

First, we split the satellite image into small blocks, and the size of the blocks determines the resolution of the cloud detection results. A smaller image block yields a higher resolution of the cloud detection results, whereas larger image blocks are more advantageous for the CNN to extract features from it. To determine the size of the division, we have referred to some datasets (MINST, CIFAR-10, CIFAR-100, SVHN, STL-10) which are commonly used in the field of CNN image classification. The image size of these datasets includes 28×28 , 32×32 and 96×96 , we chose the smallest one to obtain higher resolution of the cloud detection results.

The preprocessing includes two operations: resizing and splitting the satellite image.

Step 1 Resizing the Image

The satellite image is resized to multiples of 28×28 . Suppose that the width and height of the image are W and H ; then, the width and height of the resized image should be $\lfloor W/28 \rfloor \times 28$ and $\lfloor H/28 \rfloor \times 28$.

Step 2 Splitting the Image

After the resizing is completed, all resized satellite images will be split into $\lfloor W/28 \rfloor \times \lfloor H/28 \rfloor$ image blocks.

2.1.2. Cloud Detection Using a Convolutional Neural Network

After the preprocessing, each satellite image was converted into a 2-D array of small image blocks. Next, we use a CNN model to classify the image blocks into “cloudy” or “cloudless”. The classification process (Figure 3) comprises three stages: (1) collecting cloud and cloudless samples; (2) building the CNN model and training it using the cloud and cloudless samples; and (3) extracting the EC of all satellite images and storing these values in the database.

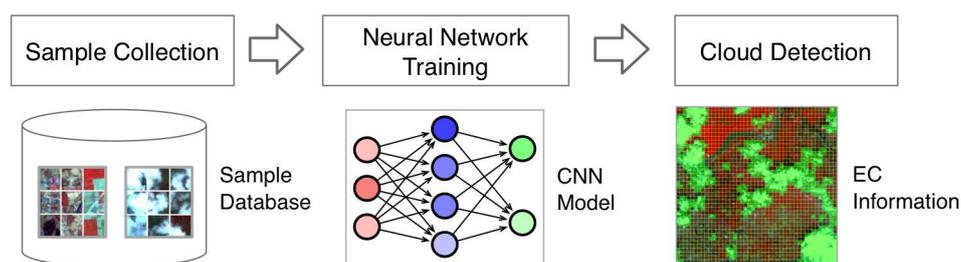


Figure 3. Cloud detection process.

Step 1 Collecting Cloud and Cloudless Samples

Because the samples largely determine the effect of the CNN model, we select many images to ensure the accuracy of the model. The samples were manually collected and stored in the sample database. By visual inspection, cloud and cloudless image blocks are selected as samples and imported into the sample database. In total, 138,690 samples were collected, including 73,602 cloudless samples and 65,088 cloud samples.

As shown in Figure 4, some samples are collected from a satellite image; the green blocks are cloud samples, and the blue blocks are cloudless samples. Each sample can be uniquely represented

by the image id and its relative position in the image; the data structure of the sample in the database is as follows (Table 1):

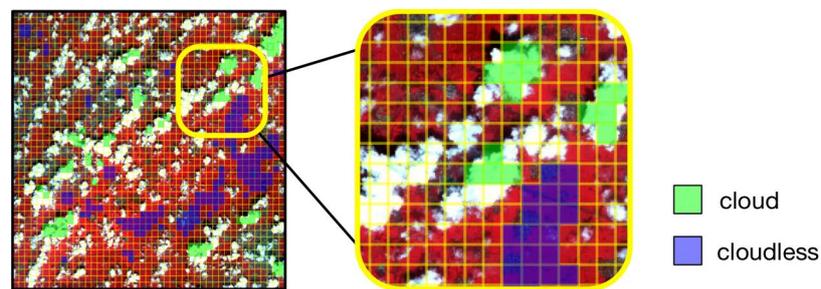


Figure 4. Collect cloud and cloudless samples from a satellite image.

Table 1. Sample data structure.

Field Name	Field Type	Example
imageid	String	"wt482nywt49ysh-20150808-gf2-pms1"
xcord	Integer	21
ycord	Integer	36
label	Integer	1 (cloudless) or 0 (cloud)

Step 2 Building and Training a CNN Model

We use TensorFlow [17] to build the CNN model. The classification of the image is in color, and there are fewer categories (two categories: cloud and cloudless) to distinguish, during the commonly used datasets. CIFAR-10 [18] is most similar to the input data in this study because its images are in color and have a smaller size (32×32), and its categories are few (ten categories); therefore, we refer to a neural network model [19] that performs well on the CIFAR-10 dataset and build the cloud detection convolution neural network model. The structure of the CNN model here is shown in Figure 5; four layers must be trained in the CNN model (not including the pooling layers): the first two layers are convolutional layers, and the last two layers are fully connected layers. The last layer of the CNN is a softmax layer and is used to complete the classification of two categories: cloud and cloudless.

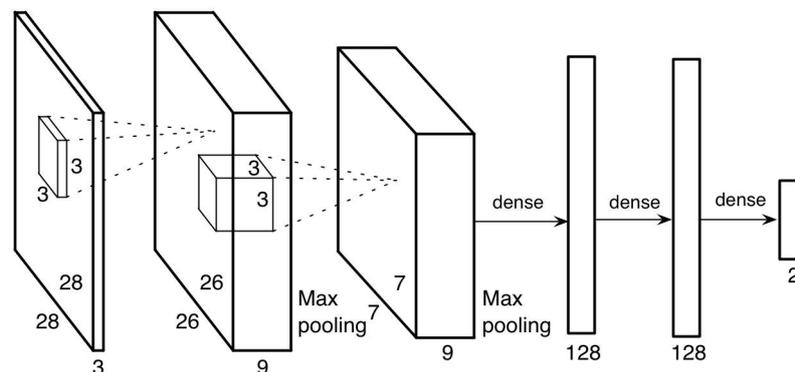


Figure 5. Structure of the convolutional neural network (CNN) model in this study.

After the CNN model is built, cloud and cloudless samples are used to train it (Figure 6). We use TensorFlow to run the training process, and the accuracy changes are shown in Figure 7.

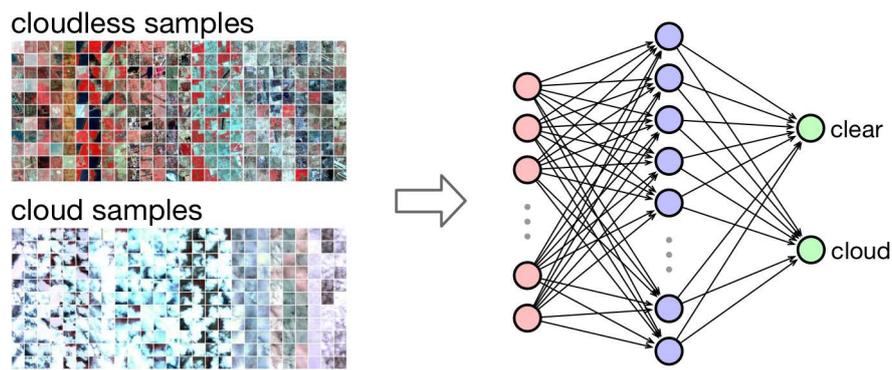


Figure 6. Training the neural network using samples.

As shown in Figure 7, in the first 500 iterations, the accuracy of the cloud detection CNN model quickly increases with the training process and becomes stable after 1000 iterations. After 10,000 iterations of training, the final accuracy of the model is stable at approximately 97%.

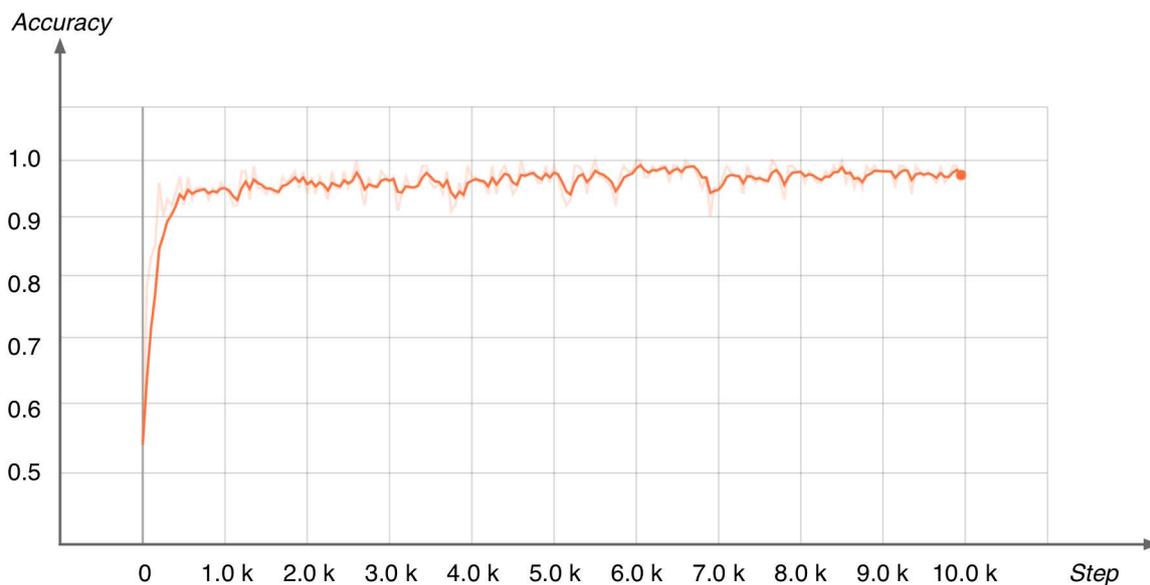


Figure 7. Accuracy curve during the CNN model training.

Step 3 Extracting the Effective Coverage

The final process of the effective coverage extraction comprises three steps (Figure 8): (1) According to the method in Section 2.1.1, each satellite image is resized and split into image blocks with a length and a width of 28; (2) The image blocks are classified by the trained CNN model into “cloud” and “cloudless”; and (3) The effective coverage of each satellite image is represented by a matrix, which is the effective coverage matrix (ECM), and saved into the satellite image database.

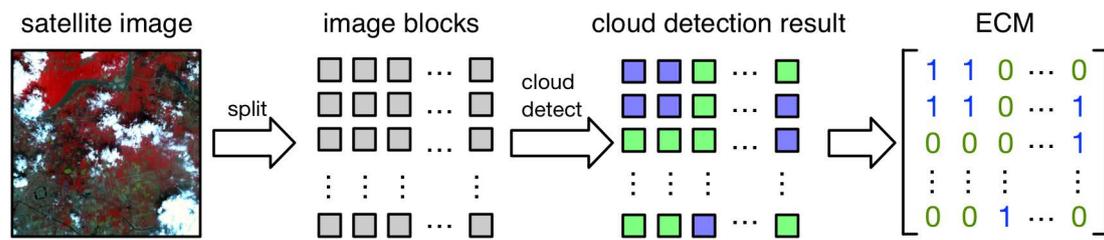


Figure 8. Effective coverage (EC) extraction process.

2.2. Effective Coverage Grid Set

The use of ECM is a simple method to store the effective coverage of each satellite image; however, the matrix does not contain geographic information, so the ECM cannot be directly applied to the satellite image screening process. We use geohash [20] coding to solve this problem. Geohash is a hierarchical spatial data structure that divides the ground surface into discrete grids, and each grid is labeled by a unique character string. It is used here to convert ECM to a new data structure, the effective coverage grid set (ECGS):

$$ECGS = \text{map}(\text{gridCode}, \text{overlapRatio}) \tag{1}$$

where gridCode is a geohash string, overlapRatio is a real number from 0 to 1, which represents the ratio of effective coverage, and map(x, y) means that ECGS is a collection of key-value pairs; in ECGS, the key is gridCode, and the value is overlapRatio. For example:

$$ECGS_{e.g} = \{ "wt4yd" : 0.36, "wt4vw" : 0.62, "wt4y1" : 1.00, \dots, "wt4yq" : 0.75 \}$$

The conversion process of ECM to ECGS is as follows:

In Figure 9, overlay refers to the geographical coverage of satellite imagery, and ECP refers to the effective coverage polygon. The conversion process of ECM to ECGS includes two steps: conversion from ECM to ECP and conversion from ECP to ECGS.

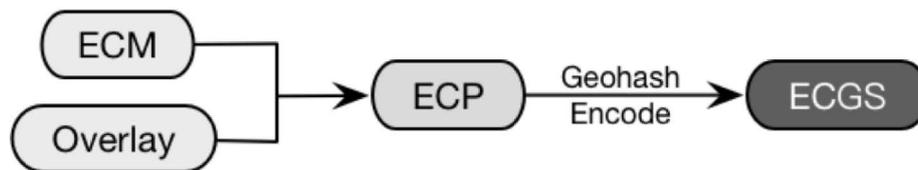


Figure 9. Conversion process of effective coverage matrix (ECM) to effective coverage grid set (ECGS).

2.2.1. Conversion from ECM to ECP

We assume that the number of rows of ECM is n_l and the number of columns of ECM is n_s . As shown in Figure 10, each element in ECM can be uniquely identified with (i, j) , and it is represented as a geographic quadrilateral in the geographic coordinate system.

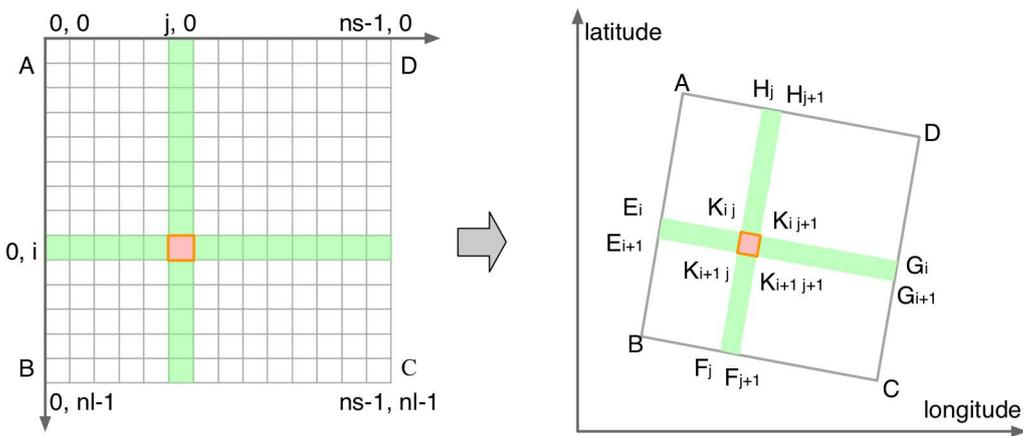


Figure 10. Convert elements of ECM to geographical polygons.

The conversion from one element of ECM (ECM_{ij}) to a geographical quadrilateral ($Cell_{ij}$) can be performed as follows:

$$\begin{cases} \frac{AH_j}{AD} = \frac{BF_j}{BC} = \frac{j}{ns} \\ \frac{AE_i}{AB} = \frac{DG_i}{DC} = \frac{i}{nl} \end{cases} \quad (2)$$

$$K_{ij} = \text{Intersection}(E_i G_i, F_j H_j) \quad (3)$$

$$Cell_{ij} = \text{Polygon}(K_{ij}, K_{i+1j}, K_{i+1j+1}, K_{ij+1}); i \in [0, nl - 2], j \in [0, ns - 2] \quad (4)$$

ECP is the union of the cells with values of 1 in ECM; therefore, ECP is expressed by

$$ECP = \text{Union}\{Cell_{ij} | ECM_{ij} = 1\} \quad (5)$$

2.2.2. Conversion from ECP to ECGS

The spatial relations between ECP and the geohash grid are shown in Figure 11; ECP is divided into small regular blocks by the geohash grids. The conversion process from ECP to ECGS is as follows.

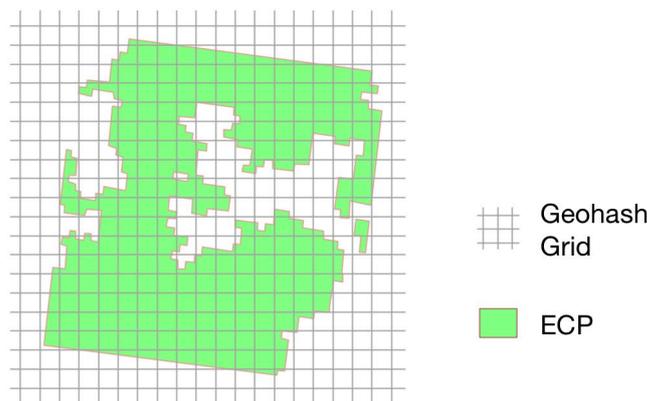


Figure 11. ECP and geohash grid.

Step 1 Determining the Precision of Geohash

The precision of geohash determines the accuracy of the effective coverage of the image (a higher precision implies a more accurate EC representation) and affects computational efficiency (a higher precision also implies greater demand for storage and lower computational efficiency).

Therefore, a moderate precision must be determined to balance efficiency and accuracy. The precision of geohash is determined to be 5. Based on this precision, the EC of the satellite images in this study is divided into more than 40 and less than 100 geohash grids, which satisfy the requirements of the follow-up accurate retrieval algorithm.

Step 2 Encoding ECP Using Geohash

As shown in Figure 12, we use the envelop box of ECP to obtain the potential grids (PGS) that may intersect with ECP, and the potential grids are marked in yellow. The calculation method of PGS is as follows:

$$\left\{ \begin{array}{l} y_i = A.y + i * \text{GridHeight}, i \in N \\ x_j = A.x + j * \text{GridWidth}, j \in N \\ r = \left\lceil \frac{A.y - B.y}{\text{GridHeight}} \right\rceil \\ c = \left\lceil \frac{D.x - A.x}{\text{GridWidth}} \right\rceil \\ \text{PGS} = \{ \text{Encode}(y_i, x_j) \mid i \in [0, r], j \in [0, c] \} \end{array} \right. \quad (6)$$

After acquiring PGS, we can obtain ECGS by calculating the overlap ratios between grids in PGS and ECP. Figure 13 shows the conversion process from ECP to ECGS.

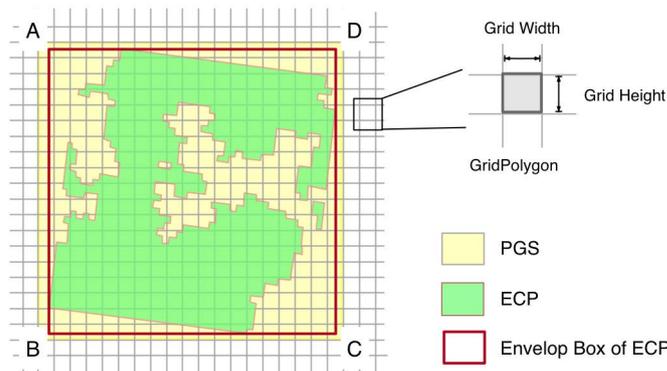


Figure 12. Potential grids that intersect with the envelope box of the effective coverage area.

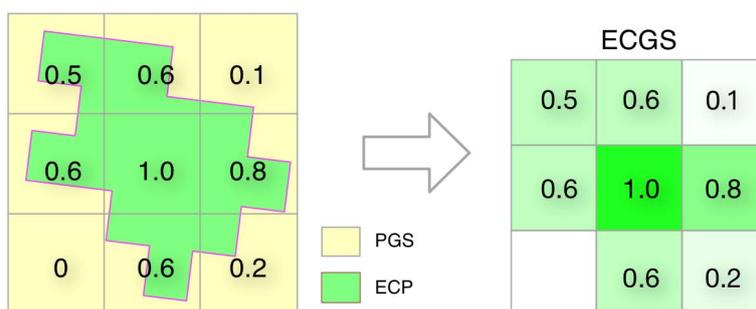


Figure 13. Convert the effective coverage area to ECGS.

ECGS is calculated by

$$\left\{ \begin{array}{l} \text{overlapRatio}_i = \frac{\text{Area}(\text{Intersection}(\text{ECP}, \text{Grid}_{\text{PGS}_i}))}{\text{Area}(\text{Grid}_{\text{PGS}_i})} \\ \text{ECGS} = \text{map}(\text{PGS}_i, \text{overlapRatio}_i) \mid \text{overlapRatio}_i > 0 \end{array} \right. \quad (7)$$

where PGS_i denotes the i -th element of PGS; $Grid_{PGS_i}$ denotes the geographical polygon represented by PGS_i ; and $overlapRatio_i$ denotes the overlap ratio between ECP and $Grid_{PGS_i}$.

2.3. Accurate Retrieval Algorithm

The final step of AA-HRSIR is to screen the images using their ECGS. To complete the screening, a satellite image accurate retrieval (SIAR) algorithm that simulates the manual-screening process is developed to screen satellite images from the PIS (Figure 14).

In the summary of the artificial screening process, we have found several rules that guide the screening process: (1) Images that cover the edge of the ROI are preferred to avoid small pieces scattered in an uncovered area; (2) Images with a large effective coverage area for the ROI are preferred to control the number of screened images; and (3) Images that properly overlap with other screened images are preferred to facilitate the image mosaic work. We apply these manual-screening rules to the SIAR algorithm as follows.

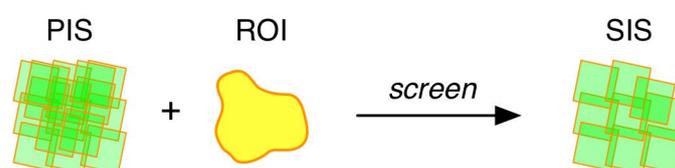


Figure 14. Process of screening images from the “Pre-selected Image Set” PIS.

2.3.1. Encoding ROI Using Geohash

To map the ECGS and ROI to a unified calculation system, we convert the ROI to the target grid set (TGS) using a similar method to the conversion from ECP to ECGS in Section 2.1.1; the only difference is that the value of each grid in the TGS is determined by whether the grid intersects the boundary of the ROI. As shown in Figure 15, there are two types of grids in the TGS: the grids on the edge area of the TGS are filled with a dark red color, which indicates that the weight of the grids is 1.0; the grids in the center area of the TGS are filled with a light red color, which indicates that the weight of the grids is 0.5.

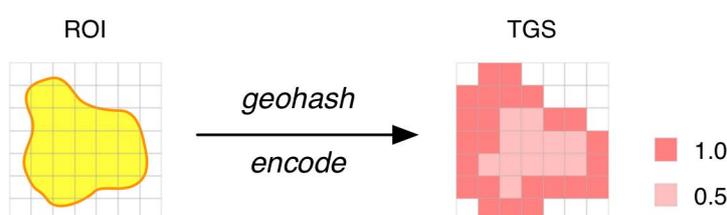


Figure 15. Encode Region of Interest (ROI) to target grid set (TGS).

We assign different weights to the grids in TGS, so that the images cover the edge of ROI are more likely selected, which is consistent with manual-screening rule 1.

2.3.2. Rating Images in the PIS

Because each image in the PIS has an intersection with the ROI, but only one image can be selected for each round of screening, we develop a rating method to evaluate the opportunity to be selected of each image in PIS, so that the images with larger effective coverage areas for the ROI are preferred, which is consistent with manual-screening rule 2.

We calculate the score of image in the PIS as follows:

$$\text{Score} = \sum \text{ECGS}_k * \text{TGS}_k, k \in \text{keys}(\text{ECGS}) \cap \text{keys}(\text{TGS}) \quad (8)$$

where ECGS_k is the value of the element in ECGS with key k ; TGS_k is the value of the element in the TGS with key k ; and the function “keys()” returns the keys of the elements in the collection, e.g.,:

$$\text{keys}(\text{ECGS}_{e.g.}) = \{“wt4yd”, “wt4vw”, “wt4y1”, \dots, “wt4yq”\}$$

2.3.3. Updating Target Grid Set

After a round of screening, the area of ROI that has not been covered by the images changes, which makes the TGS no longer suitable for the next round of screening. We update the TGS using the following steps:

Step 1 Calculating the Remaining Target Region

As shown in Figure 16, we calculate the remaining target region (RTR) as follows

$$\text{RTR} = \text{Clip}(\text{ROI}, \text{Union}(\text{SIS})) \quad (9)$$

where $\text{Union}(\text{SIS})$ is the geographical polygon that is effectively covered by images of the SIS.

Step 2 Encoding the Remaining ROI Using Geohash

After acquiring the RTR, we use geohash to encode the remaining target region (RTR) to RTGS (Figure 17) using the method in Section 2.3.1. We observe that the grids that intersect with the images in the SIS have larger weights (1.0); hence, the images that properly overlap with other screened images are preferred, which is consistent with manual-screening rule 3.

Figure 18 represents the change of PIS, SIS, and TGS during the SIAR process. The distribution of satellite images in PIS is shown in the first row. In the middle row of the figure, the yellow area denotes the ROI, and the green quadrilaterals denote the satellite images in SIS. The last row shows the change of TGS.

By using the methods mentioned in Sections 2.3.1–2.3.3, we can screen images from the “Pre-selected Image Set” using our proposed Satellite Image Accurate Retrieval (SIAR) Algorithm, which is elaborately described in Algorithm 1.

Algorithm 1. Satellite Image Accurate Retrieval (SIAR) (Figure 19)

Input: ROI: represented using a geographical polygon; PIS: $\text{PIS} = \text{map}(\text{imageid}, (\text{overlay}, \text{ECGS}))$, satellite image records searched using the conventional satellite image retrieval method;

1. Initialize SIS to an empty set and encode the ROI to the TGS using the method in Section 2.3.1.
2. Rate each image in PIS using the method in Section 2.3.2.
3. If every image’s score is 0, the algorithm is aborted and return SIS; otherwise, move the highest scoring image from PIS to SIS and update TGS using method in Section 2.3.3.
4. If TGS or PIS is empty, the algorithm is aborted and return SIS; otherwise, jump to step 2.

Output: the selected images (SIS).

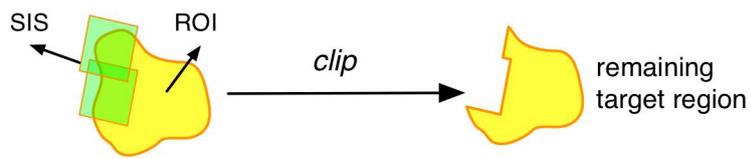


Figure 16. Obtaining the effective coverage of the “Selected Image Set” (SIS) through spatial union.

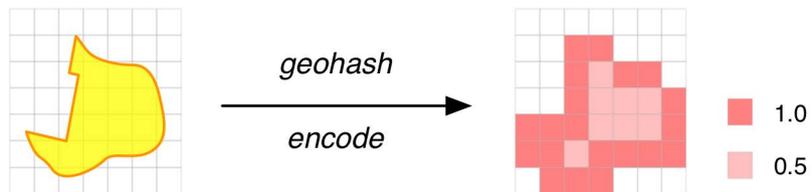


Figure 17. Encoding the remaining target region to the remaining target grid set (RTGS).

PIS				...		
SIS				...		
TGS				...		

Figure 18. Change of PIS, SIS and TGS during the SIAR process.

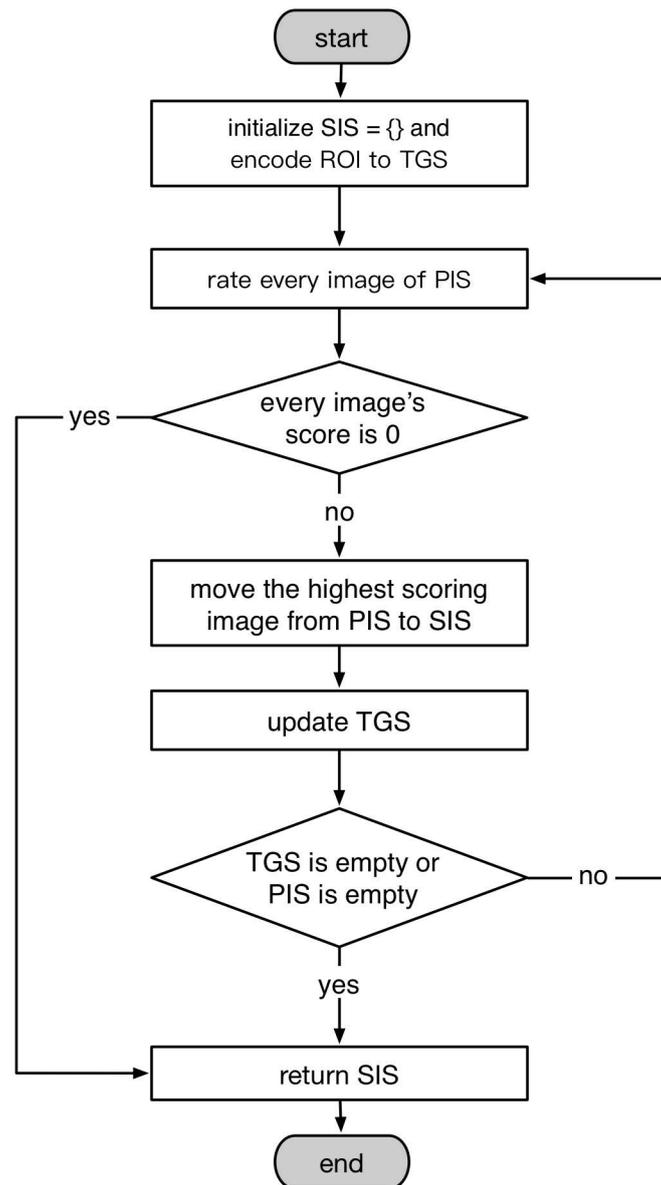


Figure 19. Flow diagram of the satellite image accurate retrieval (SIAR) algorithm.

3. Results

3.1. Cloud Detection Results

The clouds in the satellite image have many types of patterns; we classified the satellite images into five categories according to the ratio and type of cloud coverage: (a) cloud-free images: the cloud covers less than 5%; (b) less-cloud images: the cloud covers 5–30%; (c) partly cloud mages: the cloud covers 30–85%; (d) full-cloud images: the cloud covers more than 85%; (e) thin-cloud images: the images are covered by thin cloud, and the ground is vaguely visible through it. The image blocks covered by thin clouds are classified as ineffective coverage in this study.

3.1.1. Cloud Detection Test Data

The cloud detection test dataset (CDTD) was selected from the GF-1/2 satellite image database, which included 100 images. It contained the aforementioned types of images, some of which are shown in Figure 20.

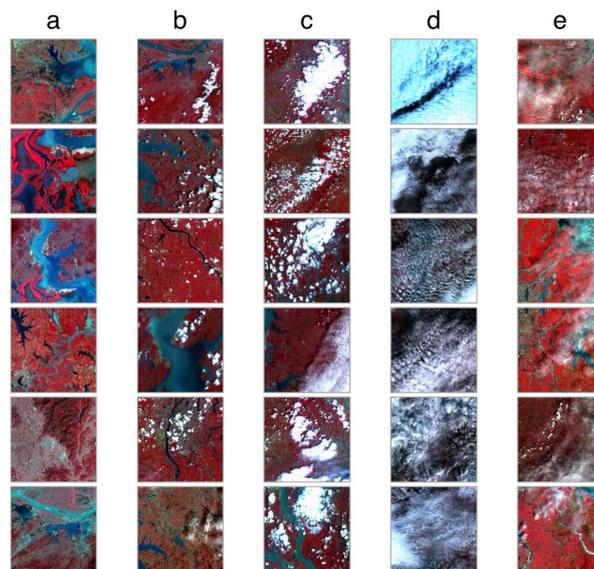


Figure 20. Some satellite images of the cloud detection test dataset (CTD).

3.1.2. Cloud Detection Test Methods

We used visual interpretation to obtain the ECM of each image in the CTD as the criterion. The specific approach is as follows:

1. Resize the test image and split it into image blocks according to Section 2.1.1.
2. Mark all image blocks covered by cloud as “cloudy”.
3. Save the manually marked result as ECM_0 according to Section 2.1.2.

After the visual interpretation process, all images in the CTD obtained the correct ECM: ECM_0 . Then, we use the trained CNN to obtain the EC of each image in the CTD; its result is recorded as ECM_t . The accuracy of cloud detection is calculated by

$$\text{Accuracy} = \frac{\sum(ECM_0 \odot ECM_t)}{\text{Size}(ECM_0)} \quad (8)$$

3.1.3. Cloud Detection Test Results

In accordance with the above method, we completed the test of the CNN cloud detection model on the CTD. The average accuracy is 96.84%, and the frequency histogram is shown in Figure 21. The pattern of the histogram is skewed right; the model accuracy on most CTD exceeds 95 and is 80–90% for few test images.

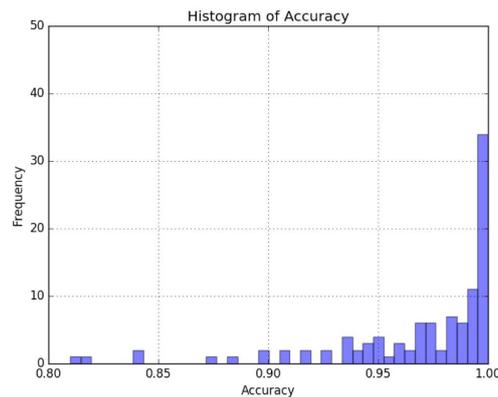


Figure 21. Histogram of the cloud detection accuracy.

In addition, we evaluated the performance of the cloud detection model on five types of images. As shown in Table 2, the average classification accuracies of the CNN model are more than 94% on type-a, -b, -c, and -d images but slightly lower on type-e images (only 88.69%).

Table 2. Cloud detection model performance on five types of satellite images.

	a	b	c	d	e
Accuracy	99.40%	97.25%	94.35%	96.28%	88.69%

3.2. Accurate Retrieval Results

3.2.1. Accurate Retrieval Evaluation Data

We arranged three experiments to show the performance of the algorithm in the satellite image retrieval tasks of regions with different sizes. In this study, we classify the ROI into three types according to the number of images required to fully cover the target area: large-area regions, medium-area regions, and small-area regions.

- Test regions (ROIs)

A large-area region denotes a region that requires at least 10 satellite images to fully cover. The number of images to be screened is large in the satellite image retrieval of the large-area region, which results in a substantial manual-screening workload. This type of satellite retrieval is generally used in mapping large-area regions. We selected Wuhan as the test large-area region (bottom right panel of Figure 22).

A medium-area region denotes a region that requires 2~10 satellite images to fully cover. Medium-area satellite image retrieval is generally used in regular mapping tasks in medium regions such as a county in central China. We selected Yanling as the test medium-area region (upper right panel of Figure 22).

A small-area region denotes a region that one satellite image can fully cover. Small-area satellite image retrieval is often used in emergency monitoring of a local area, such as for landslides or a dammed lake caused by earthquakes. We selected Tangjiashan Lake as the test small-area region (upper left panel of Figure 22).

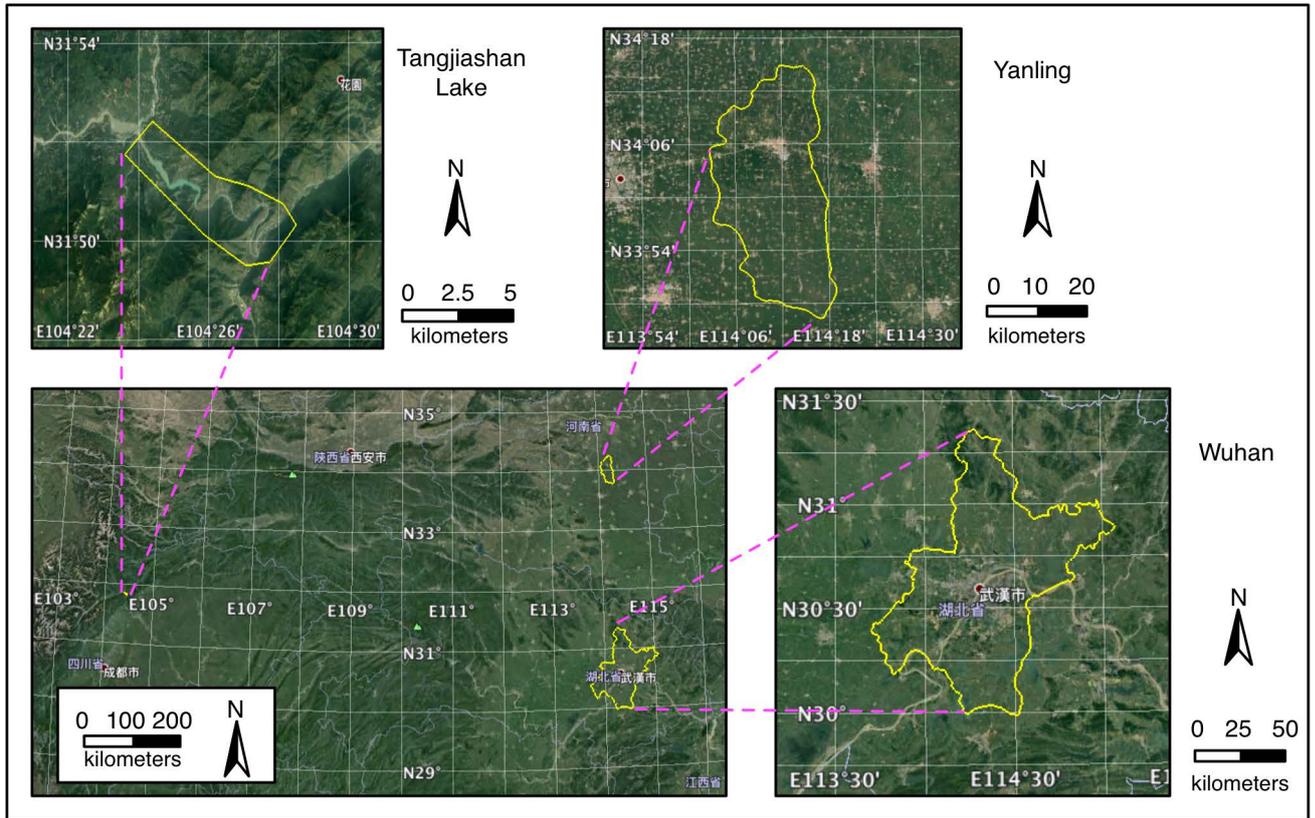


Figure 22. Three test ROIs: Wuhan, Yanling, and Tangjiashan Lake.

- Test satellite images

The test data sources are multispectral images with 8-m resolution of GF-1 satellite [21] and multispectral images with 4-m resolution of GF-2 satellite [22], which were taken in 2015 and 2016. The number of satellite images that cover these ROIs is shown in Table 3. Some examples of the test satellite images are shown in Figure 23. Some areas of the satellite images are covered by clouds.

Table 3. Statistics of the test satellite images.

ROI	Source	Time Span	Image Count
Wuhan	GF-1, GF-2	1 January 2015~31 December 2016	491
Yanling	GF-1, GF-2	1 January 2015~31 December 2016	122
Tangjiashan Lake	GF-1, GF-2	1 January 2015~31 December 2016	21

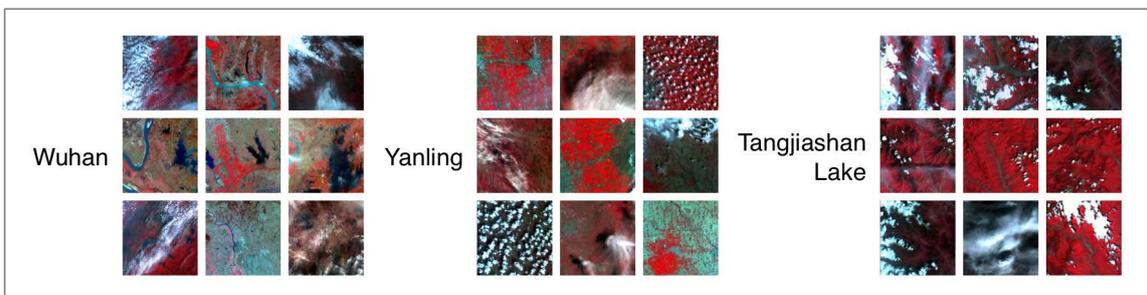


Figure 23. Satellite images in the test.

- Query conditions

We set up several different query conditions for each test ROI and executed them through spatial and attribute queries to obtain pre-selected image sets (PISs) that must be screened. The number of PISs is shown in Table 4.

Table 4. Retrieval conditions in the test.

ROI	Source	Period of Time	Number of PISs
A	Wuhan	GF-1 1 January 2015~31 December 2015	88
B		GF-1 1 January 2015~31 December 2016	117
C		GF-2 1 January 2015~31 December 2015	91
D		GF-2 1 January 2015~31 December 2016	195
E	Yanling	GF-1 1 January 2015~31 December 2015	35
F		GF-1 1 January 2015~31 December 2016	37
G		GF-2 1 January 2015~31 December 2015	27
H		GF-2 1 January 2015~31 December 2016	23
K	Tangjiashan Lake	GF-1 1 January 2015~31 December 2015	4
L		GF-1 1 January 2015~31 December 2016	7
M		GF-2 1 January 2015~31 December 2015	3
N		GF-2 1 January 2015~31 December 2016	7

3.2.2. Accurate Retrieval Evaluation Criteria

Because no other algorithm has solved the automatic satellite images screening problem, we use the results of manual screening as a standard to evaluate the performance of the SIAR algorithm.

- Effective coverage ratio

The effective coverage ratio (ECR) is used as an indicator to judge the screening results. The ECR is calculated by

$$ECR = \frac{\text{Area}(EC_{ROI})}{\text{Area}(ROI)} \quad (9)$$

where EC_{ROI} is the effective coverage to ROI, and its calculation process is as follows (as shown in Figure 24): first, we obtain the union of effective coverage polygons (ECPs, Section 2.2.1) of the SIS; then, we determine EC_{ROI} by obtaining the intersection of the union polygon and the ROI.

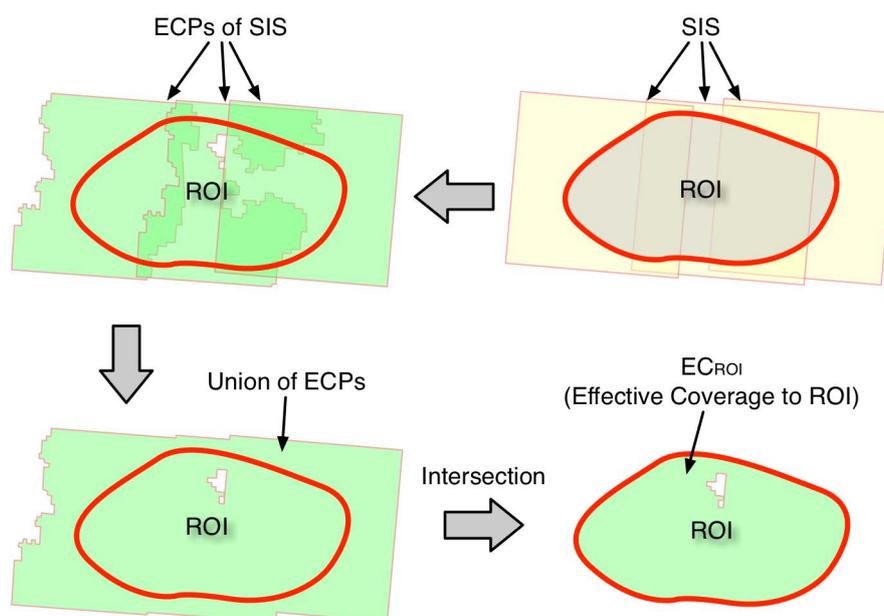


Figure 24. Calculation process of EC_{ROI} .

- Effective coverage frequency

To screen satellite images of medium and large areas, which require more than 2 images to fully cover, a qualified screening result should have a high effective coverage ratio and a suitable redundancy of coverage.

As shown in Figure 25, the SIS on the left has three satellite images, and the SIS on the right has two images; different parts of the ROI are covered with different frequencies. Obviously, SIS-I covers the ROI too frequently (most parts of the ROI is covered more than 2 times), whereas SIS-II covers the ROI with a suitable frequency.

We propose the approach of “effective coverage frequency grids” to calculate the effective coverage frequency (ECF) of each part of the ROI that is divided by the grids. As shown in Figure 26, each grid records the ECF, and different colors represent different coverage frequencies; we can easily find the spatial distribution of the ECF using this method.

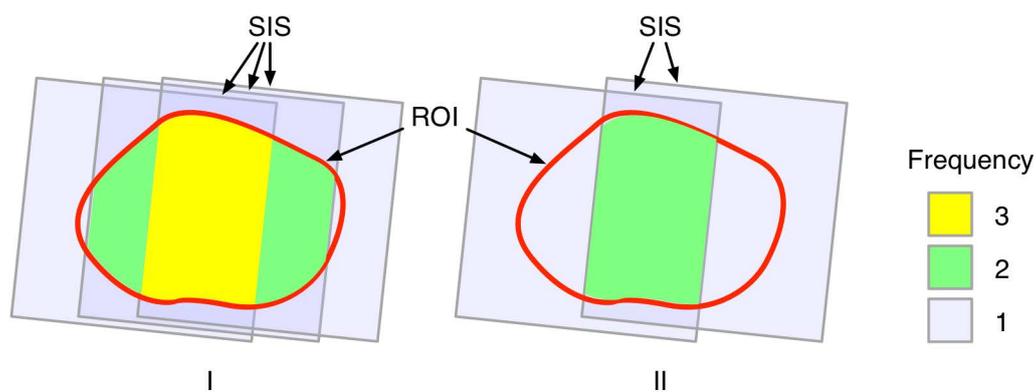


Figure 25. Two different screening results and their coverage frequency for the ROI.

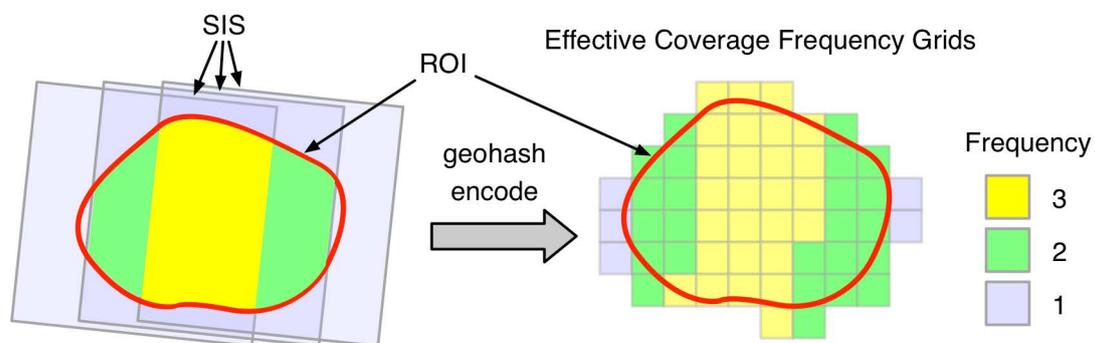


Figure 26. Conversion from SIS to effective coverage frequency grids.

Finally, a quantitative indicator, which is the average effective coverage frequency (\overline{ECF}), is used to assess the quality of the screening results. If the \overline{ECF} is greater than 2, the redundancy of the screening result is high. If the \overline{ECF} is less than 2 and greater than 1, the redundancy of the screening result is appropriate.

- Screening time consuming

Time consumption is a key indicator to assess the screening process. If the manual screening results and SIAR results have similar qualities, the method that uses less time to finish the task is undoubtedly better. We recorded the time cost of the screening processes to evaluate the efficiency improvement of the SIAR algorithm compared to manual screening.

3.2.3. Accurate Retrieval Evaluation Results

- Effective coverage ratio

We used formula 9 to calculate the ECRs of the manual and SIAR results, which are shown in Table 5. Table 5 shows that the difference in ECR of the manual and SIAR results is less than 2.5%.

The ECR of most screening results is very close to 1.0 (0.98~1.0), but the results of test cases C, G and N are not (less than 0.93), because the clouds in satellite images that match the query criteria of tests C, G and N always obscured part of the ground surface of the target region.

Table 5. Effective coverage ratios of the manual and SIAR results.

	Wuhan				Yanling				Tangjiashan Lake			
	A	B	C	D	E	F	G	H	K	L	M	N
Manual ECR	0.994	0.998	0.891	0.998	1.000	1.000	0.907	0.995	1.000	1.000	1.000	0.909
SIAR ECR	0.996	0.984	0.892	0.997	1.000	1.000	0.927	1.000	1.000	1.000	1.000	0.909

- Effective coverage frequency

The effective coverage frequencies of the manual and SIAR results are represented using the method in the above section (Figures 27 and 28). In Figures 27 and 28, the boundary of the ROI is marked by the red line; the grids are marked with different colors to denote different effective coverage frequencies: the light green grids indicate that the ECF is 1–2, the yellow ones indicate that the ECF is 2–4, the orange and red grids indicate that the ECF is 4–5, and the dark green and gray grids indicate that the ECF is less than 1.

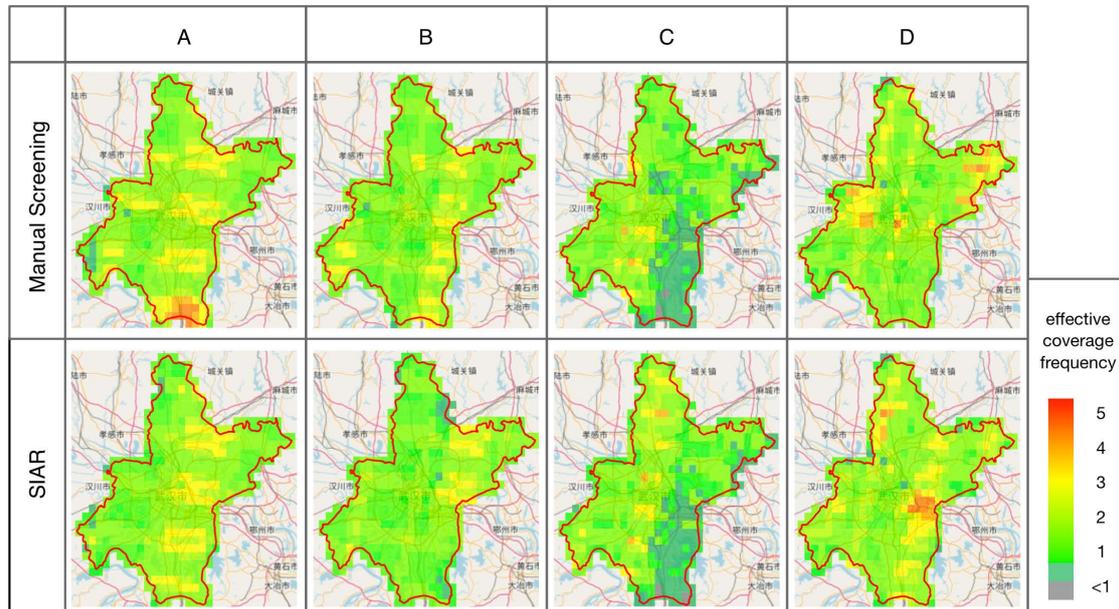


Figure 27. Effective coverage frequency grids of SISs of Wuhan test.

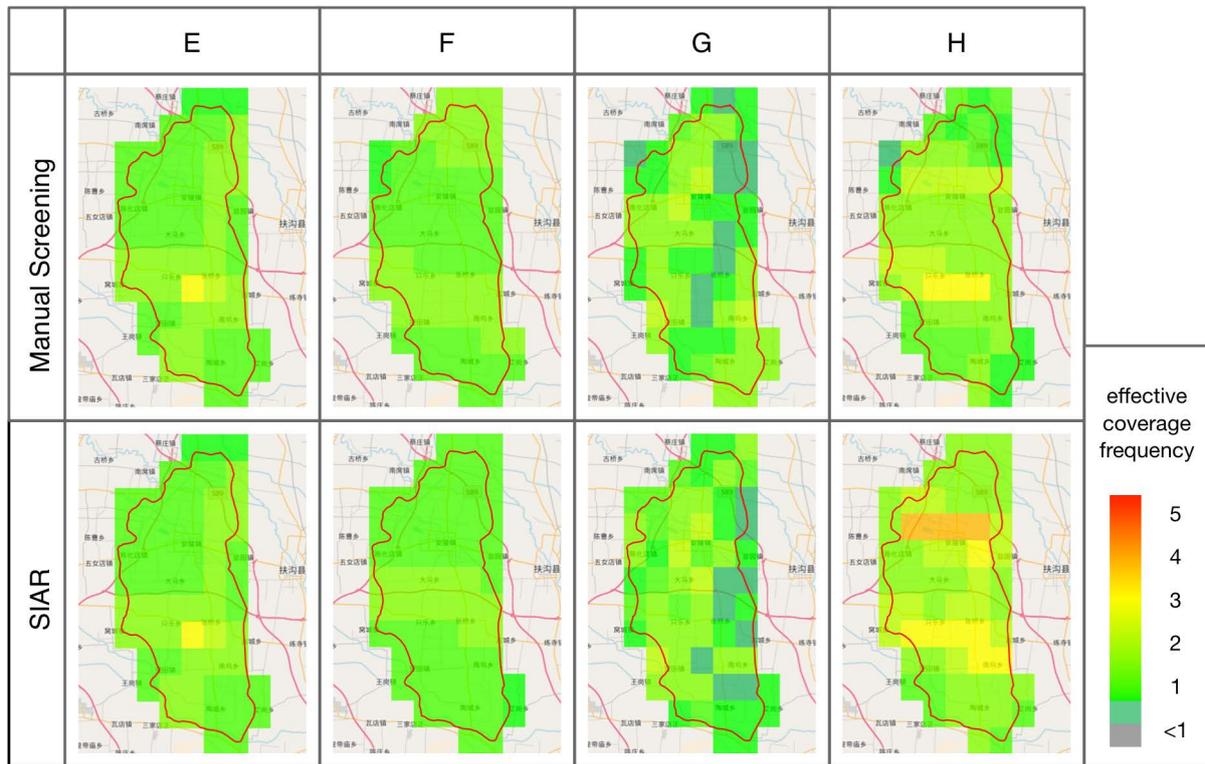


Figure 28. Effective coverage frequency grids of SISs of Yanling test.

Figures 27 and 28 show that most grids of the ROI are marked as light green in both manual screening and SIAR results, i.e., most parts of the ROI are effectively covered 1–2 times. The quantitative indicator \overline{ECF} of each result in the tests was calculated (Table 6). All \overline{ECF} s of the manual screening results are 1.0–2.0, and so are the SIAR results.

Table 6. Average effective coverage frequencies of manual screening and satellite image accurate retrieval (SIAR).

	Wuhan				Yanling			
	A	B	C	D	E	F	G	H
Manual \overline{ECF}	1.782	1.105	1.509	1.608	1.165	1.059	1.200	1.515
SIAR \overline{ECF}	1.616	1.183	1.361	1.690	1.165	1.026	1.096	1.929

- Screening time consuming

We recorded the time cost of manual screening and SIAR algorithm (Table 7). In the screening tasks of a large area (Wuhan), the manual screening costs 2–5 min, whereas the SIAR algorithm costs 4–13 s. In the screening tasks of a medium area (Yanling), the manual screening costs 15–36 s, whereas the SIAR algorithm costs approximately 0.1 s. In the screening tasks of a small area (Tangjiashan Lake), the manual screening costs 4–7 s, whereas the SIAR algorithm costs 0.01 s. The experimental test shows that the SIAR algorithm consumes much less time than the manual screening, and the SIAR algorithm is ten to hundreds of times more efficient than manual screening.

Table 7. Time cost of manual screening and SIAR.

	Wuhan				Yanling				Tangjiashan Lake			
	A	B	C	D	E	F	G	H	K	L	M	N
Manual (s)	167.34	186.99	277.40	169.38	15.42	25.87	50.41	35.07	5.38	4.49	6.39	6.83
SIAR (s)	4.57	4.00	9.87	12.67	0.11	0.07	0.14	0.16	0.01	0.01	0.01	0.01

4. Discussion

Extracting effective coverage from satellite images is a notably important first step to implement an accurate retrieval study and for application of remote sensor data. In this study, we use a convolutional neural network model to classify the image blocks that are divided from satellite images into “cloudy” and “cloud-free”. According to the results in Section 3.1, the cloud detection accuracy could exceed 95%, and the algorithm performs well on all types of satellite images (most of the accuracy is more than 90%). Accordingly, the CNN cloud detection model here can satisfy the requirement of accurate retrieval.

The management of the satellite image effective coverage information is critical to the accurate retrieval algorithm. ECGS is used in this study to organize the effective coverage information of satellite images. It has two main advantages: (1) The ECs of different satellite images are mapped into a unified coding system with geographical meaning, which makes the calculation between ECs of different satellite images convenient. (2) ECGS simplifies the representation of the effective coverage from a complex polygon to a linear set, which makes it easy to process in the accurate retrieval process.

The SIAR algorithm is the core of the automatic accurate satellite image retrieval method. It combines the ECGS with the rules of manual-screening approaches to automatically screen satellite images. To evaluate the performance of the SIAR algorithm in screening image tasks, we use the manual-screening results as the standard to determine the SIAR results and introduce three indicators to quantify such judgement: effective coverage ratio (ECR), effective coverage frequency (ECF) and screening time consumption. Three experiments cover different types (small-area, medium-area, and large-area) of satellite image retrieval and are arranged to verify the applicability of the SIAR algorithm. The experiments show that: (1) The ECRs of the SIAR algorithm slightly vary (within less than 2.5%); (2) The \overline{ECFs} of the satellite images screened by the SIAR algorithm and the manual-screening results are greater than 1 and lower than 2; and (3) The SIAR algorithm is ten to hundreds of times more efficient than the manual screening in satellite image screening tasks. The first two findings have proven that the satellite images screened by the SIAR algorithm have similar quality to manually screened images. The third finding of screening time reduction demonstrates that the SIAR algorithm can increase the gains by replacing the manual-screening work.

The proposed method has some limitations: (1) The method of extracting effective coverage information only applies to optical satellite images; and (2) The accuracy of the SIAR algorithm highly depends on the accuracy of the effective coverage information; however, a small part of cloud detection accuracy (less than 90%) is not sufficiently high, so the effective information is not sufficiently accurate for the satellite image screening process.

5. Conclusions

To achieve the automation of satellite image retrieval, this paper proposes an automatic accurate high-resolution satellite image retrieval (AA-HRSIR) approach using the effective coverage grid set (ECGS). This paper designs a CNN model to detect cloud satellite images, which can achieve good cloud detection performance on diverse satellite images. To convert the cloud detection result into effective coverage information that can be directly used in automatic accurate satellite image retrieval, this paper advocates geohash encoding to record the effective coverage of each satellite image, i.e., ECGS. We use an auto-acquired ECGS to design the satellite image accurate retrieval (SIAR) algorithm

and achieve the automation of satellite image retrieval. Many experiments show that the proposed AA-HRSIR can replace the manual screening of satellite images and significantly improve efficiency.

Retrieval is the entrance to satellite imagery applications, and the current retrieval work requires human intervention, which limits the efficiency of satellite imagery applications. The automation of retrieval is bound to improve the efficiency of satellite image retrieval and can promote continuous and periodic applications of satellite images. In future work, we will extend the proposed AA-HRSIR to more applications in the remote sensing community. For example, the proposed AA-HRSIR will be used to obtain an annual high-resolution satellite image album of one country or one region, automatically obtain historical satellite images of the affected areas after disasters, and automatically search satellite images of objects with large geographical spans (e.g., large rivers).

Acknowledgments: This work was supported by the National Key Research and Development Program of China (No. 2017YFC0405806) and the National High-Resolution Earth Observation System Projects (No. 08-Y30B07-9001-13/15).

Author Contributions: Z.F. provided the original idea for this study, conceived and designed the experiments. W.Z. analyzed the results and together they wrote the manuscript. D.Z. modified and polished the manuscript. The work was supervised by L.M., who contributed to all stages of the work.

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

References

1. China Centre for Resources Satellite Data and Application. Available online: <http://www.cresda.com> (accessed on 18 August 2017).
2. NASA Landsat Program. Available online: <https://landsat.gsfc.nasa.gov> (accessed on 18 August 2017).
3. AIRBUS. Available online: <http://www.intelligence-airbusds.com> (accessed on 18 August 2017).
4. Hagolle, O.; Huc, M.; Pascual, D.V.; Dedieu, G. A multi-temporal method for cloud detection, applied to FORMOSAT-2, VEN μ S, LANDSAT and SENTINEL-2 images. *Remote Sens. Environ.* **2010**, *114*, 1747–1755. [[CrossRef](#)]
5. Zhu, Z.; Woodcock, C.E. Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sens. Environ.* **2012**, *118*, 83–94. [[CrossRef](#)]
6. Laban, N.; Nasr, A.; ElSaban, M.; Onsi, H. Spatial Cloud Detection and Retrieval System for Satellite Images. *Int. J. Adv. Comput. Sci. Appl.* **2012**, *3*. [[CrossRef](#)]
7. Surya, S.R.; Simon, P. Automatic Cloud Detection Using Spectral Rationing and Fuzzy Clustering. In Proceedings of the 2013 2nd International Conference on Advanced Computing, Networking and Security (ADCONS), Mangalore, India, 15–17 December 2013; pp. 90–95.
8. Goodwin, N.R.; Collett, L.J.; Denham, R.J.; Flood, N.; Tindall, D. Cloud and cloud shadow screening across Queensland, Australia: An automated method for Landsat TM/ETM+ time series. *Remote Sens. Environ.* **2013**, *134*, 50–65. [[CrossRef](#)]
9. Fisher, A. Cloud and Cloud-Shadow Detection in SPOT5 HRG Imagery with Automated Morphological Feature Extraction. *Remote Sens.* **2014**, *6*, 776–800. [[CrossRef](#)]
10. Han, Y.; Kim, B.; Kim, Y.; Lee, W.H. Automatic cloud detection for high spatial resolution multi-temporal images. *Remote Sens. Lett.* **2014**, *5*, 601–608. [[CrossRef](#)]
11. Başeski, E.; Cenaras, Ç. Texture and color based cloud detection. In Proceedings of the 2015 7th International Conference on Recent Advances in Space Technologies (RAST), Istanbul, Turkey, 16–19 June 2015; pp. 311–315.
12. Zhu, Z.; Wang, S.; Woodcock, C.E. Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images. *Remote Sens. Environ.* **2015**, *159*, 269–277. [[CrossRef](#)]
13. An, Z.; Shi, Z. Scene Learning for Cloud Detection on Remote-Sensing Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 4206–4222. [[CrossRef](#)]
14. Wu, T.; Hu, X.; Zhang, Y.; Zhang, L.; Tao, P.; Lu, L. Automatic cloud detection for high resolution satellite stereo images and its application in terrain extraction. *ISPRS J. Photogramm. Remote Sens.* **2016**, *121*, 143–156. [[CrossRef](#)]

15. An, F.; Song, S.H.; Luo, X.; Pu, G.L. Cloud Index in Remote Sensing Image Based on GeoSOT. *Geogr. Geo-Inf. Sci.* **2014**, *30*, 22–25. [[CrossRef](#)]
16. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. (University of Toronto) Imagenet classification with deep convolutional neural networks. In Proceedings of the NIPS'12 Proceedings of the 25th International Conference on Neural Information Processing Systems, Lake Tahoe, NV, USA, 3–6 December 2012; pp. 1106–1114.
17. Abadi, M.; Agarwal, A.; Barham, P.; Brevdo, E.; Chen, Z.; Citro, C.; Corrado, G.S.; Davis, A.; Dean, J.; Devin, M.; et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems. 2015. Available online: <https://arxiv.org/abs/1603.04467> (accessed on 18 August 2017).
18. TheIFAR-10 Dataset. Available online: <http://www.cs.toronto.edu/~kriz/cifar.html> (accessed on 18 August 2017).
19. CIFAR-10 Network. Available online: <https://github.com/tensorflow/models/tree/master/tutorials/image/cifar10> (accessed on 18 August 2017).
20. Geohash. Available online: <http://geohash.org/> (accessed on 18 August 2017).
21. GF-1 Satellite. Available online: <http://www.cresda.com/EN/satellite/7155.shtml> (accessed on 18 August 2017).
22. GF-2 Satellite. Available online: <http://www.cresda.com/EN/satellite/7157.shtml> (accessed on 18 August 2017).



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).