

# Recognizing Global Reservoirs From Landsat 8 Images: A Deep Learning Approach

Weizhen Fang , Cunguang Wang, Xi Chen , Wei Wan , Huan Li, Siyu Zhu , Yu Fang, Baojian Liu , and Yang Hong 

**Abstract**—Man-made reservoirs are key components of terrestrial hydrological systems. Identifying the location and number of reservoirs is the premise for studying the impact of human activities on water resources and environmental changes. While complete bottom-up censuses can provide a comprehensive view of the reservoir landscape, they are time-consuming and laborious and are thus infeasible on a global scale. Moreover, it is challenging to distinguish man-made reservoirs from natural lakes in remote sensing images. This study proposes a convolutional neural network (CNN)-based framework to recognize global reservoirs from Landsat 8 imageries. On the basis of the HydroLAKES dataset, a Landsat 8 cloud-free mosaic of 2017 was clipped for each feature (reservoir or lake) and was resized into  $224 \times 224$  patches, which were collected as training and testing samples. Compared to other deep learning methods (Alexnet and VGG) and state-of-the-art traditional machine learning methods (support vector machine, random forest, gradient boosting, and bag-of-visual-words), we found that fine-tuning the pretrained CNN model, ResNet-50, could reach the highest accuracy (91.45%). Application cases in Kansas (USA, North America), Mpumalanga (South Africa, Africa), and Kostanay (Kazakhstan, Asia) resulted in classification accuracies of better than 99%, which showed the applicability of the proposed ResNet-50 model to the extraction of reservoirs from a vast amount of moderate resolution images. The framework that was developed in this paper is the first attempt to combine remote sensing big data and the deep learning technique to the recognition of reservoirs at a global scale.

**Index Terms**—Convolutional neural network (CNN), deep learning, Landsat, object recognition, reservoir.

## I. INTRODUCTION

HUMANS have a long history of building reservoirs. As the main measures of flood control, water storage, irrigation, and power generation, the reservoir acts as an irreplaceable

Manuscript received January 20, 2019; revised April 11, 2019 and May 23, 2019; accepted July 5, 2019. Date of publication August 15, 2019; date of current version September 29, 2019. This work was supported by the Key R&D Program of Ministry of Science and Technology, China, under Grant 2018YFC1506504. (*Corresponding authors:* Wei Wan; Yang Hong.)

W. Fang, X. Chen, W. Wan, H. Li, B. Liu, and Y. Hong are with the Institute of Remote Sensing and GIS, School of Earth and Space Sciences, Peking University, Beijing 100871, China (e-mail: weizhenfang@pku.edu.cn; chenxi928@pku.edu.cn; w.wan@pku.edu.cn; scarlettlee33@foxmail.com; liubaojian@pku.edu.cn; yanghong588@pku.edu.cn).

C. Wang, S. Zhu, and Y. Fang are with the Department of Hydraulic Engineering, Tsinghua University, Beijing 100084, China (e-mail: cunguangwang@gmail.com; zhusy17@mails.tsinghua.edu.cn; fangyu516@gmail.com).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JSTARS.2019.2929601

indicator to track the interactive relationships between man and nature. Man-made reservoirs have a large impact on the global water cycle, since their substantial water storage has a significant negative contribution to the sea level [1], [2]. The previous study showed that since 1990, the number and storage capacity of reservoirs in the world have been booming, which has caused a drop in the global sea level by approximately 30 mm [3]. Reservoirs also cause a seasonal change in the monthly river runoff, with an average monthly increase or decrease of 30%, and many global hydrological models are eager to introduce the impact of reservoirs [4]. Hayes *et al.* [5] predicted that the reservoir's management characteristics made the reservoir a more important regulator of climate impact than natural lakes. Reservoirs can also cause substantial environmental damage. Recent studies have shown that the construction of man-made reservoirs has destroyed river networks and has greatly distorted the natural patterns of sediment transport and seasonal variations in river flows [6].

Although reservoirs have been proven by many studies to be crucial to the environment and society, the global dataset with geographic information is largely incomplete [7]. Some researchers and organizations have created their own global and regional dam and reservoir spatial datasets, most of which were sourced primarily through identification from paper maps and by compiling attribute information from various sources, including national water resource investigation and the Internet [7]–[9]. The number, quality, and spatial resolution of the records in these databases have varied widely, from coarse coordinates to centralized countries or regions. Moreover, to date, almost all reservoir records were obtained through a ground survey, which is time consuming and of high cost. Therefore, it is necessary to seek efficient ways to acquire large-scale reservoir distribution data.

There are many definitions of reservoirs [5]. In this paper, reservoirs are defined as an artificial, human-made lake, basin or tank with a dam in which a large quantity of water can be stored. From a morphological point of view, some reservoirs have an obvious dam and slender water outlets, which have mostly elongated shapes. In comparison, natural lakes are mostly round, and the border between the lake and its surroundings is relatively smooth and natural, which is visually distinguishable. Remotely sensed images contain abundant information about the earth's surface and are available at a global scale, which brings great potential for the mapping or monitoring of reservoirs all over the world [10]–[14].

The recognition of global reservoirs can be considered to be a target detection process from remote sensing images. Many methods have been developed to detect and recognize targets from remote sensing images [15]–[17] and can be divided briefly into unsupervised and supervised ones [18]. Unsupervised algorithms are related to probability theories and do not need signature models of targets. Supervised methods, which refer to knowledge that is related to the targets and rely on signatures of the targets during the training process, usually provide better accuracy than unsupervised ones. Advancement in machine learning has brought significant progress to supervised target recognition methods, and many new methods have been developed, such as the support vector machine (SVM), random forest (RF), and gradient boosting (GB) methods [19], [20]. In the past few years, the bag-of-visual-words (BoVW) method, which is based on  $K$ -means and SVM, has been used actively in many classification and target detection applications [21], [22] and leads to better accuracy [23], [24].

Recently, deep learning has become the new state-of-the-art machine learning solution for computer visual recognition [25]. Compared to shallow classifiers, such as SVM, deep neural networks with multiple hidden nodes can be used to construct better data models [26]. Krizhevsky *et al.* [27] made convolutional neural networks (CNNs) regain focus by showing superior image classification accuracy on the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) dataset. Recently, deep CNNs have been applied successfully to many pattern recognition tasks in the remote sensing field, such as image classification [28]–[33], object detection [34]–[36], and semantic segmentation [37], [38]. Compared to the traditional machine learning methods, deep learning relies heavily on large-scale training data, since it requires a large amount of data to understand the underlying data patterns [39]. Therefore, many high-resolution remote sensing image databases have been established, such as the UC Merced Land-Use dataset, NWPU-RESISC45 dataset, AID dataset, RSSCN7 dataset, and DOTA dataset [23], [40]–[43]. However, these databases are usually limited to RGB channels under a submeter resolution, and data collection is complex and expensive. At moderate resolution (e.g., Landsat), unlabeled multispectral remotely sensed images are freely available at a global scale. These unlabeled images can be matched with other datasets to put together a global-scale labeled dataset [44]. In addition, transfer learning is an emerging method that is useful for applications when lacking enough training data. Fine-tuning a pretrained CNN model may be an effective strategy for many deep learning model applications.

In this paper, we propose a deep learning approach to recognize global reservoirs from Landsat 8 imageries. The study has the following two main contributions.

- 1) A Global Reservoir and Lake Training (GRaLT) database clipped from moderate resolution Landsat 8 images is established. The database is of free access for scientific and educational applications (<https://github.com/Weizhen-Fang/GRaLT>).
- 2) A CNN-based framework is described, modified, and applied, for the first time, to recognize reservoirs at the global scale.

The remainder of this paper is organized as follows. Section II introduces the preprocessing method to generate training and test data, and the details of the CNN-model-based framework are employed to recognize global reservoirs in this paper. Section III focuses on the experimental results and Section IV on discussions. Finally, conclusion is given in Section V.

## II. METHODOLOGY

The proposed workflow for global reservoir recognition can be described as follows (see Fig. 1).

- 1) The Google Earth Engine (GEE) platform [45] was used to generate global Landsat 8 image mosaics.
- 2) A rigorous process was performed to produce a GRaLT database for training the CNN model.
- 3) The CNN model, ResNet-50, was described and modified.
- 4) A transfer learning strategy was used for training the ResNet-50, and the best one, which was based on obtaining the highest accuracy in testing data, was saved and applied in reservoir recognition.

### A. Remote Sensing Image Preprocessing in GEE

In this study, the USGS Landsat 8 Collection 1 Tier 1 calibrated top-of-atmosphere (TOA) reflectance product [46] was used as the only remote sensing data source. After loading all Landsat 8 TOA images of 2017, a Landsat cloud-free function was executed in the GEE platform, which helped to reduce the impact of the cloud and cloud shadows. Afterward, to create a global high-quality Landsat 8 image mosaic for the year 2017, for each image, the median of all values at each pixel across the stack of all matching bands was selected and the nearest neighbor interpolation was applied to prevent null values.

Ten bands (Coastal, Blue, Green, Red, NIR, SWIR 1, SWIR 2, Cirrus, TIRS 1, TIRS 2) of the mosaic were used to compose a 30-m resolution product. The Pan band was used as a one-band 15-m resolution product. The three water-sensitive bands (Red, NIR, SWIR 1) [37] and the Pan band were also used to fuse a three-band 15-m resolution product through an Hue, Saturation, Value (HSV)-based pan-sharpening method in GEE.

### B. Acquisition of the Training and Testing Data

The HydroLAKES dataset (<http://www.hydrosheds.org>) [9], which is a global vector dataset of lakes, was used as the benchmark for generating the training and testing datasets. To specify the study within a certain size area of the reservoirs, we selected a specific area threshold ( $\geq 10 \text{ km}^2$ ) to filter the shapefile. In addition, a buffer of  $0.02^\circ$  for each object was created to prevent the effect of geographic coordinate bias in HydroLAKES to ensure that each region contained a complete water body, i.e., a lake or reservoir. All the Landsat 8 mosaic products produced in Section II-A were clipped by the shapefile in GEE.

In the HydroLAKES dataset, only reservoirs that were mostly marked in the GRaLT database [7] have been distinguished from natural lakes, and the “lake” category contains both natural lakes

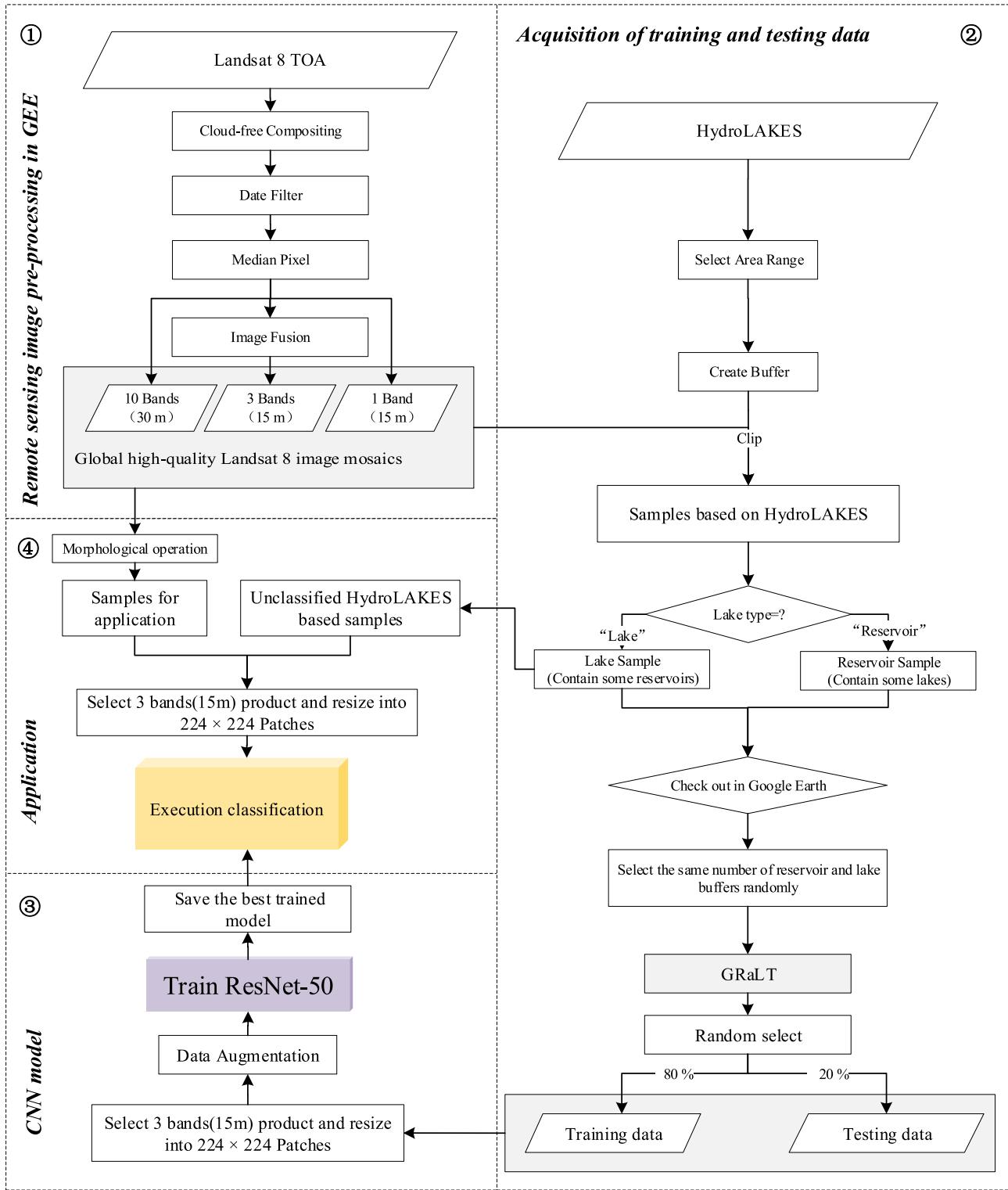


Fig. 1. Proposed workflow for reservoir recognition. TOA: Top of Atmosphere. GRaLT: Global Reservoir and Lake Training database.

and man-made reservoirs. Since the number of the “lake” category marked in the HydroLAKES is much larger than that of the reservoir, to ensure the equilibrium of the samples, we randomly selected the natural lakes in the “lake” category of HydroLAKES globally to maintain the same quantity of the reservoir samples

(i.e., using an undersampling solution). Afterward, all the samples were visually checked again through high-resolution remote sensing images in Google Earth to ensure that they were labeled correctly. Samples containing artificial facilities (dams) and having water storage abilities were categorized as reservoirs.

Different band combinations were made to constituting the GRaLT database, and 80% of the data was randomly selected for training and the remaining 20% was selected for testing.

### C. Architecture of the ResNet-50 Model

ResNet, which was the winner of the 2015 ILSVRC, is characterized by a very deep network with 34–152 layers [47]. A residual network structure is introduced in the ResNet model. By using the residual network structure, the deep CNN model not only avoids the problem of model degradation but also achieves better performance.

The structure of ResNet-50 is shown in Fig. 2. Although there are many deeper ResNet networks available, the calculation time, model size, and difficulty of training on small training sets limit their application in this study. In ResNet-50, a structure called Bottleneck is used to reduce calculations and parameter quantities. In Bottleneck architecture, the three layers are  $1 \times 1$ ,  $3 \times 3$ , and  $1 \times 1$  convolutions, where the two  $1 \times 1$  layers play the role of reducing and then increasing dimensions, which gives the  $3 \times 3$  layer the smallest input/output dimensions [47].

Similar to the VGG model [48],  $3 \times 3$  filters were mostly employed in this network. However, ResNet-50 has fewer filters and less complexity relative to VGG. Many studies have applied this model and received reasonable results [49]–[51].

ResNet was previously designed to classify the 1000-class RGB-band ImageNet database, while in this study we changed the last layer of the model to fit the two-category problem. Moreover, if the input data are not a three-band image, the depth of the first convolution was also changed to the depth of the input image.

### D. CNN Model Training and Classification

To match the architecture of ResNet-50, all images were resized into  $224 \times 224$  patches. Data augmentation was performed to artificially increase the number of training examples to avoid overfitting. Specifically, all images were rotated by four angles (i.e.,  $0^\circ$ ,  $90^\circ$ ,  $180^\circ$ , and  $270^\circ$ ), flipped horizontally and vertically, which sextupled the data quantity (see Fig. 3).

Training deep CNNs requires a significant quantity of labeled training samples, and the existing remote sensing datasets cannot provide sufficient samples. Moreover, the overfitting problem is also challenging. As a solution, in this study, we used the transfer learning strategy.

All the weights in the ResNet-50 need to be trained. The study showed that the weights trained on RGB data can also be helpful to process the remote sensing data [32]. We initialized all layers with the weights from ResNet trained on the ImageNet database [47], except that the last fully connected layer was initialized by the Xavier initialization method [52]. In addition, when not trained the model by the three-band dataset, we initialized the first convolutional layer by Xavier additionally [52].

## III. EXPERIMENTS AND RESULTS

In this section, a proposed method for reservoir recognition was evaluated. Experiments were implemented based on the

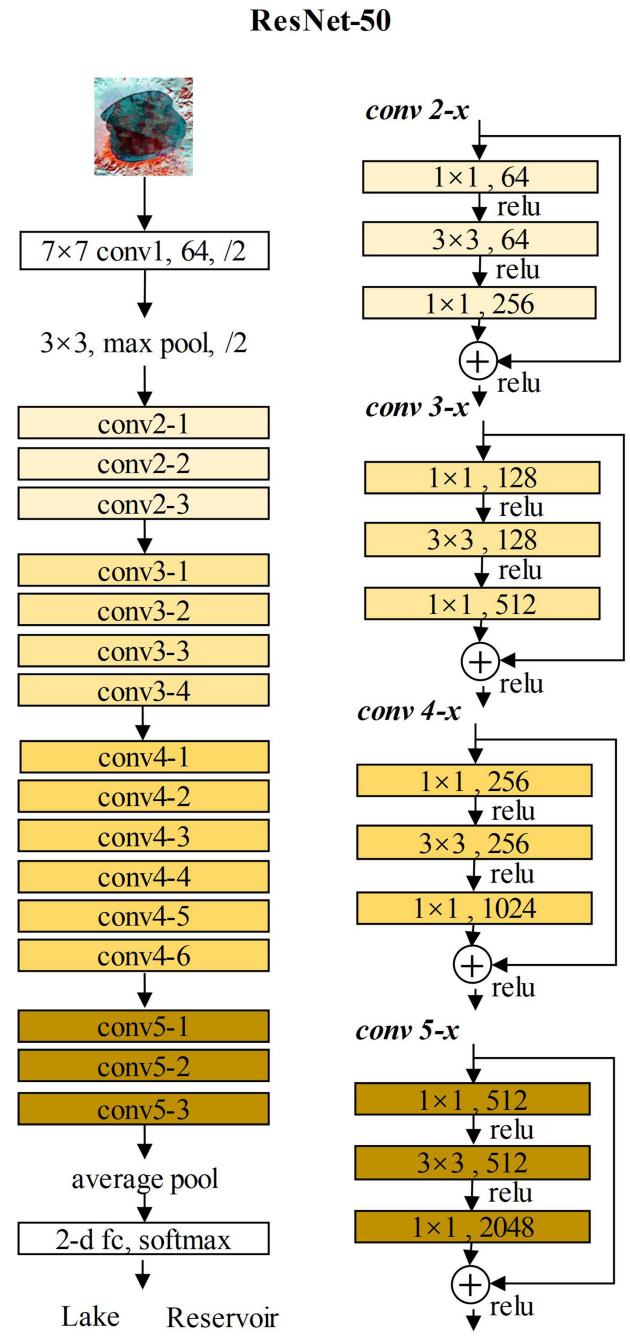


Fig. 2. Structure of the ResNet-50 used for reservoir recognition.

deep learning framework Caffe and were executed on NVIDIA GeForce GTX Titan X (Pascal).

### A. GRaLT Database

We used the approach in Section II to establish the GRaLT, which was subsequently used to train and evaluate the performance of the proposed method. The distribution of GRaLT is shown in Fig. 4. All the reservoir and lake samples were clipped from a one-year collection of Landsat 8 images of year 2017.

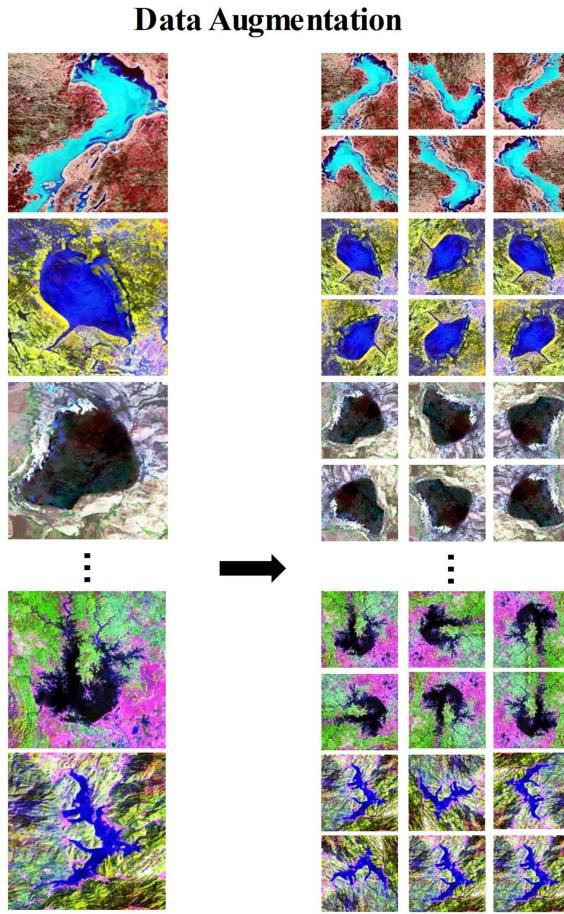


Fig. 3. Data augmentation.

Table I lists the characteristics of GRaLT. Six kinds of data combinations were tested, including Band all, Band 543, Band 654, and Water Contour (the pixel values of the Modification of Normalized Difference Water Index (MNDWI) that are larger than 0 are set to 1, and the others are set to 0) [53], [54] with 30-m resolution, and Band 8 and Pan 654 with 15-m resolution. After random selection, there were 1520 reservoirs and 1520 lakes in the training dataset and 380 reservoirs and 380 lakes in the testing dataset.

#### B. Accuracy Evaluation and Comparison

The accuracies of the experiments were evaluated from two aspects: 1) accuracy comparisons of different machine learning models; and 2) accuracy comparisons of different band combinations using fine-tuned ResNet-50.

We used four common metrics to do the evaluation, i.e., precision rate, recall rate, *F*-score, and overall accuracy. They are defined as follows:

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (1)$$

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (2)$$

$$\begin{aligned} F - \text{score} &= 2 \cdot \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \\ &= \frac{2\text{TP}}{2\text{TP} + \text{FN} + \text{FP}} \end{aligned} \quad (3)$$

$$\text{Overall accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}} \quad (4)$$

where TP denotes true positives (correctly classified reservoir images), FP denotes false positives (images mislabeled as reservoirs), TN denotes true negatives (correctly classified lake images), and FN denotes false negatives (images incorrectly labeled as lakes).

Table II lists the quantitative comparison results of seven different models, including four state-of-the-art traditional machine-learning models (SVM, RF, GB, and BoVW) and three deep learning models (Alexnet, VGG-16, and ResNet-50). For ResNet-50, we also compared the “loss function” solution [55] with the “undersampling” solution that is used in this study. The “loss function” solution used imbalanced samples (i.e., 2370 lakes and 1900 reservoirs), which were acquired additionally during the GRaLT data production.

We used Pan 654 in GRaLT to train and test the models. For SVM, RF, and GB, to acquire the best accuracy, we used 28 manually generated texture and shape features (e.g., the mean value of the image, the maximum value of the homogeneity of gray-level co-occurrence matrix in all directions of the image, the density of the region [56], and the shape index of the region [56], etc.) as inputs to optimize the models. The speeded up robust features (SURF) operator [57] was used in BoVW to extract the feature description of the image. Machine-learning methods were implemented and were optimized in scikit-learn [58].

The result shows that all of the state-of-the-art deep learning models achieve a precision of over 91%, and the fine-tuned ResNet-50 obtained the highest accuracy in three metrics, i.e., precision = 94.37%, *F*-score = 91.16%, and overall accuracy = 91.45%. In addition, the deep learning methods always show better accuracies than traditional methods.

Table III lists the accuracy evaluation results of six different band combinations. Note that five of them show good results, except for data obtained with Band 8. In terms of overall accuracy, Pan 654 data performed the best, with an accuracy of 91.45%. Band all, Band 654, and Pan 654 data achieved a better *F*-score accuracy (>91%). Pan 654 achieved the best precision (94.37%), which means that among all the samples judged to be reservoirs, the proportion of the correct classification was the highest, which makes it the most suitable one for reservoir identification. Furthermore, as shown in the results of Band all, putting all bands of Landsat 8 into the network will not improve the accuracy.

#### C. Applications

To evaluate the capability of the model in the unlabeled area, three applications in Kansas (USA), Mpumalanga (South Africa), and Kostanay (Kazakhstan) are shown in Fig. 5. Kansas, which has many large reservoirs, is located in the middle of North

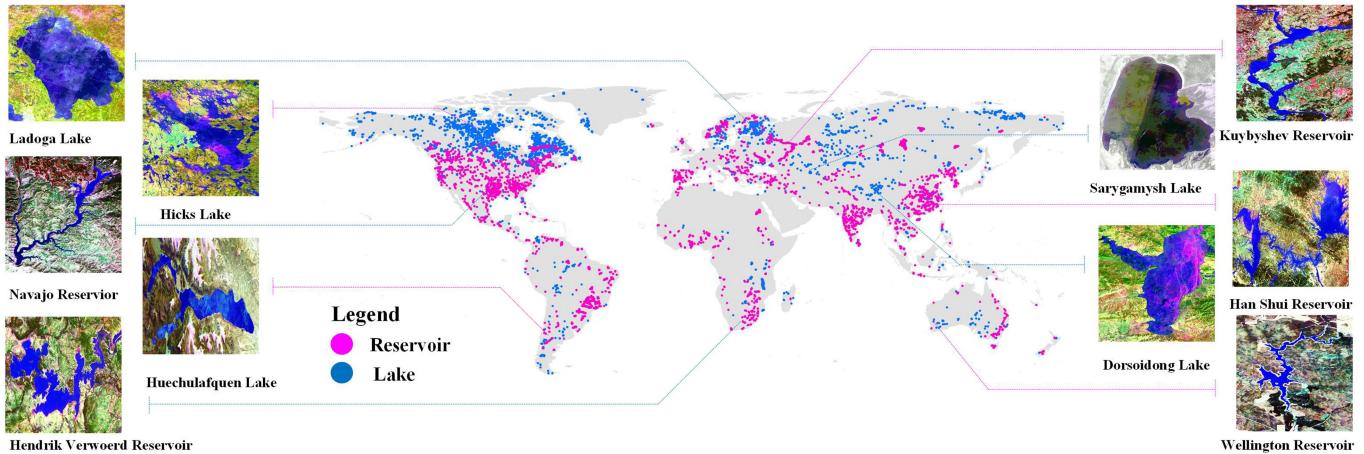


Fig. 4. Distribution of the GRaLT database. (The pink dots are reservoirs and the blue dots are lakes.) Training and testing data (examples after resizing into  $224 \times 224$  patches are given on the left and right sides of the figure). Lakes (Ladoga, Hicks, Taciula, Sarygamysh, Dorsoidong) and reservoirs (Navajo, Hendrik Verwoerd, Kuybyshev, Han Shui, Wellington) were all extracted from Landsat 8 TOA images with the buffered HydroLAKES layer.

TABLE I  
CHARACTERISTICS OF THE GRaLT

Name	Band Combination	Resolution (m)	Number of training samples	Number of testing samples
Band all	Coastal \ Blue \ Green \ Red \ NIR \ SWIR 1 \ SWIR 2 \ Pan \ Cirrus \ TIRS 1 \ TIRS 2			
Band 543	Green \ Red \ NIR	30	3040	760
Band 654	Red \ NIR \ SWIR 1			
Water Contour	MNDWI > 0			
Band 8	Pan			
Pan 654	Red \ NIR \ SWIR 1	15		

TABLE II  
ACCURACY EVALUATIONS OF DIFFERENT METHODS

Methods	Precision	Recall	F-score	Overall Accuracy
SVM	84.47%	83.38%	83.92%	83.82%
RF	84.21%	81.84%	83.01%	82.76%
GB	85.05%	82.37%	83.69%	83.95%
SURF+BOVW	77.75%	80.00%	78.86%	78.55%
Fine-tuned Alexnet	92.12%	<b>89.21%</b>	90.64%	90.79%
Fine-tuned VGG-16	91.85%	88.95%	90.37%	90.53%
Fine-tuned ResNet-50 (loss function)	92.27%	87.89%	90.03%	90.26%
Fine-tuned ResNet-50	<b>94.37%</b>	88.16%	<b>91.16%</b>	<b>91.45%</b>

TABLE III  
ACCURACY COMPARISON BETWEEN DIFFERENT BAND COMBINATIONS

Band combination	Precision	Recall	F-score	Overall Accuracy
Band all	92.20%	90.26%	91.22%	91.32%
Band 543	93.80%	87.63%	90.61%	90.92%
Band 654	91.32%	<b>91.32%</b>	<b>91.32%</b>	91.32%
Band 8	84.33%	70.79%	76.97%	78.82%
Pan 654	<b>94.37%</b>	88.16%	91.16%	<b>91.45%</b>
Water Contour	92.09%	85.79%	88.83%	89.21%

America. The western two-thirds of the state, lying in the great central plain of the USA, have a generally flat or undulating surface. Mpumalanga is located in the southern part of Africa, with the largest shallow freshwater lake, Lake Chrissie, located there. Kostanay is a region located in northern Kazakhstan and in the middle of Asia. It is supplied with water from Aman-geldinsky water basins with many lakes. The processing steps of the applications on the three regions are listed as follows.

- 1) The buffered lake and reservoir shapefiles from HydroLAKES in Kansas, Mpumalanga, and Kostanay were used to clip the region from the global high-quality Landsat 8 15-m resolution band 654 combined mosaics of 2017 in GEE [see Fig. 5(a1)–(a3)].
- 2) The water parts of the image were extracted through three water-sensitive bands [37] to find the water (marked as 1) and the nonwater (marked as 0) pixels in the image [see Fig. 5(b1)–(b3)].
- 3) The morphological operation and manual selection were performed to clip the water bodies which area is greater than or equal to  $10 \text{ km}^2$  in Kansas and Kostanay, and in Mpumalanga is greater than or equal to  $2.5 \text{ km}^2$  to ensure enough water body samples extracted [see Fig. 5(c1)–(c3)].

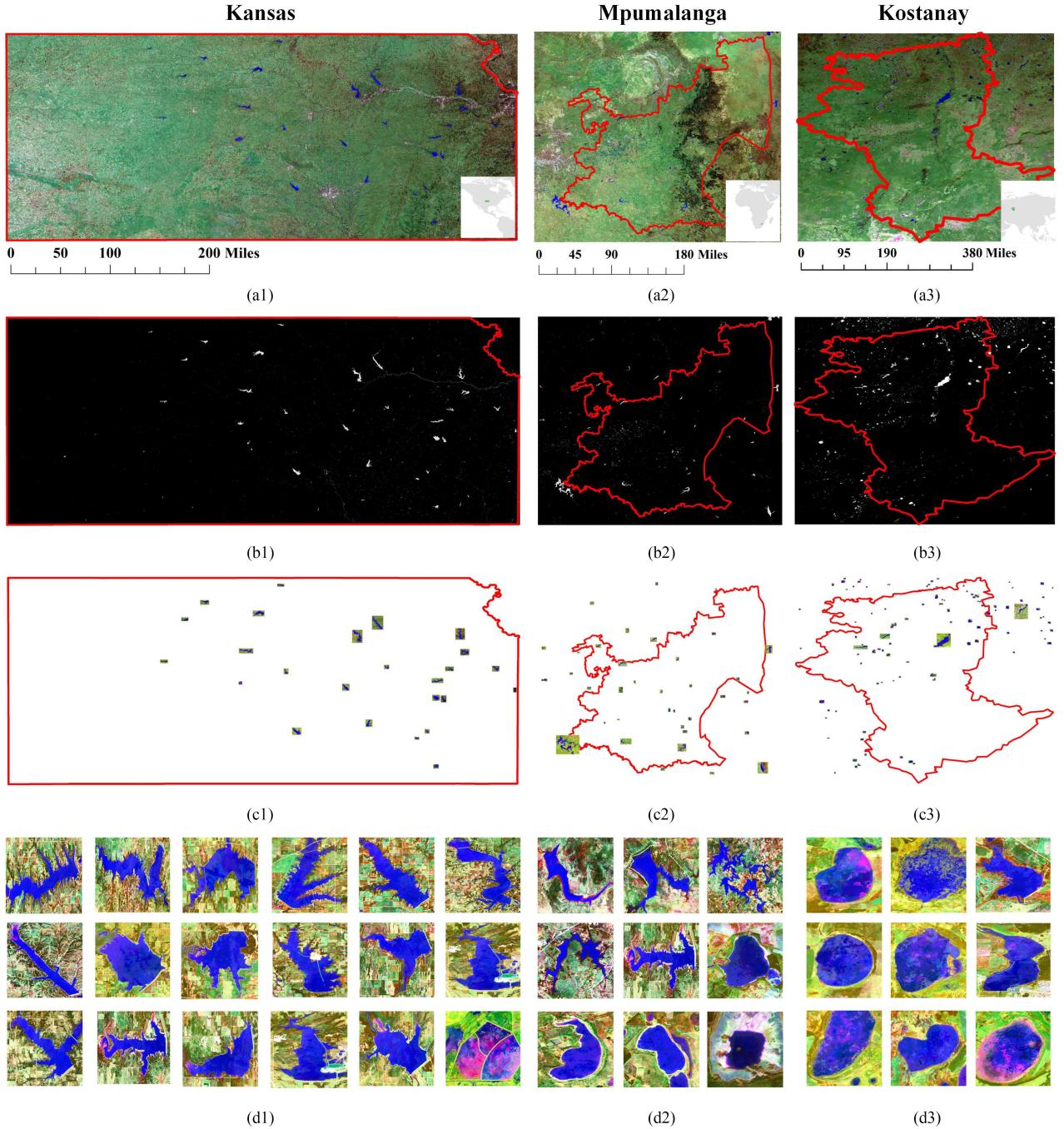


Fig. 5. Applications of the proposed method in Kansas, USA; Mpumalanga, South Africa; and Kostanay, Kazakhstan. (a1)–(a3) High-quality Landsat 8 15-m resolution Band 654 combined mosaics in 2017. (b1)–(b3) Water (white pixels) and the nonwater (black pixels) parts in the image. (c1)–(c3) Clips of the specific water body. (d1)–(d3) Part of the classification data.

- 4) Before classifying with trained CNN model, we performed the same preprocessing operations on all images as the training data (see Section II-B), which is shown in Fig. 5(d1)–(d3).
- 5) All images were classified with the best ResNet-50 model trained in GRAILT.

The accuracy comparison between different application regions is shown in Table IV. Note that all the reservoirs and lakes in Kansas and Mpumalanga were accurately classified,

TABLE IV  
ACCURACY COMPARISON BETWEEN DIFFERENT APPLICATION REGIONS

Region	Reservoir Number	Lake Number	Area	Overall Accuracy
Kansas	25	1	$\geq 10 \text{ km}^2$	100%
Mpumalanga	33	6	$\geq 2.5 \text{ km}^2$	100%
Kostanay	4	107	$\geq 10 \text{ km}^2$	99.1%

whereas in Kostanay only one sample was misclassified. The result shows that the model can accurately pick up the reservoirs in the regions regardless of where the regions have more reservoirs or lakes. The result also shows the ability of the model to classify smaller reservoirs [e.g., in Mpumalanga, there are some water bodies whose areas are smaller than the training data ( $\geq 10 \text{ km}^2$ )]. These application examples demonstrate the practical application value of the trained model as well as its potential to be applied to accurately extract reservoirs from a random moderate resolution image.

#### IV. DISCUSSION

##### A. Advantages of the Proposed Framework

For this application (i.e., small dataset), the key advantage of ResNet-50 is that its residual structure helps to solve problems such as gradient disappearance and gradient explosion, allowing us to train deeper networks while maintaining good performance. Nevertheless, there is still some risk of overfitting when deepening the layer depth, even though we used data augmentation and fine-tuned training strategies.

ResNet-50 is also proven to be an efficient and effective approach for this application. First, the number of parameters of ResNet-50 is minimal (ResNet-50: 25 million; VGG-16: 138 million; Alexnet: 60 million) due to its specific “residual” structure. This makes ResNet-50 an efficient approach for this application. Second, in many cases, it is difficult or even impossible to collect large dataset (such as ImageNet [59]) for training. A network (such as ResNet) with fewer parameters would have unique advantage of dealing with small dataset [60]. This makes ResNet-50 an effective approach for this application.

Based on the results of this study (see Table II), there is probably no direct relationship between the depth of the layer and the accuracy. For example, compared to Alexnet, the layer depth of VGG-16 is deeper (16 versus 8), but the accuracy is lower (overall accuracy: 90.53% versus 90.79%; *F*-score: 90.37% versus 90.64%).

##### B. Limitations of the Data and Method

The proposed method has limitations when dealing with water body images with man-made facilities rather than dams. Morphological structure is a key feature to distinguish reservoirs and lakes, and straight structures can usually be found near the dam of the reservoir. Around the natural lakes in some cities, the establishment of man-made facilities (i.e., roads and bridges) has significantly changed the natural state of the lake, making it difficult to be distinguished from man-made reservoirs in moderate resolution remote sensing images.

Besides, the HydroLAKES data are collected from multiple sources in different time scales. Some reservoirs were demolished and some lakes may have changed in their extents (or even disappeared) in recent times or may have undergone strong seasonal fluctuations, causing the clips to be improperly represented by the temporal snapshot. In addition, when creating the HydroLAKES database, the distinction between lakes, rivers, and wetlands is a difficult and important issue [9]. Despite the

rigorous review, there are still some inevitable mistakes, which may lead to diminished accuracy.

Further, for evaluating results of six different band combinations in Table III, due to limitations of the structural characteristics of CNN, we cannot keep the network structures exactly the same. Although the best band combination was found for application, the manner in which different band combinations influence the accuracy remains uncertain.

#### V. CONCLUSION AND FUTURE WORK

In this paper, we presented a CNN-based framework to recognize reservoirs at the global scale with high accuracies. A GRALT database clipped from Landsat 8 TOA using HydroLAKES database was generated for training the CNN model. The results draw the following conclusions.

- 1) The fine-tuned deep learning model, Resnet-50, has better accuracy than used traditional models in reservoir recognition.
- 2) The band combination can influence the classification accuracy. It verifies that the Landsat 8 band combination of SWIR 1, NIR, and Red band (Pan 654, 15-m resolution after fusion using Pan band) achieves the best precision (94.37%) and overall accuracy (91.45%).
- 3) Application experiments in Kansas, USA; Mpumalanga, South Africa; and Kostanay, Kazakhstan further confirm the superior capability of the deep learning model and its potential of application for recognizing smaller reservoirs.

For future work, we will focus on extending the capability of the trained deep learning model to classify the unlabeled lakes in HydroLAKES to obtain the number and distribution of the man-made reservoirs and natural lakes, thus providing an advanced remote-sensing-based dataset of the reservoirs.

#### APPENDIX

The Global Reservoir and Lake Training database is offered for free for scientific and educational applications at <https://github.com/Weizhen-Fang/GRALT>.

#### ACKNOWLEDGMENT

The authors would like to thank the United States Geological Survey and Messager *et al.* [9] for providing Landsat 8 and HydroLAKES data.

#### REFERENCES

- [1] D. L. Sahagian, F. W. Schwartz, and D. K. Jacobs, “Direct anthropogenic contributions to sea level rise in the twentieth century,” *Nature*, vol. 367, pp. 54–57, Jan. 6, 1994.
- [2] F. C. Benjamin, “Anthropogenic impact on global geodynamics due to reservoir water impoundment,” *Geophys. Res. Lett.*, vol. 22, no. 24, pp. 3529–3532, Dec. 15, 1995.
- [3] B. F. Chao, Y. H. Wu, and Y. S. Li, “Impact of artificial reservoir water impoundment on global sea level,” *Science*, vol. 320, pp. 212–214, Apr. 11, 2008.
- [4] I. Haddeland, T. Skaugen, and D. P. Lettenmaier, “Anthropogenic impacts on continental surface water fluxes,” *Geophys. Res. Lett.*, vol. 33, no. 8, 2006, Art. no. L08406.
- [5] N. M. Hayes, B. R. Deemer, J. R. Corman, N. R. Razavi, and K. E. Strock, “Key differences between lakes and reservoirs modify climate signals: A case for a new conceptual model,” *Limnol. Oceanogr. Lett.*, vol. 2, no. 2, pp. 47–62, 2017.

- [6] N. L. Poff and J. C. Schmidt, "How dams can go with the flow," *Science*, vol. 353, no. 6304, pp. 1099–1100, 2016, doi: 10.1126/science.aah4926.
- [7] B. Lehner *et al.*, "High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management," *Frontiers Ecol. Environ.*, vol. 9, no. 9, pp. 494–502, 2011.
- [8] B. Lehner and P. Döll, "Development and validation of a global database of lakes, reservoirs and wetlands," *J. Hydrol.*, vol. 296, no. 1–4, pp. 1–22, 2004.
- [9] M. L. Messager, B. Lehner, G. Grill, I. Nedeva, and O. Schmitt, "Estimating the volume and age of water stored in global lakes using a geo-statistical approach," *Nature Commun.*, vol. 7, Dec. 15, 2016, Art. no. 13603.
- [10] L. N. Rodrigues, E. E. Sano, T. S. Steenhuis, and D. P. Passo, "Estimation of small reservoir storage capacities with remote sensing in the Brazilian Savannah region," *Water Resour. Manage.*, vol. 26, no. 4, pp. 873–882, Mar. 2012.
- [11] H. Gao, C. Birkett, and D. P. Lettenmaier, "Global monitoring of large reservoir storage from satellite remote sensing," *Water Resour. Res.*, vol. 48, no. 9, 2012, Art. no. W09504.
- [12] X. Yang and X. Lu, "Drastic change in China's lakes and reservoirs over the past decades," *Sci. Rep.*, vol. 4, Aug. 13, 2014, Art. no. 6041.
- [13] J.-F. Crétaux, S. Biancamaria, A. Arsen, M. Bergé-Nguyen, and M. Becker, "Global surveys of reservoirs and lakes from satellites and regional application to the Syrdarya river basin," *Environ. Res. Lett.*, vol. 10, no. 1, 2015, Art. no. 015002.
- [14] H. Zhang, S. M. Gorelick, P. V. Zimba, and X. Zhang, "A remote sensing method for estimating regional reservoir area and evaporative loss," *J. Hydrol.*, vol. 555, pp. 213–227, 2017.
- [15] J. Yang, L. Tang, B. Zhao, C. Deng, and B. Wang, "Airport target detection algorithm in remote sensing images based on JPEG2000 compressed domain," in *Proc. IET Int. Radar Conf.*, 2013, pp. 1–4.
- [16] Y. Lin, H. He, H.-M. Tai, F. Chen, and Z. Yin, "Rotation and scale invariant target detection in optical remote sensing images based on pose-consistency voting," *Multimedia Tools Appl.*, vol. 76, no. 12, pp. 14461–14483, 2016.
- [17] F. Xu, J. Liu, M. Sun, D. Zeng, and X. Wang, "A hierarchical maritime target detection method for optical remote sensing imagery," *Remote Sens.*, vol. 9, no. 3, p. 280, 2017.
- [18] N. Poojary, H. D. Souza, M. R. Puttaswamy, and G. H. Kumar, "Automatic target detection in hyperspectral image processing: A review of algorithms," in *Proc. 12th Int. Conf. Fuzzy Syst. Knowl. Discovery*, 2015, pp. 1991–1996.
- [19] C. Huang, L. S. Davis, and J. R. G. Townshend, "An assessment of support vector machines for land cover classification," *Int. J. Remote Sens.*, vol. 23, no. 4, pp. 725–749, 2010.
- [20] S. Mahdavi, B. Salehi, J. Granger, M. Amani, B. Brisco, and W. Huang, "Remote sensing for wetland classification: A comprehensive review," *GISci. Remote Sens.*, vol. 55, no. 5, pp. 623–658, 2017.
- [21] M. Dinesh Kumar, M. Babaie, S. Zhu, S. Kalra, and H. Tizhoosh, "A comparative study of CNN, BoVW and LBP for classification of histopathological images," in *Proc. IEEE Symp. Ser. Comput. Intell.*, 2017, pp. 1–7.
- [22] N. Aldeborgh, G. Ouzounis, and K. Stamatou, "Unsupervised object detection on remote sensing imagery using hierarchical image representations and deep learning," in *Proc. Conf. Big Data Space: Publications Office Eur. Union*, 2017, pp. 255–258.
- [23] Y. Yang and S. Newsam, "Bag-of-visual-words and spatial extensions for land-use classification," presented at the Proc. 18th SIGSPATIAL Int. Conf. Advances Geographic Information Systems, San Jose, CA, USA, 2010.
- [24] F. Hu, G. Xia, Z. Wang, L. Zhang, and H. Sun, "Unsupervised feature coding on local patch manifold for satellite image scene classification," in *Proc. IEEE Geosci. Remote Sens. Symp.*, 2014, pp. 1273–1276.
- [25] L. Zhang, L. Zhang, and B. Du, "Deep learning for remote sensing data: A technical tutorial on the state of the art," *IEEE Geosci. Remote Sens. Mag.*, vol. 4, no. 2, pp. 22–40, Jun. 2016.
- [26] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 28, 2015.
- [27] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," presented at the Proc. 25th Int. Conf. Neural Information Processing Systems, Lake Tahoe, NV, USA, 2012, vol. 1.
- [28] J. Geng, H. Wang, J. Fan, and X. Ma, "SAR image classification via deep recurrent encoding neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 4, pp. 2255–2269, Apr. 2018.
- [29] Y. Liu, Y. Liu, and L. Ding, "Scene classification based on two-stage deep feature fusion," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 2, pp. 183–186, Feb. 2018.
- [30] C. Zhang, I. Sargent, X. Pan, A. Gardiner, J. Hare, and P. M. Atkinson, "VPRS-based regional decision fusion of CNN and MRF classifications for very fine resolution remotely sensed images," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 8, pp. 4507–4521, Aug. 2018.
- [31] J. Li, X. Zhao, Y. Li, Q. Du, B. Xi, and J. Hu, "Classification of hyperspectral imagery using a new fully convolutional neural network," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 2, pp. 292–296, Feb. 2018.
- [32] D. Marmanis, M. Datcu, T. Esch, and U. Stilla, "Deep learning earth observation classification using ImageNet pretrained networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 1, pp. 105–109, Jan. 2016.
- [33] L. Zhang, A. Li, Z. Zhang, and K. Yang, "Global and local saliency analysis for the extraction of residential areas in high-spatial-resolution remote sensing image," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 7, pp. 3750–3763, Jul. 2016.
- [34] P. Zhang, X. Niu, Y. Dou, and F. Xia, "Airport detection on optical satellite images using deep convolutional neural networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 8, pp. 1183–1187, Aug. 2017.
- [35] J. Zuo, G. Xu, K. Fu, X. Sun, and H. Sun, "Aircraft type recognition based on segmentation with deep convolutional neural networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 2, pp. 282–286, Feb. 2018.
- [36] Z. Zou and Z. Shi, "Ship detection in spaceborne optical image with SVD networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 10, pp. 5832–5845, Oct. 2016.
- [37] H. Xie, X. Luo, X. Xu, H. Pan, and X. Tong, "Evaluation of Landsat 8 OLI imagery for unsupervised inland water extraction," *Int. J. Remote Sens.*, vol. 37, no. 8, pp. 1826–1844, 2016.
- [38] R. Kemker, C. Salvaggio, and C. Kanan, "Algorithms for semantic segmentation of multispectral remote sensing imagery using deep learning," *ISPRS J. Photogramm. Remote Sens.*, vol. 145, pp. 60–77, 2018.
- [39] C. Tan, F. Sun, T. Kong, W. Zhang, C. Yang, and C. Liu, "A survey on deep transfer learning," V. Kůrková, Y. Manolopoulos, B. Hammer, L. Iliadis, and I. Maglogiannis, Eds., *Artificial Neural Networks and Machine Learning - ICANN 2018*. Lecture Notes in Computer Science, Springer, Cham, vol. 11141, 2018.
- [40] G. Sheng, W. Yang, T. Xu, and H. Sun, "High-resolution satellite scene classification using a sparse coding based multiple feature combination," *Int. J. Remote Sens.*, vol. 33, no. 8, pp. 2395–2412, 2011.
- [41] Q. Zou, L. Ni, T. Zhang, and Q. Wang, "Deep learning based feature selection for remote sensing scene classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 11, pp. 2321–2325, Nov. 2015.
- [42] G. S. Xia *et al.*, "DOTA: A large-scale dataset for object detection in aerial images," in *Proc. IEEE/CVF Conf. Comput. Vision Pattern Recognit.*, 2018, pp. 3974–3983.
- [43] G.-S. Xia *et al.*, "AID: A benchmark data set for performance evaluation of aerial scene classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 7, pp. 3965–3981, Jul. 2017.
- [44] F. Isikdogan, A. C. Bovik, and P. Passalacqua, "Surface water mapping by deep learning," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 11, pp. 4909–4918, Nov. 2017.
- [45] N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore, "Google Earth engine: Planetary-scale geospatial analysis for everyone," *Remote Sens. Environ.*, vol. 202, pp. 18–27, 2017.
- [46] G. Chander, B. L. Markham, and D. L. Helder, "Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors," *Remote Sens. Environ.*, vol. 113, no. 5, pp. 893–903, 2009.
- [47] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vision Pattern Recognit.*, 2016, pp. 770–778.
- [48] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *ICLR*, 2015.
- [49] Z. Tinghe, L. Yuhui, Y. Qiankun, H. Hong, and F. Tao, "Integrating saliency and ResNet for airport detection in large-size remote sensing images," in *Proc. 2nd Int. Conf. Image, Vision Comput.*, 2017, pp. 20–25.
- [50] E. Rezende, G. Ruppert, T. Carvalho, F. Ramos, and P. de Geus, "Malicious software classification using transfer learning of ResNet-50 deep neural network," in *Proc. 16th IEEE Int. Conf. Mach. Learn. Appl.*, 2017, pp. 1011–1014.
- [51] H. Jung, M.-K. Choi, J. Jung, J.-H. Lee, S. Kwon, and W. Y. Jung, "ResNet-based vehicle classification and localization in traffic surveillance systems," in *Proc. IEEE Conf. Comput. Vision Pattern Recognit. Workshops*, 2017, pp. 934–940.
- [52] G. Xavier and B. Yoshua, "Understanding the difficulty of training deep feedforward neural networks," in *Proc. Proc. 13th Int. Conf. Artif. Intell. Statist.*, Mar. 31, 2010, pp. 249–256. [Online]. Available: <http://proceedings.mlr.press/v9/glorot10a.html>

- [53] H. Xu, "Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery," *Int. J. Remote Sens.*, vol. 27, no. 14, pp. 3025–3033, 2007.
- [54] L. Ji, X. Geng, K. Sun, Y. Zhao, and P. Gong, "Target detection method for water mapping using Landsat 8 OLI/TIRS imagery," *Water*, vol. 7, no. 12, pp. 794–817, 2015.
- [55] T. Lin, P. Goyal, R. Girshick, K. He, and P. Dollar, "Focal loss for dense object detection," *IEEE Trans. Pattern Analysis Mach. Intell.*, p. 1, 2018, doi: [10.1109/TPAMI.2018.2858826](https://doi.org/10.1109/TPAMI.2018.2858826).
- [56] X. Huang, C. Xie, X. Fang, and L. Zhang, "Combining pixel- and object-based machine learning for identification of water-body types from urban high-resolution remote-sensing imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 5, pp. 2097–2110, May 2015.
- [57] H. Bay, T.uytelaars, and L. Van Gool, "SURF: Speeded up robust features," in *Proc. Eur. Conf. Comput. Vision*, 2006, pp. 404–417.
- [58] R. Garreta and G. Moncecchi, *Learning Scikit-Learn: Machine Learning in Python*. Birmingham, U.K.: Packt Publ., 2013, p. 118.
- [59] O. Russakovsky *et al.*, "ImageNet large scale visual recognition challenge," *Int. J. Comput. Vision*, vol. 115, no. 3, pp. 211–252, 2015.
- [60] R. Keshari, M. Vatsa, R. Singh, and A. Noore, "Learning structure and strength of CNN filters for small sample size training," Mar. 1, 2018, *arXiv:1803.11405*. [Online]. Available: <https://ui.adsabs.harvard.edu/abs/2018arXiv180311405K>



**Weizhen Fang** received the B.S. degree from the College of Geological Engineering and Geomatics, Chang'an University, Xi'an, China, in 2017. He is currently working toward the master's degree with the Institute of Remote Sensing and GIS, School of Earth and Space Sciences, Peking University, Beijing, China.

His current research focuses on water resources remote sensing and deep learning application in remote sensing images.



**Cunguang Wang** received the B.S. degree in hydraulic engineering in 2016 from Tsinghua University, Beijing, China, where he has been working toward the Ph.D. degree with the Department of Hydraulic Engineering since 2016.

His current research focuses on remote sensing precipitation and deep learning application in geo-science field.



**Xi Chen** received the Ph.D. degree from the Faculty of Geographical Science, Beijing Normal University, Beijing, China, in 2017.

He is currently a Postdoc Researcher with the Institute of Remote Sensing and GIS, School of Earth and Space Sciences, Peking University, Beijing. His current research focuses on fusing multisource geographical information for decision support, including the fusion of remote sensing images, marine acoustic data, and statistical information.



**Wei Wan** received the M.E. degree from Nanjing University, Nanjing, China, in 2010, and the Ph.D. degree from Peking University, Beijing, China, in 2014, both in photogrammetry and remote sensing.

From 2014 to 2017, she was a Postdoctoral Researcher with Tsinghua University, Beijing. Since 2017, she has been an Assistant Research Fellow with the Institute of Remote Sensing and GIS, Peking University. Her research interests include GNSS remote sensing and satellite remote sensing of lakes and reservoirs.



**Huan Li** received the B.S. degree in remote sensing science and technology from the School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China, in 2010, and the M.S. degree in surveying engineering and the Ph.D. degree in cartography and geographic information engineering from the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, in 2012 and 2016, respectively.

She was a Postdoctoral Researcher with Tsinghua University, Beijing, China, in 2018. She is currently a Postdoctoral Researcher with the School of Earth and Space Sciences, Peking University, Beijing. Her research interests include remote sensing and GIS big data for water storage and polar sea ice.



**Siyu Zhu** received the M.E. degree in hydraulics and hydropower engineering from Tsinghua University, Beijing, China, in 2019. He is currently working toward the Ph.D. degree with the School of Earth and Space Science, Peking University, Beijing.

His research interests include satellite remote sensing of lakes and reservoirs, and the combination of remote sensing and artificial intelligence.



**Yu Fang** received the B.S. degree from the Minzu University of China, Beijing, China, in 2009, the M.S. degree from the Chinese Academy of Sciences, Beijing, in 2012, and the Ph.D. degree from the University of Florida, Gainesville, FL, USA, in 2016.

She is currently a Postdoctoral Researcher with the Department of Hydraulic Engineering, Tsinghua University, Beijing. Her research interests include hydrology and water resources, human–water interaction, remote sensing application, ecosystem services, and sustainability.



**Baojian Liu** received the B.S. degree in 2017 from the School of Earth and Space Science, Peking University, Beijing, China, where he is currently working toward the Ph.D. degree.

His research interests include GNSS-reflectometry remote sensing and cryosphere remote sensing.



**Yang Hong** received the B.S. and M.S. degrees from Peking University, Beijing, China, in 1996 and 1999, respectively, and the Ph.D. degree from The University of Arizona, Tucson, AZ, USA, in 2003.

He is currently a Professor with the Institute of Remote Sensing and GIS, Peking University. His research interests include radar/satellite remote sensing of water cycle, hydrology and water resource, hydrometeorology, and hydroclimatology.