An Automatic Method for Water Extraction From High Spatial Resolution GF-1 Imagery Based On A Deep Learning Algorithm

Jing Kang, Tianyu Ci, Anrong Dang and Yongming Wang

ABSTRACT

Traditional remote sensing methods for water extraction generally rely on complex spectral analysis and subjective thresholds, which do not satisfy the requirements for automated high-accuracy water detection from high spatial resolution images. In this paper, a novel framework of Convolutional Neural Networks (CNNs) is proposed for automatic water extraction based on Densely Connected Convolutional Networks (DenseNet) and a U-net structure. The proposed CNN model is trained on 1300 GF-1 images and tested on another 50 GF-1 images. Six of the GF-1 images are used for experiments to demonstrate the accuracy of the results and compare them with those from NDWI and supervised classification methods. The results show that our method has better computation efficiency and classification accuracy. The new framework has remarkable advantages for automatic water extraction from high-resolution images, especially in maintaining the integrity of small water bodies and in overcoming interference from clouds, shadows in built-up-areas, and mountainous areas with snow cover.

INDEX TERMS
Water Extraction, Remote Sensing, CNNs, GF-1, Densenet, U-net.

INTRODUCTION

Extracting water bodies on the Earth's surface has always been a fundamental task for remote sensing image analysis. As a key component of the hydrologic cycle, water extent is essential for monitoring extreme events and understanding the mechanisms of climate change. At regional scales, water variations have significant impacts on both ecological systems and human lives. The precise perception of water status is critical for successful quantitative evaluations associated with spatiotemporal monitoring, change detection and flooding warnings. Therefore, it is crucial to efficiently extract the water extent from images to further monitor water dynamics or quantifying water volume[1].

Many traditional remote sensing methods have been developed to distinguish water
from non-water features based on its spectral characteristics. A simple and early single-band threshold method was used to develop water maps based on the substantially lower reflectance of water in the infrared band. The most widely used method for water extraction is water indices that are calculated from two or more bands. The normalized difference water index (NDWI)[2] and modified NDWI (MNDWI)[3] can be considered the first generation of water indices. MNDWI uses the shortwave infrared (SWIR) band to replace the near-infrared (NIR) band in NDWI, which has been widely accepted to be more stable and reliable[4, 5]. Similar studies have used the automated water extraction index (AWEI)[6] and a new WATER INDEX[7]. However, the performance of water index methods can sometimes be affected by the complexity of the images, including urban built-up areas shadows, mountains and snow cover[8]. Therefore, several studies have also used combinations of several indices to extract water. For example, the normalized difference built-up index (NDBI) and the normalized difference snow index (NDSI) are commonly used to improve the accuracy of extracted water bodies. The classic normalized difference vegetation index (NDVI), which is mainly used for vegetation extraction, has also been used to detect water bodies and the extent of flood inundation in several studies[9]. In addition, water can be extracted as a typical land cover when using either supervised or unsupervised classification methods[10]. The optimal thresholds of these traditional methods are usually determined using training samples and manual adjustments that rely on prior knowledge and spectral feature analysis, which is subjective and has a low level of automation. In particular, in studies of detecting water bodies from high-resolution images, including QuickBird[11], IKONOS[12], RapidEye[13], GF-1[14], ZY-3[15] and Worldview-2[16, 17], the presence of shadows and large computational issues are two major defects[18]. In addition, the effect of cloud cover is a widespread issue for optical sensors[19].

In recent years, deep learning (DL) methods have achieved state-of-the-art performance in the interpretation of high-resolution satellite images[20]. Compared with traditional remote sensing methods that were followed by classifiers based on handcrafted features, DL methods emphasize automatic feature learning and high-level feature representation from a huge data set through the organization of multiple layers due to their powerful ability of abstraction. Extracting water bodies using deep learning methods has been studied. Yang et al. developed a new method for water extraction based on spectral characteristics and a BP neural network[21]. As one of the most popular DL techniques, CNNs have been widely applied in scene classification, object detection and image retrieval with remarkable results. Martin Lagkvist et al. presented a CNN-based method for classification, including water, that outperformed the existing classification approaches[22]. Fu et al. explored a fully convolutional network model, which is a variant of CNNs, for high spatial resolution imagery classification[10]. Yu et al. proposed a hybrid CNN and logistic regression (LR) classifier for water body extraction from Landsat imagery[23]. However, CNNs have not yet been applied to extract water from high spatial resolution images.

In this paper, we propose a CNN-based method for automatic water extraction from GF-1 images with high accuracy. DenseNet are combined with the U-net architecture to build a novel framework for water extraction that overcomes cloud and shadow interference and improves the computational efficiency. The water body
training set was established from 1300 GF-1 images. The refined output is readily applied to other GF-1 data for automatically extracting water.

DATA AND METHOD

GF-1 DATA AND TRAINING SAMPLES

The data used in this paper are multispectral images of GF-1 PMS data with a resolution of 8 meters. The GF-1 satellite has accumulated abundant data since it was launched in September 2013. Studies of water extraction from GF-1 data have mostly been performed using traditional remote sensing methods.

![Figure 1. Three typical labeled training samples from the 1300 GF-1 images.](image)

Training a powerful CNN would require a sufficient amount of labeled training samples and appropriate techniques. In this study, 1300 GF-1 images from 2014 to 2015 are employed to build training samples that contain rivers, lakes, sea water and urban surface water. Three samples are used to explain the training strategy. Three typical scene images were taken into consideration, which contained the main non-water interferences, clouds, shadows, urban areas, and mountainous areas, as shown in Figure 1. All of the images were acquired in China.

The landcover types shown in sample a) are mainly rivers, lakes, urban areas and rugged mountainous areas. Therefore, the difficulty of extracting water bodies from this image is in distinguishing water bodies from mountain shadows and urban built-up areas. Sample b) is mainly covered by scattered and different water types as well as urban regions. The difficulties and key point for high accuracy water extraction is differentiating small water bodies and built-up areas. Sample c) contains significant cloud cover, but it does not cover the river areas. The difficulty of the water extraction in this image is to eliminate the influences of clouds and cloud shadows.
FRAMEWORK BASED ON CNNS

A typical CNN architecture consists of an input layer, one or more combinations of convolutional and pooling layers (also called subsampling layers), a fully connected layer and an output layer. The pooling layer, which follows a convolutional layer, usually includes mean-pooling and max-pooling. As CNNs have become increasingly popular in DL approaches, several improvements and new network structures have been developed. Each structure has advantages and limitations. In this paper, a novel framework is proposed based on two deep CNNs: DenseNet[24] and U-net[25].

The U-net architecture was originally designed for biomedical image segmentation[25] and has major computational advantages in that it usually requires very few annotated images and uses a reasonable amount of training time. It modified and extended the fully convolutional network that designed the “upsampling part”, yielding a U-shaped architecture.

![Figure 2. The architecture of the U-Net—DenseNet framework (U-DenseNET) for extracting water based on CNNs.](image)

DenseNet, which was developed in 2016, is a new successful CNN with dense connections[24]. This network is designed so there is a direct connection between any two layers. The input of each layer of DenseNet is the union of the outputs of all of the previous layers, and the feature map learned by each previous layer will also be used as an input and is directly transmitted to all of the layers behind. DenseNet is more efficient than other networks due to its decreased computing capacity per layer of the network and the reuse of features, which greatly decreases the number of parameters.

A sufficient number of features from the datasets affects the accuracy of water extraction when using CNNs. The network and training strategy we developed in
this paper relies on the strong use of data augmentation to use the available annotated water samples (1300 images) from GF-1, as shown in Figure 2. By combining the features of U-net and DenseNet, we proposed a new framework, U-DenseNET, for extracting water from high-resolution GF-1 images. The structure of each block is modified by the dense block of DenseNet as shown in Figure 2(a). As shown in Figure 2(b), the entire architecture was followed by U-net, which consists of multiple such blocks. The parameters of U-DenseNet are adjusted to optimal values by considering both accuracy and running time based on numerous experiments, as shown in Figure 2(a).

**RESULTS AND DISCUSSION**

We demonstrate the application of the new U-DenseNET framework to 50 test samples for water extraction from GF-1 images. The test images include various scenes for extracting water, cloud cover, mountainous areas, urban areas and unavoidable shadows. To confirm the effectiveness of U-DenseNET, we perform experiments using the NDWI method and the supervised classification method for comparison.

For visual interpretation, composite infrared, red and green band images were used to distinguish the water due to its lower reflectance in the infrared band, as shown in Figure 1 and Figure 3. The overall accuracy (OA), Kappa coefficient and F-1 score, which are three commonly used indicators of statistical measures, were employed to assess the accuracy. The corresponding GF-1 images were used as the ground truth images. Six representative images were used to perform experiments to make a valid comparison. Figure 3 shows the experimental results of our method compared with those from the NDWI and supervised classification methods. Each set of examples consists of four images, including the original multispectral GF-1 data and the results of water extraction using the three methods, which are binary vectors.

The results show that water bodies could be extracted using the three methods with generally good overall accuracy. However, the water extraction from these examples using NDWI has several misclassifications due to clouds and cloud shadows (Figure 3(2)), mountainous areas, especially with snow cover (Figure 3(4)), urban land (Figure 3(1)(3)(5)), and non-negligible shadows in each image. Therefore, the NDWI results have lower Kappa coefficients and F-1 scores than the others. The supervised classification method requires a sufficient number of training areas using separate manual selection for each image. Although all of the accuracy indicators are high, it is not applicable for water extraction from large numbers of high-resolution images. The average accuracy values of the 6 images show that our results achieved satisfactory accuracy, as shown in Table 1. The water
features extracted using our method have good accuracy, especially in maintaining the integrity of small water bodies (Fig. 3(1)), and the results are not affected by the usual interference factors that affect the NDWI method. A variety of water types in the images were automatically accurately extracted, including rivers, lakes, sea water and small water bodies.

Figure 3. Experimental results. (a) contains many small branches of river; (b) contains certain clouds cover and cloud shadows; (c) contains urban areas and lake. (d) contains river and mountain with snow cover; (e) contain sea water and little cloud cover; (f) contains small body waters in built-up areas.
CONCLUSIONS

In this paper, we proposed a novel deep learning-based approach for automatically extracting water from high-resolution GF-1 images that modified and improved two classical CNNs. Compared with the traditional remote sensing index method, the accuracy and automation are significantly improved.

By combining the architecture of U-net and the dense connections of DenseNet, we developed the U-DenseNET framework. A total of 1300 high-quality GF-1 images were labeled as training samples with professional interpretation based on experience, which required a large amount of manual work. This framework would also be applicable to other GF-1 images.

This work is a preliminary investigation of using CNNs for water extraction from high-resolution images, which has tremendous potential based on deep learning. In the future, we will apply the high-quality water training samples to other high-resolution images, such as GF-2 and ZY-3 images, to make the technique applicable for water extraction or classification from other types of high-resolution images.

ACKNOWLEDGMENT

This study is funded this research is supported by Natural Science Foundation of Beijing “Research on the Planning Decision Making Supporting Approaches of Healthy City Planning of Beijing Based on the Analysis of Social Sensing Data” (No.8182027), and open fund of Institute for China Sustainable Urbanization, Tsinghua University; “Pre-study on new urban development strategy integrating multi-source big data” (TUCSU-K-17026-01). The GF-1 dataset is supported by China Centre for Resources satellite data and application and State key Laboratory of Geo-information engineering.

Corresponding author: Anrong Dang

REFERENCES

1. C. Huang, Y. Chen, S. Zhang, and J. Wu, "Detecting, Extracting, and Monitoring Surface Water From Space Using Optical Sensors: A Review," Reviews of Geophysics, vol. 56, pp. 333-360, 2018.
2. S. K. Mcefeeters, "The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features," International Journal of Remote Sensing, vol. 17, pp. 1425-1432, 1996.
3. X. U. Han-Qiu, "A Study on Information Extraction of Water Body with the Modified Normalized Difference Water Index (MNDWI)," Journal of Remote Sensing, vol. 9, pp. 589-595, 2005.
4. Y. Chen, B. Wang, C. A. Pollino, S. M. Cuddy, L. E. Merrin, and C. Huang, "Estimate of flood inundation and retention on wetlands using remote sensing and GIS," Ecohydrology, vol. 7, pp. 1412-1420, 2015.
5. A. Mohammadi, J. F. Costelloe, and D. Ryu, "Application of time series of remotely sensed normalized difference water, vegetation and moisture indices in characterizing flood dynamics of large-scale arid zone floodplains," Remote Sensing of Environment, vol. 190, pp. 70-82, 2017.
6. G. L. Feyisa, H. Meilby, R. Fensholt, and S. R. Proud, "Automated Water Extraction Index: A new technique for surface water mapping using Landsat imagery," Remote Sensing of Environment, vol. 140, pp. 23-35, 2014.
7. A. Fisher, N. Flood, and T. Danaher, "Comparing Landsat water index methods for automated water classification in eastern Australia," Remote Sensing of Environment, vol. 175, pp. 167-182, 2016.
8. H. Choi and R. Bindschadler, "Cloud detection in Landsat imagery of ice sheets using shadow matching technique and automatic normalized difference snow index threshold value decision," Remote Sensing of Environment, vol. 91, pp. 237-242, 2004.
9. C. Domenikiotis, Loukas, A., & Dalezios, N. R., "The use of NOAA/AVHRR satellite data for monitoring and assessment of forestfires and flood," Natural Hazards and Earth System Sciences, vol. 3, pp. 115-128, 2003.
10. G. Fu, C. Liu, R. Zhou, T. Sun, and Q. Zhang, "Classification for High Resolution Remote Sensing Imagery Using a Fully Convolutional Network," Remote Sensing, vol. 9, p. 498, 2017.
11. K. Xu, J. Zhang, M. Watanabe, and C. Sun, "Estimating River Discharge from Very-High-Resolution Satellite Data: A Case Study in the Yangtze River, China," Hydrological Processes, vol. 18, pp. 1927-1939, 2004.
12. R. C. Sidle, A. D. Ziegler, and J. B. Vogler, "Contemporary changes in open water surface area of Lake Inle, Myanmar," Sustainability Science, vol. 2, pp. 55-65, 2007.
13. S. Klemenjak, B. Waske, S. Valero, and J. Chanussot, "Unsupervised river detection in RapidEye data," in Geoscience & Remote Sensing Symposium, 2012.
14. H. B. Zhang, J. S. Li, N. Xiang, S. Qian, F. F. Zhang, and W. Liang, "An automatic method of monitoring water bodies based on GF-1 data," in Ocean Remote Sensing & Monitoring from Space, 2014.
15. F. Yao, W. Chao, D. Di, J. Luo, Z. Shen, and K. Yang, "High-Resolution Mapping of Urban Surface Water Using ZY-3 Multi-Spectral Imagery," Remote Sensing, vol. 7, pp. 12336-12355, 2015.
16. C. Xie, X. Huang, W. Zeng, and X. Fang, "A novel water index for urban high-resolution eight-band WorldView-2 imagery," International Journal of Digital Earth, vol. 9, pp. 925-941, 2016.
17. R. Malinowski, G. Groom, W. Schwanghart, and G. Heckrath, "Detection and Delineation of Localized Flooding from WorldView-2 Multispectral Data," Remote Sensing, vol. 7, pp. 14853-14875, 2015.
18. K. E. Sawaya, L. G. Olmanson, N. J. Heinert, P. L. Brezonik, and M. E. Bauer, "Extending satellite remote sensing to local scales: land and water resource monitoring using high-resolution imagery," Remote Sensing of Environment, vol. 88, pp. 144-156, 2003.
19. Y. Sheng, Y. Su, and Q. Xiao, "Challenging the cloud-contamination problem in flood monitoring with NOAA/AVHRR imagery," Photogrammetric Engineering & Remote Sensing, vol. 64, pp. 191-198, 1998.
20. X. X. Zhu, D. Tuia, L. Mou, G. S. Xia, L. Zhang, F. Xu, et al., "Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources," IEEE Geoscience & Remote Sensing Magazine, vol. 5, pp. 8-36, 2018.
21. W. L. Yang, M. H. Yang, and Q. I. Hong-Xia, "Water body extracting from TM image based on BPNN," Science of Surveying & Mapping. 2012.
22. M. Längkvist, A. Kiselev, M. Ailerazae, and A. Loutfi, "Classification and Segmentation of Satellite Orthoimagery Using Convolutional Neural Networks," Remote Sensing, vol. 8, pp. 329-, 2016.
23. D. A. Van Dyk and X. Meng, "The Art of Data Augmentation," Journal of Computational and Graphical Statistics, vol. 10, pp. 1-50, 2001.
24. H. Gao, L. Zhuang, and K. Q. Weinberger, "Densely Connected Convolutional Networks," 2016.
25. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in International Conference on Medical Image Computing & Computer-assisted Intervention, 2015.