

Change Detection of Surface Water in Remote Sensing Images Based on Fully Convolutional Network

Authors: Ahram Song, Yeji Kim, and Yongil Kim

Source: Journal of Coastal Research, 91(sp1) : 426-430

Published By: Coastal Education and Research Foundation

URL: <https://doi.org/10.2112/SI91-086.1>

BioOne Complete (complete.BioOne.org) is a full-text database of 200 subscribed and open-access titles in the biological, ecological, and environmental sciences published by nonprofit societies, associations, museums, institutions, and presses.

Your use of this PDF, the BioOne Complete website, and all posted and associated content indicates your acceptance of BioOne's Terms of Use, available at www.bioone.org/terms-o-use.

Usage of BioOne Complete content is strictly limited to personal, educational, and non - commercial use. Commercial inquiries or rights and permissions requests should be directed to the individual publisher as copyright holder.

BioOne sees sustainable scholarly publishing as an inherently collaborative enterprise connecting authors, nonprofit publishers, academic institutions, research libraries, and research funders in the common goal of maximizing access to critical research.

Change Detection of Surface Water in Remote Sensing Images Based on Fully Convolutional Network

Ahram Song[†], Yeji Kim^{†‡}, and Yongil Kim^{†*}

[†]Department of Civil Environmental Engineering
Seoul National University
Seoul, Republic of Korea

[‡]Satellite Operation and Application Center
Korea Aerospace Research Institute
Daejeon, Republic of Korea



www.cerf-jcr.org



www.JCRonline.org

ABSTRACT

Song, A.; Kim, Y., and Kim, Y., 2019. Change detection of surface water in remote sensing images based on fully convolutional network. In: Lee, J.L.; Yoon, J.-S.; Cho, W.C.; Muin, M., and Lee, J. (eds.), *The 3rd International Water Safety Symposium. Journal of Coastal Research*, Special Issue No. 91, pp. 426-430. Coconut Creek (Florida), ISSN 0749-0208.

This study presents a new approach based on fully convolutional networks (FCN) to detect changes in surface water. The proposed method can be divided into three steps: (1) training the FCN using color-infrared (CIR) images from the Coastwide Reference Monitoring System (CRMS) dataset with two classes, such as water and land; (2) passing the multitemporal images respectively through the pre-trained FCN and generating a difference image (DI) from score maps of the last prediction layers; and (3) determining optimal threshold values using fuzzy entropy and discriminating between changed and unchanged pixels in the DI. This method has the advantage of effectively learning the spatial and spectral characteristics of water bodies from large remote-sensing datasets, and it would be helpful to analyze and monitor changes in newly obtained images without ground truth. The experimental results obtained using the multitemporal CRMS data demonstrated the effectiveness of this deep-learning approach for detecting changes in remote-sensing images, as compared other traditional methods for change detection.

ADDITIONAL INDEX WORDS: *Deep learning, surface water, change detection, fully convolutional network, CRMS dataset.*

INTRODUCTION

Surface water includes oceans, rivers, lakes, and wetlands. Analysis of the spatial distribution of surface water and the assessment of any changes in surface water are crucial for watershed analysis, flood and drought mapping, and environmental monitoring (Rokni *et al.*, 2015). Because remote-sensing data have various spatial, spectral, and temporal resolutions and can be obtained from various sensors, such as satellites, unmanned aerial vehicles, and piloted airborne platforms, they have been widely used in recent years for surface change detection (Rokni *et al.*, 2015; Xu, Lin, and Meng, 2017).

Change detection is an important technique for many remote-sensing applications, and a large number of methodologies for the change detection using remote-sensing images have been developed. In general, there are two major categories for change detection. The first one is a post-classification method that separately classifies different-time-phase images and compares the classification results to discriminate between changed and unchanged pixels. Although this method has the advantage of avoiding radiation effects, it is highly dependent on the accuracy of classification. The second one is a comparative analysis in which a difference image (DI) is constructed between target images, and then the changes are detected. In this case, the DI quality is very important to acquiring accurate change-detection results. The comparative analysis has been considered as the

mainstream change-detection method, which has shown good performances in many studies (Gong *et al.*, 2016).

Regarding surface water change detection, post-classification methods have been generally used to detect changes in water bodies (Rokni *et al.*, 2015). However, in many remote-sensing fields, it is hard to gain the prior knowledge that is required for supervised classification methods. On the other hand, the methods of comparative analysis, various water indices (Li *et al.*, 2018; Rokni *et al.*, 2015) and change vector analysis (CVA) (Sarp and Ozcelik, 2017) have been successfully applied to the detection of changes in surface water. However, because those methods are based on the spectral information of multiple bands, it is difficult to consider the spatial structure of surface water.

Deep learning is the fastest growing trend in big data analysis and has extensively advanced the area of image segmentation. With the development of various sensors, remote-sensing data have been explosively collected, and data-based approaches are needed to effectively manage the large amounts of data, without prior knowledge. However, deep learning algorithms are rarely considered in surface water change detection. In many applications, renowned large datasets that consist of everyday pictures—such as ImageNet and PASCAL VOC2012—have been successfully used for the classification of remote-sensing data (Audebert, Le, and Lefèvre, 2016). However, there are some differences between the everyday pictures and remote-sensing data. First, everyday pictures only have red, green, and blue spectral information, whereas many remote-sensing data generally contain near-infrared (NIR) spectral information as well as red, green, and blue spectral information. In particular, the NIR region is useful to discriminate water bodies from other areas because water has the characteristics of high absorption in this

DOI: 10.2112/SI91-086.1 received 9 October 2018; accepted in revision 14 December 2018.

*Corresponding author: yik@snu.ac.kr

©Coastal Education and Research Foundation, Inc. 2019

region. Second, each pixel of remote-sensing data has semantic meaning, whereas datasets containing general digital photos have meaningless backgrounds with a few foreground objects of interest (Long, Shelhamer, and Darrell, 2015). Therefore, weights are trained with a remote-sensing dataset instead of everyday pictures, in order to effectively classify surface water in remote-sensing data with fully convolutional network (FCN).

In this paper, a new approach based on FCN is introduced for the detection of surface water changes in multitemporal remote-sensing images. FCN is state of the art in semantic segmentation, and they accept arbitrary sizes as input and retain spatial information by replacing the fully connected layers with convolutional layers.

The proposed method can be divided into three steps: (1) training parameters of the proposed FCN using color-infrared (CIR) images of the Coastwide Reference Monitoring System (CRMS) dataset; (2) generating a DI from the score map by passing the target images through the pre-trained FCN; and (3) calculating the optimal threshold using fuzzy entropy and detecting changed areas. To demonstrate the effectiveness of the proposed method, we conducted the quantitative analysis of change detection using various evaluation criteria—such as post-classification comparison (PCC) and the Kappa coefficient—with traditional change-detection methods.

THEORETICAL BACKGROUND

FCN Architecture and Mechanism

FCN assigns a label to every pixel by replacing the fully connected layer of convolutional neural networks with convolutional layers containing kernels that cover the entire input region. As a result, FCN can accept arbitrary sizes as input (fully connected layers require fixed sizes) and a classification network can output a heat map (Isikdogan, Bovik, and Passalacqua, 2017).

The architecture of the FCN discussed in this paper is demonstrated in Figure 1. The proposed FCN has an encoder–decoder structure that downsamples inputs and then upsamples them again, and it is mainly composed of three components: a convolution layer, a max-pooling layer, and a deconvolution layer.

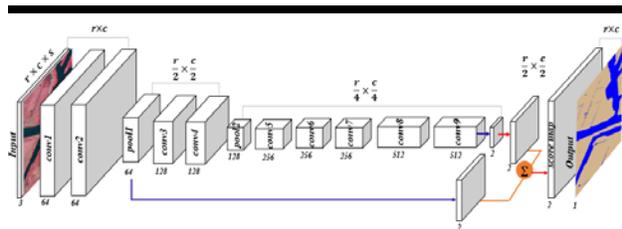


Figure 1. Overall architecture of the FCN. The blue line represents the process of prediction, and the red line represents the process of upsampling.

$r \times c \times s$ represents the size of the feature map. r^l and c^l are the row and column, and s^l is the number of feature maps in the layer l . $x_i^l \in \mathbb{R}^{r^l \times c^l}$ is the i th feature map in layer l , so x_i^{l+1} can be calculated through Equation (1) (Li *et al.*, 2018):

$$x_i^{l+1} = \theta(\sum_{i \in S^l} x_i^l * w_{l+1} + b_{l+1}) \tag{1}$$

where $*$ is the convolution operation and $\theta(\cdot)$ denotes the activation function. In this paper, rectified linear units (ReLU) are used as the nonlinear activation function. w_{l+1} and b_{l+1} are the convolution kernels and bias of the layer $l + 1$, respectively. The convolution layers extract multiscale features. First, a few layers extract general features that are not specific to a particular dataset; then, the high layers extract more abstract and complex features (Huang *et al.*, 2016).

The deconvolution layers enlarge the feature maps to generate dense feature maps that are the same size with input. Deconvolution is executed with bilinear interpolation or learned from the network as follows:

$$x_i^{l+1} = \theta(\sum_{i \in S^l} x_i^l * dw_{l+1} + db_{l+1}) \tag{2}$$

where dw_{l+1} and db_{l+1} are the deconvolution kernel and bias of the layer $l + 1$, respectively.

Because the input image passes several convolution and pooling layers, the spatial information of feature maps is significantly reduced. Typically, the FCN uses skip connection, which integrates low- and high-layer prediction to obtain finer multiscale spatial features (Isikdogan, Bovik, and Passalacqua, 2017). The FCN makes it possible to retain high-level semantic information through combining prediction by the low and high layers. The output number of the last convolution layer (also called the prediction layer) is equal to the number of classes to be discriminated. The feature maps of the prediction layer represent heat maps for all classes, and the pixel in feature map is the score of the corresponding class. In this paper, prediction layers were added after pool1 and conv9. The conv9 prediction was upsampled using the deconvolution layer with pixel stride of 2. Then conv9 upsampling and pool1 prediction were fused. The score map was generated by adding a deconvolution layer after the fusion feature. This process provides local information at a finer scale.

In order to predict class distribution, the softmax function is performed on the score map and gives a distribution for each class. Then, the optimal w_l, b_l are determined through the minimization of the cross entropy loss:

$$\text{Loss} = \frac{-1}{N} \sum_{i=1}^N \sum_{k=1}^K [y_{ik} \log \tilde{y}_{ik} + (1 - y_{ik}) \log(1 - \tilde{y}_{ik})] \tag{3}$$

where N is the number of pixels in a patch, and K is the number of classes. y_{ik} denotes its label and \tilde{y}_{ik} denotes the predicted label.

The network parameters $W = \{w_l, b_l | l \in (1, 2, \dots, L)\}$ are updated using the stochastic gradient descent with momentum. L is the number of layers. In the t th iteration, the network parameters are updated as follows (Fu *et al.*, 2017):

$$W_{t+1} = W_t - \Delta W_{t+1} \tag{4}$$

$$\Delta W_{t+1} = \eta \cdot \left(d_w \cdot W_t + \frac{\partial \text{Loss}}{\partial W_t} \right) + m \cdot \Delta W_t \tag{5}$$

where η is the learning rate, and d_w and m are the weight

decay and momentum, respectively.

METHODS

Proposed Method for Change Detection

In general, image registration between multitemporal images should be implemented before change detection because slight offsets could result significant errors. The procedure of the surface water change-detection method for remote-sensing data is shown in Figure 2. In the first step, the parameters in the FCN were trained on the CRMS dataset to distinguish water bodies from land surface.

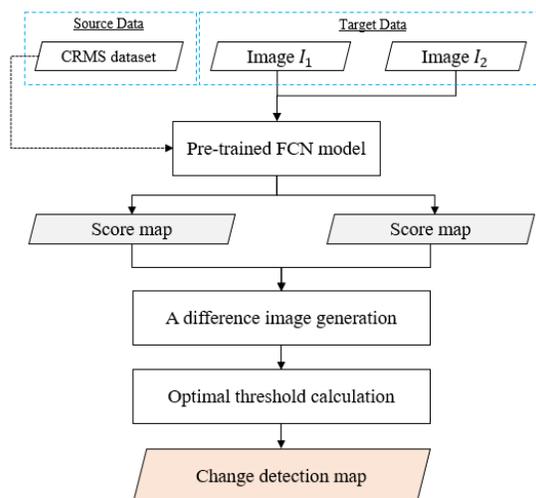


Figure 2. Change detection framework based on the proposed method.

For a wetland restoration effort, CMRS has taken CIR aerial photos at a resolution of 1 m and provided ground-truth maps for about 390 observation sites in Louisianan, USA. The CIR images were obtained from digital orthophoto quarter quadrangles produced by the United States Geological Survey every 3 years. The ground-truth map contained two classes: water and land. In the second step, multitemporal images passed through the pre-trained FCN to generate the DI. Figure 3 shows the process of generating DIs from target images. Each image produced score maps, which were stacked heat maps, for water and land classes because the pre-trained network was classifying two classes. The value of pixels along the channel axis in the score maps represents the score of the corresponding class (Fu *et al.*, 2017). For example, the water score map was equal to the quantity of the water class. Then, the DIs were generated by subtracting each score map.

If the target images were different from those in the CRMS dataset and require new classes (beyond water and land), the pre-trained FCN should be fine-tuned on these target images. Fine-tuning is the process in which parameters of an already trained network are adapted for the new task.

Finally, appropriate threshold values that distinguish the changed and unchanged pixels were determined using multi-level fuzzy entropy (Sarker *et al.*, 2014), and the final output was a binary change map.

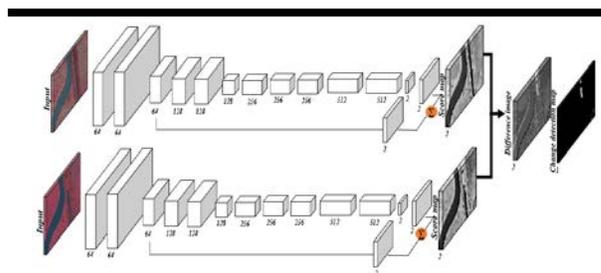


Figure 3. Change-detection architecture of the proposed method.

Quality Assessments of Change Detection

There are various methods for evaluating change-detection accuracy. In this paper, the false-negative (FN) ratio, false-positive (FP) ratio, overall error (OE), overall accuracy (OA), and Kappa coefficient were calculated. FNs refer to undetected changed pixels, and FPs are unchanged pixels that are wrongly detected as changed pixels. The OE and OA are given in the following equations:

$$OE = FN + FP \quad (6)$$

$$OA = \frac{TP+TN}{TP+FP+TN+FN} \quad (7)$$

where TP stands for true positive, and the true-positive ratio is the number of pixels with detected changed regions in both input and ground-truth images. TN stands for true negative, and the true-negative is the number of pixels with detected unchanged regions in both input and ground-truth images. For effective change detection, the $FN:FP$ ratio should be close to 0% and the OA should be 100%. The Kappa coefficient is a measure of classification accuracy based on the difference between the error matrix and chance agreement. Kappa is a metric that compares an observed accuracy with an expected accuracy.

Materials

A total of 356 1000 × 1000-pixel CIR images from the CRMS dataset were used for training. Among them, 284 images were used as training sets and 72 images were used as validation sets. The input images and the corresponding ground-truth images were subset into smaller patches with 300 × 300 pixels in order to consider graphics processing unit (GPU) memory limitations.

The basic parameters for the training network are batch size = 1, base learning rate = 10^{-8} , $m = 0.99$, $d_w = 0.0005$, and the policy of the learning rate adjustment was selected to be “fixed.” The proposed method is implemented on the basis of the Caffe library, which is a representative deep learning framework, and is performed on a CentOS Linux operating system with an NVIDIA GeForce GTX 1080 GPU installed.

RESULTS

Multitemporal images of the CRMS dataset, which show the changes in sites 1 and 2 from 2005 to 2008, were not included in the training and validation sets.

The experimental results for the proposed method were

compared with those of traditional change-detection methods—such as CVA, principal component analysis (PCA), normalized difference water index (NDWI), and the FCN-PCC. Figures 4 and 5 show multitemporal CIR images of sites 1 and 2 in 2005 and 2008 and their score maps. Because vegetation generally has high reflectance in the NIR region, it appears red in CIR images, while water relatively black. The multitemporal images passed through the pre-trained network, and the score maps were produced (Figure 4 and 5c–f). The score maps of water showed relatively higher pixel intensities on river areas than on land areas. On the other hand, the score maps of land have high pixel intensities in areas with bare soil and vegetation. The DI images were generated by subtracting the score maps generated in 2008 and 2005.

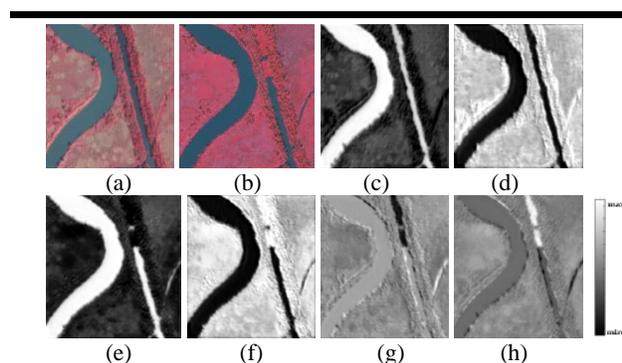


Figure 4. The results of site 1. The CRMS multitemporal CIR images were acquired in (a) 2005 and (b) 2008. The score maps represent (c) water in 2005, (d) land in 2008, (e) water in 2008, and (f) land in 2008. Relatively higher pixel intensities, which is close to white on the score maps. By subtracting each score map, the DIs of the (g) water class and (h) land class were generated.

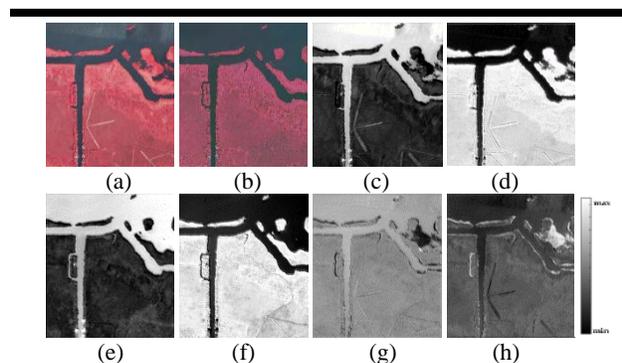


Figure 5. The results of site 2. The CRMS multitemporal CIR images were acquired in (a) 2005 and (b) 2008. The score maps represent (c) water in 2005, (d) land in 2008, (e) water in 2008, and (f) land in 2008. Relatively higher pixel intensities, which is close to white on the score maps. By subtracting each score map, the DIs of the (g) water class and (h) land class were generated.

In each score map, the value indicating the corresponding component has a high positive value, and the opposite component has a high negative value. Therefore, the unchanged pixels have close to zero $\pm\sigma$ and close to grey on the DI images. On the other hand, the changed pixels have relatively extreme positive or negative values and close to white and black. For example, in the DI of water (Figure 4g), the pixels with high negative values were water in 2005 and land in 2008. On the other hand, the pixels with high positive values represent changes from land to water from 2005 to 2008.

The change-detection results were generated by the proposed method, and four comparative methods are presented in Figure 6 and 7. The changed areas were determined on the basis of thresholds determined through fuzzy entropy. Although the CVA generally detected changed areas, there are many spot noises. This noise occurs because CVA and PCA directly calculate the reflectance of multispectral bands; thus, land areas containing other materials with dark reflectance, which are similar to that of water, could be misclassified as water. On the contrary, NDWI and FCN-PCC detected changed areas without spot noise. However, in general, misdetections occurred in the edges of the shoreline.

The proposed method shows optimal results, as compared with traditional methods. This method could detect changed areas clearly, and spot noise is rarely presented. Through training on CIR images, the proposed method has the advantage of being able to learn the spatial structures of surface water, and it can consider spectral information at the NIR region, which is significant for water bodies. A quantitative comparison of the five methods is shown in Tables 1 and 2. In both cases, the proposed method shows the best results, with low FN and FP rates and low OE, and high OA and Kappa coefficients.

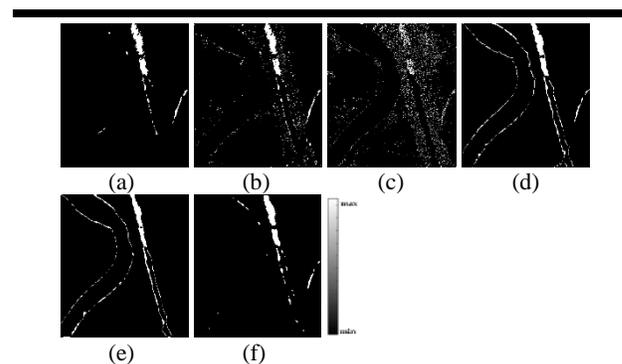


Figure 6. Changes in site 1 from 2005 to 2008, generated using the methods of (a) ground truth, (b) CVA, (c) PCA, (d) NDWI, (e) FCN-PCC, and (f) the proposed method.

CONCLUSIONS

Not only spectral values but also spatial structures are crucial information for surface water change detection. This paper has presented a novel method for surface water change detection based on an FCN. The FCN was trained on the CRMS dataset and classified water bodies, distinguishing them from the background.

Then, the multitemporal images passed through the pre-trained FCN, and the score maps and DIs were generated. The experiments on the two sites demonstrated the effectiveness of the proposed method. Compared with traditional methods, the proposed method exhibits better performance. In conclusion, the proposed method has proven the effectiveness of detecting changes in surface water based on remote sensing images, which were collected from different area of test sites, by considering both the spectral values and spatial structures. Accordingly, the method may be useful in the management of big data and the monitoring of surface waters in other sites around the world.

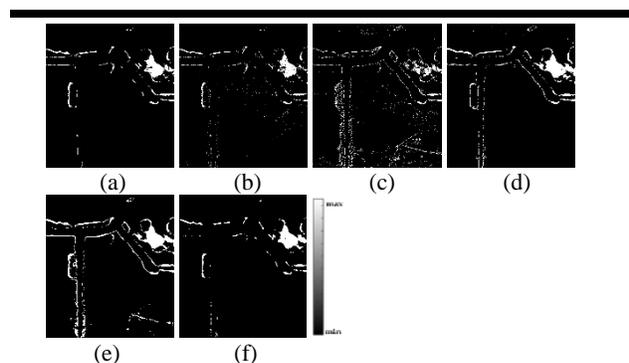


Figure 7. Changes in site 2 from 2005 to 2008, generated using the methods of (a) ground truth, (b) CVA, (c) PCA, (d) NDWI, (e) FCN-PCC, and (f) the proposed method.

Table 1. Change-detection results on site 1.

| Method (%) | CVA | PCA | NDWI | FCN-PCC | Proposed method |
|------------|-------|-------|-------|---------|-----------------|
| FP | 0.127 | 0.680 | 0.410 | 0.428 | 0.266 |
| FN | 0.012 | 0.045 | 0.024 | 0.018 | 0.006 |
| OE | 0.139 | 0.724 | 0.434 | 0.447 | 0.272 |
| OA | 98.51 | 94.16 | 96.78 | 97.28 | 98.81 |
| Kappa | 70.94 | 16.78 | 42.88 | 46.17 | 72.13 |

Table 2. Change-detection results on site 2.

| Method (%) | CVA | PCA | NDWI | FCN-PCC | Proposed method |
|------------|-------|-------|-------|---------|-----------------|
| FP | 0.413 | 0.586 | 0.239 | 0.079 | 0.227 |
| FN | 0.001 | 0.029 | 0.019 | 0.028 | 0.008 |
| OE | 0.414 | 0.615 | 0.258 | 0.107 | 0.235 |
| OA | 98.57 | 95.30 | 97.36 | 97.06 | 98.53 |
| Kappa | 72.13 | 34.10 | 63.92 | 65.66 | 76.56 |

ACKNOWLEDGMENTS

This work was supported by the National Research Foundation of Korea (NRF) funded by the Korean government (MSIT) (grant no. NRF-2016R1A2B4016301) and by the Satellite Information Utilization Center Establishment program funded by Ministry of

Land, Infrastructure and Transport of the Korean government (grant no. 18SIUE-B148326-01).

LITERATURE CITED

- Audebert, N.; Le Saux, B., and Lefèvre, S., 2016. Semantic segmentation of earth observation data using multimodal and multi-scale deep networks. *Proceedings of the Asian Conference on Computer Vision* (Taipei, Taiwan), pp. 180-196.
- Fu, G.; Liu, C.; Zhou, R.; Sun, T., and Zhang, Q., 2017. Classification for high resolution remote sensing imagery using a fully convolutional network. *Remote Sensing*, 9(5), 498p.
- Gong, M.; Zhao, J.; Liu, J.; Miao, Q., and Jiao, L., 2016. Change detection in synthetic aperture radar images based on deep neural networks. *IEEE transactions on neural networks and learning systems*, 27(1), 125-138.
- Huang, C.; Zan, X.; Yang, X., and Zhang, S., 2016. Surface water change detection using change vector analysis. *Proceedings of 2016 Geoscience and Remote Sensing Symposium* (Beijing, China), pp. 2834-2837.
- Isikdogan, F.; Bovik, A. C., and Passalacqua, P., 2017. Surface Water Mapping by Deep Learning. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(11), 4909-4918.
- Jiao, L.; Liang, M.; Chen, H.; Yang, S.; Liu, H., and Cao, X., 2017. Deep fully convolutional network-based spatial distribution prediction for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 55(10), 5585-5599.
- Li, J.; Zhao, X.; Li, Y.; Du, Q.; Xi, B., and Hu, J., 2018. Classification of Hyperspectral Imagery Using a New Fully Convolutional Neural Network. *IEEE Geoscience and Remote Sensing Letters*, 15(2), 292-296.
- Long, J.; Shelhamer, E., and Darrell, T., 2015. Fully convolutional networks for semantic segmentation. *Proceedings of the IEEE conference on computer vision and pattern recognition* (Boston, MA), pp. 3431-3440.
- Rokni, K.; Ahmad, A.; Solaimani, K., and Hazini, S., 2015. A new approach for surface water change detection: Integration of pixel level image fusion and image classification techniques. *International Journal of Applied Earth Observation and Geoinformation*, 34, 226-234.
- Sarkar, S.; Paul, S.; Burman, R.; Das, S., and Chaudhuri, S. S., 2014. A fuzzy entropy based multi-level image thresholding using differential evolution. *Proceedings of International Conference on Swarm, Evolutionary, and Memetic Computing* (Bhubaneswar, India), pp. 386-395.
- Sarp, G. and Ozcelik, M., 2017. Water body extraction and change detection using time series: A case study of Lake Burdur, Turkey. *Journal of Taibah University for Science*, 11(3), 381-391.
- Xu, Y.; Lin, L., and Meng, D., 2017. Learning-Based Sub-Pixel Change Detection Using Coarse Resolution Satellite Imagery. *Remote Sensing*, 9(7), 709.