

# RGB-BASED DEEP SURFACE WATER CONTOUR DETECTION

AbdulRahman Alsamman<sup>a</sup>, Mohammad Syed<sup>a\*</sup>

<sup>a</sup>Department of Electrical and Computer Engineering, University of New Orleans – (aalsamma, mbsyed)@uno.edu

**KEY WORDS:** Surface Water Detection, Deep Learning, Waterbody Detection, Water Contour Detection, Data Collection

## ABSTRACT:

The application of remote monitoring of surface water has focused primarily on the detection of water bodies using expensive multi-spectral IR sensors. However, critical information about surface water bodies, particularly the dynamic behavior, is better derived from water contours. We show that water body detection is inadequate in accurately capturing the contours. Furthermore, we argue that RGB-based detection should be sufficient for accurate water detection. We present a new global dataset of remote sensing images obtained from Sentinel-2 and Landsat-8 missions and contour labeled to assist in this effort. We propose a unique UNet-style contour detection system that utilizes multiscale filters to detect contours accurately. Comparisons between our proposed system, existing water detection, and other segmentation and contour detection systems show the system's effectiveness in detecting water.

## 1. INTRODUCTION

Water is an important ecosystem and a critical resource for life, agriculture, commerce, and industry needs. Additionally, rapid changes in water levels through major weather events threaten life and economic developments in their proximity. Remote sensing can be used to provide accurate, continuous, automatic, and real-time monitoring of surface water bodies. Remote sensing technologies are typically deployed via satellites, aircraft, and drones and provide data at variable temporal resolution (hours/days), spatial resolution (<1m/pixel 1km/pixel), and spectral channels (see (Huang et al. 2018) for a review).

Early remote sensing solutions utilized multi-spectral satellite data to formulate rule-based metrics such as mNDWI (Xu, 2006), WI (Fisher et al. 2016), AEWI (Feyisa et al. 2014), MBWI (Wang et al. 2108), each designed to account for variations in sensor accuracy, water temperature, turbidity, and cloud coverage. Improved water detection was further demonstrated with machine learning techniques involving decision trees (Friedl et al. 1997; Mueller et al. 2016), SVM (Aung et al. 2018), and clustering (Cordeiro et al. 2021) applied to the same multi-spectral satellite data.

Deep learning based fully convolution networks (FCN) (Krizhevsky et al., 2012; Long et al., 2015) have been proven to be extremely successful in applications of object detection and segmentation (see (Hoeser et al, 2020) for a review of techniques and applications to remote sensing). Deep water map (Isikdogan et al. 2017; Isikdogan et al. 2019) is a UNet (Ronneberger et al. 2015) based FCN water body detection system that is used for accurate detection of water bodies from multi-spectral satellite data.

In this paper we focus on the detection of water contours using data in the visible range only (RGB). Water contours are more dynamic and allow for better analysis of changes that are happening at the surface and in the sediment. Our analysis shows that more accurate detection of water bodies doesn't necessarily

lead to the accurate detection of the contour. Contours are more difficult to detect due to the spectral blending that occurs at the boundary. Additionally, there is a need to address the huge imbalance between contour targets and the non-target to improve the detection of contour.

Humans are very skilled at detecting water in the visible spectrum, so we propose that machines can be trained to do the same. IR-based detection of water from sensing data is complicated by the water body type and the surrounding environment. The movement, depth, and temperature of water, as well as shadows from structures and clouds affects the absorption of IR bands. IR sensing technologies are more expensive and require careful calibration and vary in sensitivity, number of spectral bands, and spatial resolution even within the same sensor system. For example, the Sentinel-2 mission provides data at 10m spatial resolution for the visible range, and the Near Infra-red (NIR) range only, the rest of the data is at a lower spatial resolution.

An extensive literature review revealed several datasets (Carroll et al. 2009; Homer et al. 2004; Lehner et al. 2004; Prigent et al. 2012; Verpoorter et al. 2014; Yamazaki et al. 2015) with labeled water bodies. However, these were not useful mainly due to inaccurate contour labeling due to the use of polygons, or aggregate/cumulative water labeling, low resolution of satellite images (>60m/pixels), and difficulty in recognizing the satellite source image used for labeling.

As such we developed our own dataset based on Landsat-8<sup>1</sup> and Sentinel-2<sup>2</sup> missions. The dataset contains of over 1M images for Landsat and over 400K for Sentinel of which 14K have been hand selected for the purpose of fully supervised training. We also present a unique Fully Convolutional Network (FCN) UNet-style network for automatic and accurate detection of water contours from aerial RGB images. Our network outperforms other architectures in contour detection, even though it has fewer parameters compared to the other architectures. We also show that existing waterbody detection systems like Deep water map (Isikdogan et al. 2017; Isikdogan et al. 2019) and water detect (Cordeiro et al. 2021) that use RGB+IR channels even with an F-score > 0.98 do not translate into accurate contour detection.

\* Corresponding Author – (mbsyed@uno.edu)

(1) [www.usgs.gov/landsat-missions](http://www.usgs.gov/landsat-missions)

(2) [sentinel.esa.int/web/sentinel/missions](http://sentinel.esa.int/web/sentinel/missions)

The use of visible bands (RGB) makes it a cost-effective detection system that can be deployed for all remote sensing solutions (fixed camera, drones, etc.). Rest of the paper is organized as follows. Section 2 details the data collection and labeling process. Section 3 has the UNet style architecture used for contour detection. Section 4 contains the results of the proposed system and comparative results to other systems in the literature, and section 5 has the concluding remarks.

## 2. DATA COLLECTION AND LABELLING

Data for both Landsat and Sentinel satellites were acquired from locations around the globe. Both Landsat and Sentinel satellite locations were obtained from (Pekel et al. 2016) and downloaded via Google Earth Engine (GEE) (Gorelick et al., 2017). Additional Sentinel satellite locations were obtained from bluedotwater<sup>3</sup>; data was acquired using Sentinel's API<sup>4</sup>. The dataset was created using the following process.

1. Machine-label the water using a successful waterbody labeling technique.
2. Apply image processing to extract the contour.
3. Split the images into smaller sets.
4. Hand-select the tiles with the best accuracy.
5. Generate meta-data about each tile, including satellite image source and water content percentage.

For step 1, we used the normalized difference water index (NDWI) (Gao, 1996) as it was most successful in producing accurate and continuous contours when applied to various water bodies from around the globe. To extract the contour (step 2), each image was then binarized to show only water, and subtracted from its morphological dilation to yield a contour. While this method generated good contours generally, it also created bad contours within the same image. To improve the yield of good contours, the images were split into  $128 \times 128$  tiles and visually inspected (steps 3 and 4). Tiles with heavy cloud coverage as well as bad contours were removed.

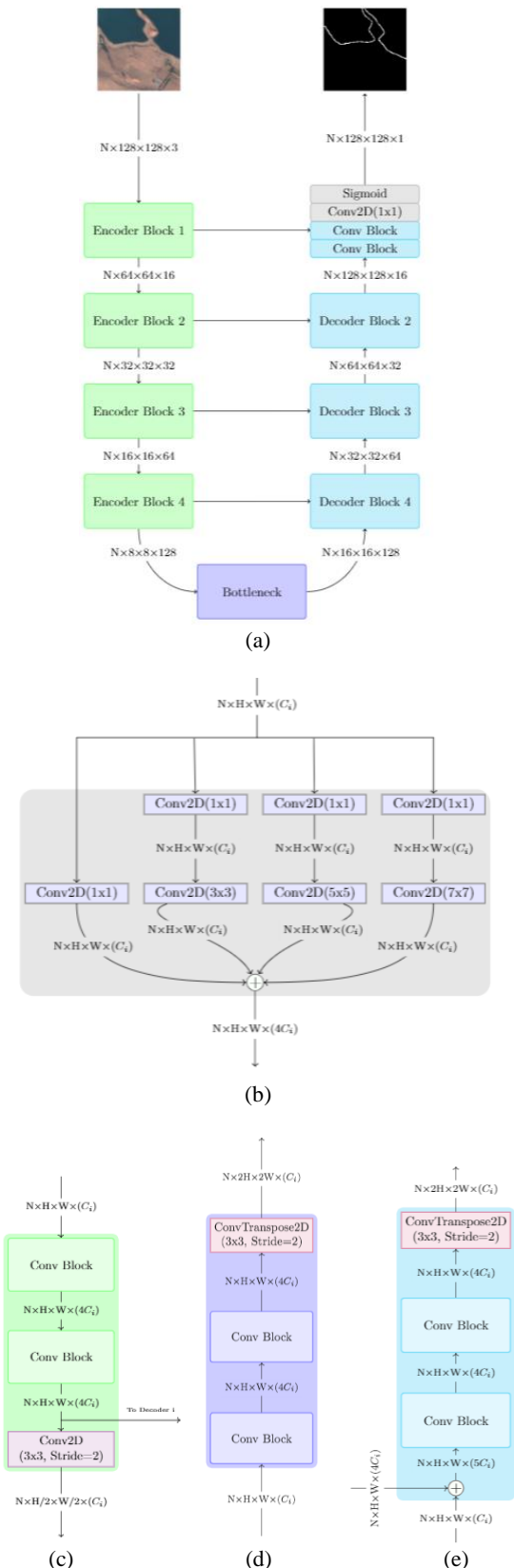
Our current repository<sup>5</sup> contains over 1M tiles for Landsat-8 data and 400K+ tiles for Sentinel-2 of visually unconfirmed data. The hand-selecting process is extremely time consuming with a very low yield. After visually inspecting 100K tiles, we created two datasets each with 7K+ accurately labeled tiles for both Landsat and Sentinel. The datasets were "balanced" to prevent the overabundance of tiles with little to no contours. The data also contains JRC (Pekel et al. 2016) water labels for each Landsat tile for reference. Each tile in the dataset is stored with 6 channels (16-bit raw satellite data), organized in the following order: blue (b1), green (b2), red (b3), NIR (b4), SWIR1 (b5), SWIR2 (b6). Though our work emphasizes RGB detection we included the IR bands in the dataset for others to use freely. Additionally, metadata containing the image source file/satellite, cloud coverage, and approximate water content in each tile is also provided.

## 3. DEEP CONTOUR DETECTION

### 3.1. Preprocessing

The datasets are 16-bit data and require sensor-specific preprocessing to convert them into regular true-color RGB data (TCI). For Landsat, this involved subtracting the min and dividing by the max for each color channel of every image, while for Sentinel, the RGB channels were clipped at 3558.

(3) [www.blue-dot-observatory.com/](http://www.blue-dot-observatory.com/)  
(4) [github.com/sentinelsat/sentinelsat](https://github.com/sentinelsat/sentinelsat)



**Figure 1.** UNET Architecture, (a) full architecture and (b) encoder, (c) decoder, (d) bottleneck, and (e) conv blocks. The  $\oplus$  operator represents the channel concatenation operation.

(5) [github.com/mbsyed/Deep-Surface-Water-Contour-Detection](https://github.com/mbsyed/Deep-Surface-Water-Contour-Detection)

Model	Data	Waterbody Detection			Contour Detection		
		Precision	Recall	F-score	Precision	Recall	F-score
Pre-trained DWM	Landsat (RGB+IR)	0.9929	0.9980	0.9955	0.6757	0.6508	0.6631
	Sentinel (RGB+IR)	0.9831	0.9879	0.9855	0.7416	0.6298	0.6811
WaterDetect	Landsat (RGB+IR)	0.9976	0.9677	0.9824	0.5734	0.6353	0.6028
	Sentinel (RGB+IR)	0.9652	0.9539	0.9764	0.6471	0.4745	0.6436

**Table 1.** Contour detection as a function of waterbody detection.

Model	Time (per epoch.)	Param (M)	Landsat (RGB)		Sentinel (RGB)	
			F-score	AP	F-score	AP
DeepLabV3+	21.4	26.68	0.4801	0.3616	0.5739	0.5125
Unet-Resnet	23.5	32.52	0.6450	0.6391	0.6683	0.7068
PAN	20.7	24.26	0.4483	0.3376	0.5297	0.4677
Unet++	60.7	48.99	0.6579	0.6559	0.6634	0.7059
DWM	17	37.21	0.5529	0.4797	0.6237	0.5806
WaterDetect	N/A	N/A	0.6028	N/A	0.6436	N/A
Unet (Ours)	44	22.85	0.6920	0.7106	0.7379	0.7980

**Table 2.** Comparison of UNet Multiscale(Ours) against the state of the art models.

### 3.2. Architecture

Our proposed UNet-based water contour detector model with unique specifications can be seen in Figure 1. Only RGB channels are used as input. The encoder/decoder layers rely on multiscale convolution layers using sizes of 1x1, 3x3, 5x5, and 7x7 filters. The 1x1 filters are useful in controlling the size of the parameters and help provide weights for each channel. The other size filters are successful in capturing edges at various scales. Batch-normalization (BN) with each convolution was also found to improve contour detection. Also unique to our design is the use of strided convolution for down-sampling in the encoder network and transpose convolution for up-sampling in the decoder network. While this adds parameters to the model, it barely affects processing time and improves the contour detection compared to (un)pooling. "Skip" connections between corresponding encoder and decoder blocks are a general attribute of UNet systems that have been shown to improve training and provide better localization in the output. A sigmoid output is used to classify each pixel output as a contour or non-contour (i.e., 1 or 0).

### 3.3. Loss Function

Binary cross entropy (BCE) was used to capture the pixel-level loss in the image. Intersection over union (IoU) and Dice losses were used to capture the object-related loss.

Due to the significant imbalance between contour and non-contour pixels, the pixels are first weighted by the ratio of non-water pixels to the number of pixels in a  $n \times n$  border around the labeled contour. The weighing makes errors closest to the contour count more heavily than those outside the  $n \times n$  border. We found a border of  $9 \times 9$  to be optimum.

### 3.4. Training

The training was done separately for Landsat and Sentinel datasets. Adam optimizer was used with an initial learning rate of 0.003; default beta1 and beta2 values were used. Each iteration had a mini-batch of 32 images. All the weights in the network were initialized using Kaiming He's initialization method. No augmentation was applied to the training images.

## 4. RESULTS

Each of the 7K+ Landsat and 7K+ Sentinel datasets were randomly split into 5K training tiles and 2K testing tiles and used for experimentation.

### 4.1. Evaluation Metrics

There are two evaluation metrics that we use. F-score, and the Average Precision (AP).

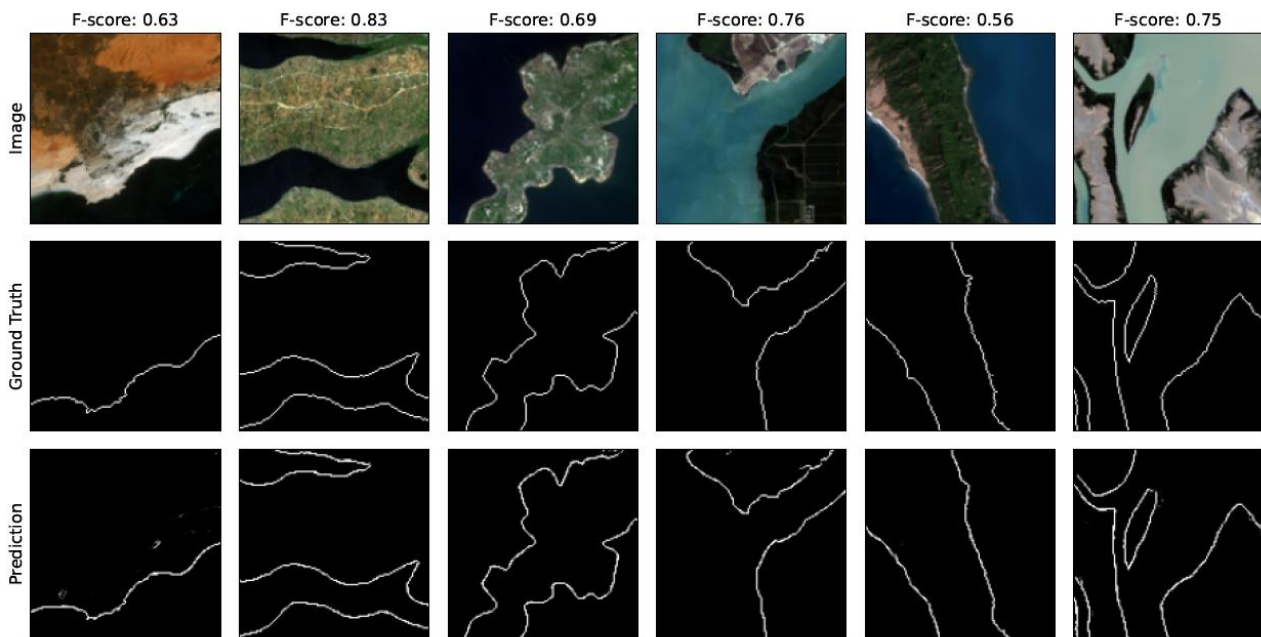
$$Fscore = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (1)$$

where precision and recall are defined as below

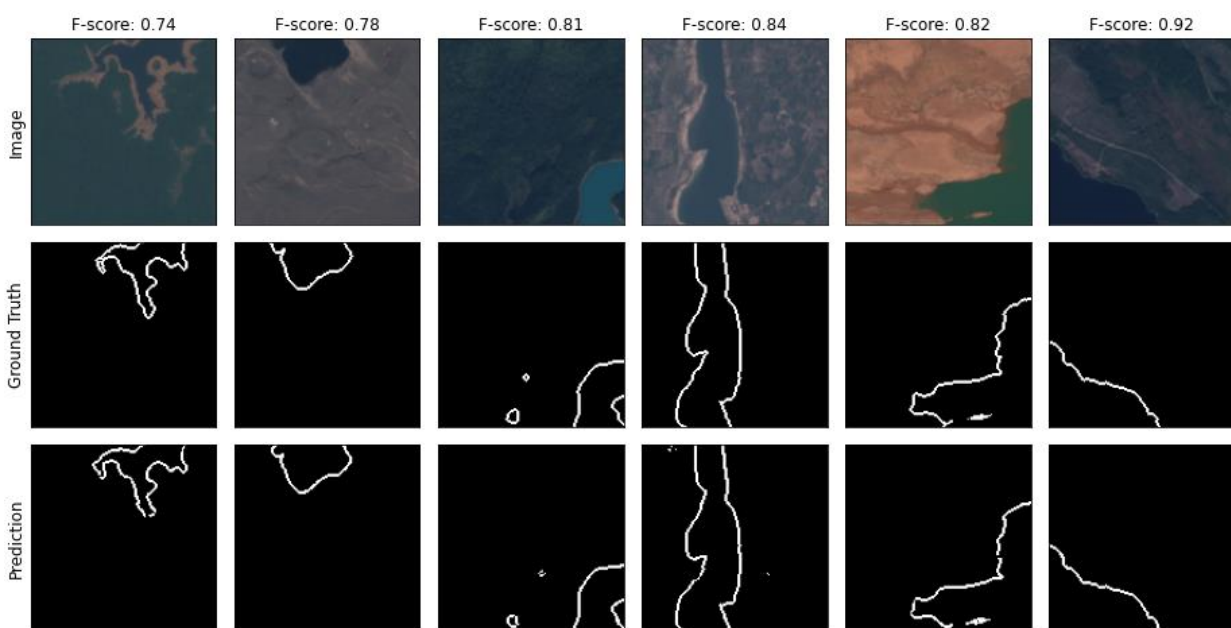
$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

F-score gives us an idea of the model's accuracy for the data, it is the harmonic mean of the precision and recall. The AP is a measure of the accuracy of the prediction of true positives. It is calculated using the precision recall plot's area under the curve (AUC).



**Figure 2.** Results from Landsat test data. The F-score for each image is at the top. The first row contains the original RGB image. The second row has the ground truth for each corresponding image. The bottom row has the predictions from the model trained on Landsat data.



**Figure 3.** Results from Sentinel test data. The F-score for each image is at the top. The first row contains the original RGB image. The second row has the ground truth for each corresponding image. The bottom row has the predictions from the model trained on Sentinel data.

#### 4.2. Numerical Analysis

In Table 1, we demonstrate that successful detection of water bodies does not translate into accurate detection of water contours. The contour detection measures were derived by performing edge detection on the detected water bodies. We compare our model to the pre-trained DeepWaterMap (Isikdogan et al. 2019) (DWM) model, and Waterdetect model. The pre-trained DeepWaterMap is a UNet style deep waterbody detector that is trained on RGB+IR Landsat data. Waterdetect (Cordeiro

et al., 2021) is another waterbody detection tool that applies hierarchical clustering on RGB+IR data. Even with extremely high F-scores (>0.97) both models fail to accurately capture the contour accurately even with IR channels available as input. Compared to this our model has superior performance only using RGB channels for contour detection. This further emphasizes the need for a contour detection system, which will be necessary in water resource management. The table includes performance measures for waterbody detection and contour detection.

Performance metrics for our proposed contour detector are compared to popular image segmentation or waterbody detection systems in the literature, the results are summarized in Table 2. DeeplabV3+ (Chen et al. 2018), UNet-Resnet (Ronneberger et al. 2015), UNet++ (Zhou et al. 2018), and PAN (Li et al. 2018) are popular DL-based segmentation techniques. DWM is a DL waterbody detection model that was retrained on our data specifically for contour detection. Waterdetect (Cordeiro et al. 2021) is also a water body detector but relies on hierarchical clustering of rule-based metrics. As such, it does not have a training phase. The results show that our model outperforms several popular segmentation architectures with the least number of parameters. This can be attributed to the multiscale nature of the convolution block and the weighting of each channel that only keeps the most relevant information. Our system uses fewer parameters, has a faster training time, and is more accurate at detecting water contours than other systems.

### 4.3. Qualitative Analysis

We present some output results of our proposed system in Figure 2 and Figure 3 to show that a low F-score does not necessarily indicate bad output. Visually, the prediction of contours on the RGB images in Figures 2 and 3 are mostly accurate, but the F-scores are low. This can be attributed to a minute misalignment between the prediction and the ground truth. Due to the ground truth being an extremely thin contour, even minor misalignment can cause a low F-score.

## 5. CONCLUSION

Existing and state-of-the-art remote sensing systems rely on water body detection that does not necessitate accurate contour detection. Additionally, there is a need for inexpensive RGB based detection. Due to the lack of accurate water contour data, we presented a technique for collecting and labeling data from Sentinel and Landsat missions. We also proposed a unique UNet-style detector that is very effective in water contour detection. We demonstrated a system that can train faster using fewer parameters by focusing on contours.

### ACKNOWLEDGMENT

This research was made possible through funding from the Navy Research Laboratory.

### REFERENCES

Carroll, M. L., Townshend, J. R., DiMiceli, C. M., Noojipady, P., & Sohlberg, R. A. (2009). A new global raster water mask at 250 m resolution. *International Journal of Digital Earth*, 2(4), 291–308. [doi.org/10.1080/17538940902951401](https://doi.org/10.1080/17538940902951401)

Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018). Encoder-decoder with atrous separable convolution for Semantic Image segmentation. *Computer Vision – ECCV 2018*, 833–851. [doi.org/10.1007/978-3-030-01234-2\\_49](https://doi.org/10.1007/978-3-030-01234-2_49)

Cordeiro, M. C. R., Martinez, J.-M., & Peña-Luque, S. (2021). Automatic water detection from multidimensional hierarchical clustering for sentinel-2 images and a comparison with level 2A processors. *Remote Sensing of Environment*, 253, 112209. [doi.org/10.1016/j.rse.2020.112209](https://doi.org/10.1016/j.rse.2020.112209)

Feyisa, G. L., Meilby, H., Fensholt, R., & Proud, S. R. (2014). Automated Water Extraction Index: A new technique for surface

water mapping using landsat imagery. *Remote Sensing of Environment*, 140, 23–35. [doi.org/10.1016/j.rse.2013.08.029](https://doi.org/10.1016/j.rse.2013.08.029)

Fisher, A., Flood, N., & Danaher, T. (2016). Comparing landsat water index methods for Automated Water Classification in Eastern Australia. *Remote Sensing of Environment*, 175, 167–182. [doi.org/10.1016/j.rse.2015.12.055](https://doi.org/10.1016/j.rse.2015.12.055)

Friedl, M. A., & Brodley, C. E. (1997). Decision Tree Classification of land cover from remotely sensed data. *Remote Sensing of Environment*, 61(3), 399–409. [doi.org/10.1016/s0034-4257\(97\)00049-7](https://doi.org/10.1016/s0034-4257(97)00049-7)

Gao, B.-cai. (1996). NdwI—a normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, 58(3), 257–266. [doi.org/10.1016/s0034-4257\(96\)00067-3](https://doi.org/10.1016/s0034-4257(96)00067-3)

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202, 18–27. [doi.org/10.1016/j.rse.2017.06.031](https://doi.org/10.1016/j.rse.2017.06.031)

Hoese, T., & Kuenzer, C. (2020). Object detection and image segmentation with deep learning on earth observation data: A review-part I: Evolution and recent trends. *Remote Sensing*, 12(10), 1667. [doi.org/10.3390/rs12101667](https://doi.org/10.3390/rs12101667)

Homer, C., Huang, C., Yang, L., Wylie, B., & Coan, M. (2004). Development of a 2001 national land-cover database for the United States. *Photogrammetric Engineering & Remote Sensing*, 70(7), 829–840. [doi.org/10.14358/pers.70.7.829](https://doi.org/10.14358/pers.70.7.829)

Huang, C., Chen, Y., Zhang, S., & Wu, J. (2018). Detecting, extracting, and monitoring surface water from space using optical sensors: A Review. *Reviews of Geophysics*, 56(2), 333–360. [doi.org/10.1029/2018rg000598](https://doi.org/10.1029/2018rg000598)

Isikdogan, F., Bovik, A. C., & 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. [doi.org/10.1109/jstars.2017.2735443](https://doi.org/10.1109/jstars.2017.2735443)

Isikdogan, L. F., Bovik, A., & Passalacqua, P. (2020). Seeing through the clouds with deepwatermap. *IEEE Geoscience and Remote Sensing Letters*, 17(10), 1662–1666. [doi.org/10.1109/lgrs.2019.2953261](https://doi.org/10.1109/lgrs.2019.2953261)

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional Neural Networks. *Communications of the ACM*, 60(6), 84–90. [doi.org/10.1145/3065386](https://doi.org/10.1145/3065386)

Li, Hanchao & Xiong, Pengfei & An, Jie & Wang, Lingxue. (2018). Pyramid Attention Network for Semantic Segmentation. Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). [doi.org/10.1109/cvpr.2015.7298965](https://doi.org/10.1109/cvpr.2015.7298965)

Moh Aung, E. M., & Tint, T. (2018). Ayeyarwady River regions detection and extraction system from Google Earth imagery. 2018 IEEE International Conference on Information Communication and Signal Processing (ICICSP). [doi.org/10.1109/icicsp.2018.8549806](https://doi.org/10.1109/icicsp.2018.8549806)

Mueller, N., Lewis, A., Roberts, D., Ring, S., Melrose, R., Sixsmith, J., Lymburner, L., McIntyre, A., Tan, P., Curnow, S., & Ip, A. (2016). Water observations from space: Mapping Surface Water from 25 years of landsat imagery across Australia. *Remote Sensing of Environment*, 174, 341–352. [doi.org/10.1016/j.rse.2015.11.003](https://doi.org/10.1016/j.rse.2015.11.003)

Pekel, J.-F., Cottam, A., Gorelick, N., & Belward, A. S. (2016). High-resolution mapping of global surface water and its long-term changes. *Nature*, 540(7633), 418–422. [doi.org/10.1038/nature20584](https://doi.org/10.1038/nature20584)

Prigent, C., Papa, F., Aires, F., Jimenez, C., Rossow, W. B., & Matthews, E. (2012). Changes in land surface water dynamics since the 1990s and relation to population pressure. *Geophysical Research Letters*, 39(8). [doi.org/10.1029/2012gl051276](https://doi.org/10.1029/2012gl051276)

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. *Lecture Notes in Computer Science*, 234–241. [doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28)

Verpoorter, C., Kutser, T., Seekell, D. A., & Tranvik, L. J. (2014). A global inventory of lakes based on high-resolution satellite imagery. *Geophysical Research Letters*, 41(18), 6396–6402. [doi.org/10.1002/2014gl060641](https://doi.org/10.1002/2014gl060641)

Wang, X., Xie, S., Zhang, X., Chen, C., Guo, H., Du, J., & Duan, Z. (2018). A Robust Multi-band water index (MBWI) for automated extraction of surface water from Landsat 8 Oli imagery. *International Journal of Applied Earth Observation and Geoinformation*, 68, 73–91. [doi.org/10.1016/j.jag.2018.01.018](https://doi.org/10.1016/j.jag.2018.01.018)

Xu, H. (2006). Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing*, 27(14), 3025–3033. [doi.org/10.1080/01431160600589179](https://doi.org/10.1080/01431160600589179)

Yamazaki, D., Trigg, M. A., & Ikeshima, D. (2015). Development of a global ~90m water body map using multi-temporal landsat images. *Remote Sensing of Environment*, 171, 337–351. [doi.org/10.1016/j.rse.2015.10.014](https://doi.org/10.1016/j.rse.2015.10.014)

Zhou, Z., Rahman Siddiquee, M. M., Tajbakhsh, N., & Liang, J. (2018). UNet++: A nested U-Net Architecture for Medical Image segmentation. *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, 3–11. [doi.org/10.1007/978-3-030-00889-5\\_1](https://doi.org/10.1007/978-3-030-00889-5_1)