HIGH SCHOOL EDITION
Journal of Student Research

Volume 13 Issue 4 (2024)

Impact of Environmental Factors on Water Body Segmentation: A Study of Deep Learning Models

Siddharth Karthikeyan[1] and Vidya Rao[#]

[1]Ardrey Kell High School, USA
[#]Advisor

ABSTRACT

Water body segmentation is a critical task in remote sensing and environmental monitoring, with applications ranging from flood management to natural resource assessment. This task is complex due to many factors such as water color, vegetation, and land type variations. One major and often neglected challenge is the impact of night lighting conditions. In this paper, we present a novel dataset specifically designed to enhance performance under diverse conditions especially with nighttime scenarios. Utilizing nighttime imagery offers advantages such as reduced solar reflection, minimized glare, and clearer detection of water boundaries, which can be particularly useful during periods of haze, cloud cover, or when daytime observations are limited. Our methodology includes the collection, annotation, augmentation, and performing selective slicing on this dataset, followed by the training and evaluation of advanced segmentation models. The results demonstrate significant improvements in model performance, including accuracy, IoU, and precision, across various scenarios, particularly in handling previously challenging conditions. This work helps advance the state of water body segmentation and provides a valuable resource for future research in environmental monitoring and remote sensing.

Introduction

Water body segmentation is a fundamental task in remote sensing and environmental monitoring, playing a crucial role in applications such as flood management, natural resource assessment, and ecological studies [1-3]. Accurate segmentation of water bodies from satellite and aerial imagery enables better analysis of water distribution, quality, and changes over time. However, achieving reliable segmentation in diverse and challenging conditions remains an obstacle [4]. Fig 1 illustrates these challenges, showing examples of segmentation difficulties caused by varying water colors, vegetation along the water's edge, different land types, and night lighting conditions. Each of these factors introduces unique complexities that can hinder accurate detection and delineation of water bodies. Addressing these challenges requires sophisticated methods and specialized datasets.

Figure 1. Examples Illustrating the Challenges in Water Body Segmentation Across Diverse Conditions

The most commonly used dataset for water body segmentation is derived from Sentinel-2 satellite imagery [5-6], which has black and white mask generated by the Normalized Water Difference Index (NWDI) to differentiate water bodies from other features. While NWDI is effective for detecting vegetation, its adaptation for water detection often relies on a higher threshold, which can result in inaccurate or incomplete segmentation, especially for more complex water bodies [1]. The masks may not adequately represent the variability in water color, land types, and vegetation, nor do they sufficiently cover nighttime conditions where reduced illumination and increased noise affect image quality. These limitations in existing datasets contribute to the difficulties faced by current segmentation methods, as they are not trained on data that fully captures the range of conditions encountered in real-world applications. As a result, models trained on such datasets may struggle to generalize effectively, leading to reduced performance in practical scenarios.

The challenge at hand is to develop a diverse dataset that can overcome these limitations and improve performance across diverse conditions. To address this, we created a novel dataset that includes a wider range of environmental conditions, such as varying water colors, different land types, and vegetation patterns, as well as nighttime imagery. By incorporating these diverse scenarios, our dataset aims to provide a better foundation for training and evaluating segmentation models. This approach seeks to enhance model robustness, allowing for more accurate and reliable detection of water bodies in real-world applications, where traditional datasets often fall short.

Methodology

Dataset Creation

In this work, we aimed to create a novel dataset that overcomes the limitations of existing datasets by including diverse environmental conditions such as different water colors, land types, vegetation, and nighttime lighting. To achieve this, we sourced high-resolution satellite images from Google Earth Pro, which provides imagery captured by the Landsat and Copernicus satellites [7]. This dataset offers greater variability in land cover and lighting conditions

HIGH SCHOOL EDITION
Journal of Student Research
Volume 13 Issue 4 (2024)

compared to datasets relying solely on pre-existing segmentation masks. The primary goal was to collect images that represent challenging segmentation scenarios, enabling robust model training.

Data Collection

Google Earth Pro was used as the primary platform for obtaining satellite imagery, using Landsat and Copernicus's multispectral imaging capabilities [8]. We specifically focused on regions with significant environmental diversity to capture a broad range of water body characteristics. These images included a variety of landscapes such as coastal areas, desert regions with intermittent water bodies, dense vegetation zones, and urban areas with infrastructure. Additionally, we collected imagery under different lighting conditions, including nighttime scenarios, to address the often-neglected challenge of night lighting.

Data Annotation

The next step involved generating high-precision segmentation masks for each of the satellite images. To ensure the accuracy and quality of these annotations, we used the Computer Vision Annotation Tool (CVAT) to manually generate masks [9], which is illustrated in Fig 2. CVAT allowed for detailed, pixel-level annotations, ensuring that the boundaries of water bodies were accurately captured, even in cases where water color, land types, or vegetation made segmentation more difficult. Each image was carefully annotated by identifying and labeling the water bodies, which were marked in a binary mask format (water as white and non-water areas as black). Manual annotations were necessary to overcome the limitations of automated mask generation methods like NWDI, which often struggle with edge cases such as complex land formations, shadows, or vegetation overlapping with water bodies.

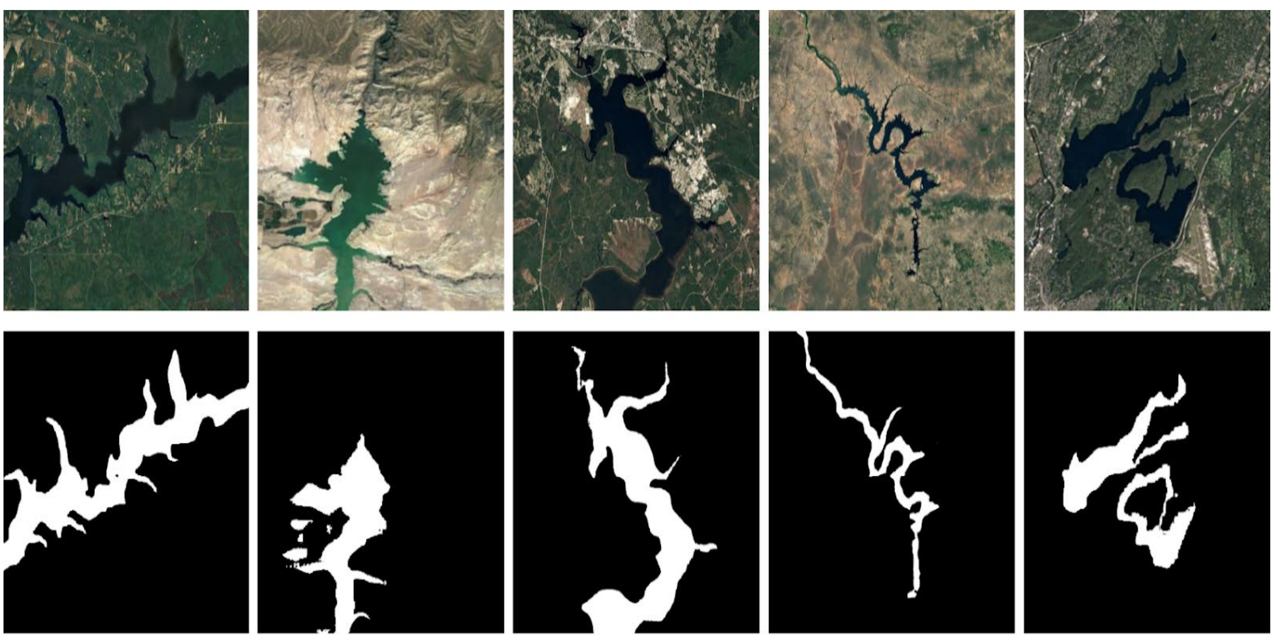

Figure 2. Creating Masks for Images using CVAT

Data Augmentation

To further enhance the dataset, data augmentation techniques were applied [10]. These included transformations such as rotations, scaling, and brightness adjustments as shown in Fig 3. The purpose of these augmentations was to artificially increase the diversity of the dataset, allowing the model to generalize better under various lighting conditions and land formations. For nighttime imagery, we also simulated different noise levels and illumination variations to make the dataset even more comprehensive.

Figure 3. Transforming Images Using Rotations, Flipping, Scaling and Lighting Adjustments

Selective Slicing

Selective slicing is a crucial component of our methodology to enhance the performance of deep learning models for water body segmentation. We divided each high-resolution satellite image into non-overlapping slices of 256x256 pixels, creating a balance between preserving contextual information and adhering to input size constraints. This technique ensures that essential features along water body edges are retained, allowing the model to learn from localized patterns effectively.

For labeling the slices, we utilized the corresponding segmentation masks and applied a water pixel threshold of 10% to classify each slice as containing water or not. Fig 4 illustrates the process of selective slicing applied to high-resolution satellite imagery. Each 256x256 pixel slice is outlined on the original image and the slices are also color-coded to indicate their classification based on the water pixel threshold. This approach not only increases dataset diversity but also enhances the model's ability to generalize across various environmental conditions. Ultimately, our selective slicing strategy boosts both the volume and quality of training data, improving the accuracy and reliability of models in real-world applications.

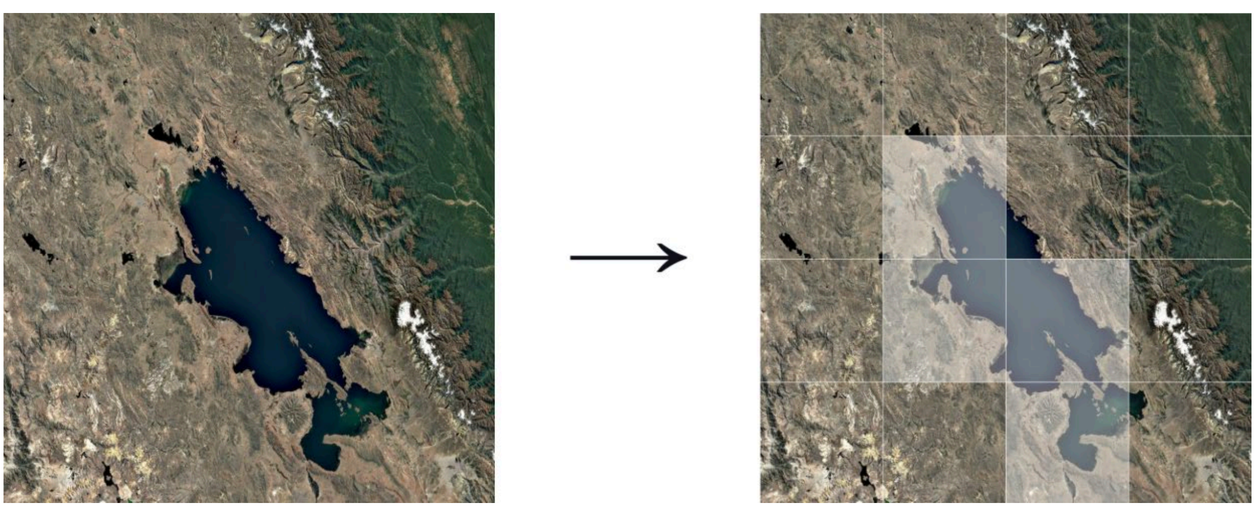

Figure 4. Visualization of the Selective Slicing

Segmentation Models

For the segmentation task, we employed advanced deep learning models, leveraging architectures such as FCN, U-Net, Link-Net and DeepLabV3+ [11-15]. These models were selected for their ability to capture fine-grained spatial details and perform well on pixel-level classification tasks. Modifications were made to these architectures to enhance their performance in detecting water bodies under challenging conditions such as low lighting or areas with overlapping vegetation. The models were trained on the newly created dataset, using a combination of convolutional layers and attention mechanisms to focus on the complex boundaries of water bodies.

Results and Analysis

Dataset

First: Sentinel-2 Dataset

The first dataset was sourced from Kaggle, titled "Images and Masks of Satellite Images of Water Bodies for Image Segmentation." It consists of 2,841 daytime images captured by the Sentinel-2 satellite. The water body masks were automatically generated by calculating the Normalized Water Difference Index (NWDI), which detects water by comparing near-infrared and visible green reflectance. While this approach works in standard conditions, it has limitations as it can create noise when handling complex environments, such as varying water colors, land types, and overlapping vegetation.

A significant challenge with this dataset is the presence of noise in some images [16]. Specifically, there are random images where the content is either completely black or where the water mask is entirely filled. These anomalies can adversely affect the performance of the model by introducing incorrect training examples, which may lead to overfitting or reduced generalization capability.

Second: Custom Google Earth Pro Dataset

To overcome the limitations of the first dataset, we created a custom dataset using high-resolution satellite imagery from Google Earth Pro, utilizing images captured by the Landsat and Copernicus satellite programs. This dataset consists of 110 high-resolution images consisting of both daytime and nighttime imagery providing a wider range of environmental conditions. For this custom dataset, we manually annotated the water body masks using the Computer Vision Annotation Tool (CVAT). This tool allowed us to create precise, pixel-level annotations, ensuring the accurate delineation of water bodies in challenging scenarios. We then applied selective slicing, rotations, scaling, and horizontal and vertical flipping resulting in 3,306 images.

Evaluation Metrics

To quantitatively assess and compare the performance of our segmentation models, we employed four key evaluation metrics: Accuracy, Precision, Recall, and Intersection over Union (IoU). These metrics are fundamental for understanding different aspects of model performance:

Equation 1: Equations of Evaluation Metrics

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$IoU = \frac{TP}{TP + FP + FN}$$

Here, TP, TN, FP, and FN represent True Positives, True Negatives, False Positives, and False Negatives, respectively. Accuracy measures the overall correctness of the model, while Precision indicates how often positive predictions are correct. Recall evaluates the model's ability to detect all actual positives. Intersection over Union (IoU) assesses the spatial overlap between predicted and actual segmentations. IoU was chosen as the primary metric for its ability to accurately evaluate the alignment between predicted and ground truth segmentations.

Results

We compare the performance of various segmentation models using both the existing Sentinel-2 dataset and our newly created dataset. This comparative analysis highlights how the models perform under different conditions, including the challenges posed by night lighting, diverse water colors, vegetation, and land types. By evaluating the models across these datasets, we aim to demonstrate the improvements and advantages provided by our new dataset in addressing previously difficult segmentation scenarios.

Performance on Sentinel-2 Dataset

Table 1 presents the results of various models trained on the "Satellite Images of Water Bodies" dataset sourced from Kaggle. The table includes key performance metrics: accuracy, precision, recall, and Intersection over Union (IoU). Each metric was evaluated on the test dataset, which comprises 10% of the whole dataset, highlighting the effectiveness of each model in segmenting water bodies from satellite images. Among the models, Link-Net achieved the highest accuracy (0.7633) while DeepLabV3+ excelled in precision (0.9642), recall (0.6551) and IoU (0.7365). These results demonstrate the varied strengths of each approach in tackling the segmentation challenge.

Table 1. Performance Comparison of Segmentation Models with Sentinel-2 Dataset

Models	**Metrics**			
	Accuracy	**Precision**	**Recall**	**IoU**
FCN	0.7263	0.9255	0.5803	0.5275
U-Net	0.7050	0.8797	0.5657	0.5297

Link-Net	0.7633	0.9183	0.6365	0.6663
DeepLabsV3+	0.7605	0.9642	0.6551	0.7365

Fig 5 illustrates the performance of various models, trained with Sentinel-2 dataset, when predicting nighttime images of the same water body. It is evident that the FCN and U-Net models perform poorly, struggling to accurately capture the features in the nighttime context. In contrast, Link-Net shows improved performance compared to FCN and U-Net, while DeepLabV3+ stands out as the most effective model, demonstrating superior accuracy in predicting nighttime conditions.

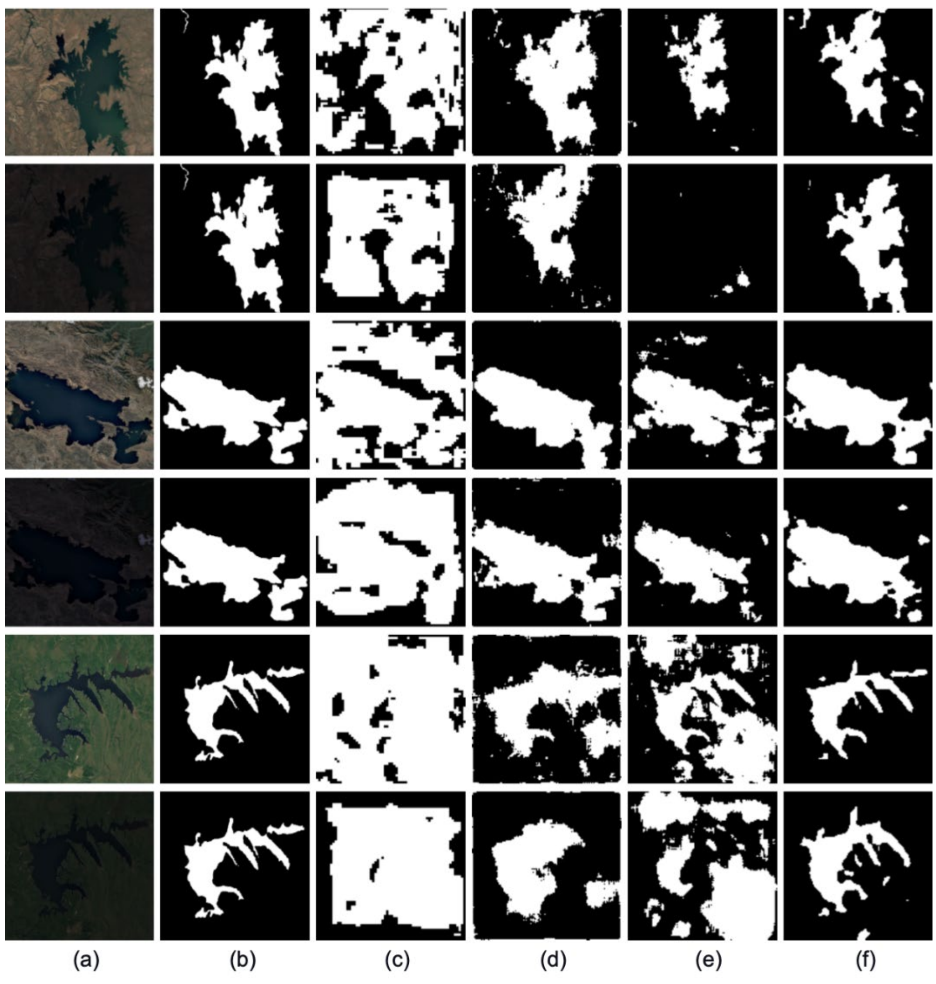

Figure 5. Performance Comparison of Segmentation Models with Sentinel-2 Dataset. (a) Images, (b) Ground-truth, (c) FCN, (d) U-Net, (e) Link-Net (f) DeepLabsV3+

Performance on Our Dataset

Table 2 displays the results of models trained on our custom dataset, which utilizes Landsat and Copernicus satellite imagery sourced from Google Earth Pro. The table summarizes key performance metrics: batch size, accuracy, precision, recall, and Intersection over Union (IoU). All models were evaluated with a batch size of 16. Notably, DeepLabV3+ achieved the highest accuracy (0.9916) and IoU (0.9733), while Link-Net closely followed with impressive precision (0.9874) and recall (0.9871). These results underscore the strong performance of each model in the context of water body segmentation, showcasing their effectiveness in leveraging high-resolution satellite imagery.

Table 2. Performance Comparison of Segmentation Models with Our Dataset

Models	**Metrics**			
	Accuracy	**Precision**	**Recall**	**IoU**
FCN	0.9746	0.9700	0.9618	0.9078
U-Net	0.9830	0.9712	0.9835	0.9419
Link-Net	0.9905	0.9874	0.9871	0.9682
DeepLabsV3+	0.9916	0.9875	0.9901	0.9733

Fig 6 illustrates the performance of various models, trained with our custom dataset after implementing the selective slicing technique. The consistent and precise results across the board highlight the effectiveness of this dataset and approach of training. By using this technique, all models achieve significant improvements, further emphasizing the importance of data processing strategies in enhancing model performance.

Figure 6. Performance Comparison of Segmentation Models with Our Dataset. (a) Images, (b) Ground-truth, (c) FCN, (d) U-Net, (e) Link-Net (f) DeepLabsV3+

Discussion

This study highlights the superior performance of segmentation models when trained on our newly developed dataset, which contains nighttime imagery and diverse environmental conditions. Compared to the traditional Sentinel-2 dataset, our dataset led to significantly higher accuracy, precision, recall, and IoU scores, as seen in table 1 and table 2. Models like FCN and U-Net, while already struggling with daytime images in the traditional dataset, showed a particularly steep decline in performance with nighttime images, as demonstrated in Fig 5. In contrast, DeepLabV3+ consistently excelled across both datasets, with the inclusion of nighttime imagery making a clear difference in its robustness. Link-Net, while performing moderately well, still trailed behind DeepLabV3+ in complex scenarios. The Sentinel-2 dataset's lack of variation in lighting conditions was evident in its lower performance metrics.

The application of the selective slicing and merging technique, detailed in Fig 6, significantly enhanced all models' precision and IoU, proving especially valuable for nighttime segmentation. The ability to break down large images into smaller slices allowed the models to focus on specific details, improving overall segmentation accuracy. Despite this, there is room for improvement: expanding the dataset to include more balanced day and night images, as well as optimizing models like Link-Net, could further increase robustness in real-world, unpredictable environments.

Conclusion

This study presents significant advancements in water body segmentation by introducing a novel dataset that enhances model performance across a range of challenging scenarios. By comparing models trained on both the existing Sentinel-2 dataset and our new dataset, we have demonstrated the advantages of incorporating diverse and comprehensive data. The improved segmentation accuracy observed with our dataset underscores its potential for better handling real-world complexities, such as night lighting and varying environmental conditions.

In summary, our work highlights the importance of dataset quality and diversity in enhancing segmentation models. Future research should focus on further refining data collection and augmentation techniques to address additional challenges and continue advancing the state of water body segmentation. This research not only improves current methodologies but also provides a foundation for developing more robust and accurate models for environmental monitoring.

Limitations

The limitations of this study are primarily related to the manual annotation process for the 3,306 images was time-intensive, which may restrict scalability and hinder the model's ability to generalize to larger or more diverse datasets. Expanding the dataset to include other satellite programs and a wider range of environmental conditions could enhance the model’s robustness. Future research should focus on optimizing the balance between day and night images, addressing extreme environmental conditions, and exploring semi-automated annotation methods to facilitate the creation of larger, high-quality datasets, thereby increasing the model's applicability across different sensors and conditions.

Acknowledgments

We would like to express our gratitude to Ardrey Kell High School for their support and resources throughout this research project.

References

Miao, Z., Fu, K., Sun, H., Sun, X., & Yan, M. (2018). Automatic water-body segmentation from high-resolution satellite images via deep networks. IEEE geoscience and remote sensing letters, 15(4), 602-606. https://doi.org/10.1109/LGRS.2018.2794545

Wang, Z., Gao, X., Zhang, Y., & Zhao, G. (2020). MSLWENet: A novel deep learning network for lake water body extraction of Google remote sensing images. Remote Sensing, 12(24), 4140. https://doi.org/10.3390/rs12244140

Lv, S., Meng, L., Edwing, D., Xue, S., Geng, X., & Yan, X. H. (2022). High-performance segmentation for flood mapping of hisea-1 sar remote sensing images. Remote Sensing, 14(21), 5504. https://doi.org/10.3390/rs14215504

Zhang, Z., Lu, M., Ji, S., Yu, H., & Nie, C. (2021). Rich CNN features for water-body segmentation from very high resolution aerial and satellite imagery. Remote Sensing, 13(10), 1912. https://doi.org/10.3390/rs13101912

Yuan, K., Zhuang, X., Schaefer, G., Feng, J., Guan, L., & Fang, H. (2021). Deep-learning-based multispectral satellite image segmentation for water body detection. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 14, 7422-7434. https://doi.org/10.1109/JSTARS.2021.3098678

Wieland, M., Fichtner, F., Martinis, S., Groth, S., Krullikowski, C., Plank, S., & Motagh, M. (2023). S1S2-Water: A global dataset for semantic segmentation of water bodies from Sentinel-1 and Sentinel-2 satellite images. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. https://doi.org/10.1109/JSTARS.2023.3333969

Yang, X., Chen, Y., & Wang, J. (2020). Combined use of Sentinel-2 and Landsat 8 to monitor water surface area dynamics using Google Earth Engine. Remote Sensing Letters, 11(7), 687-696. https://doi.org/10.1080/2150704X.2020.1757780

Li, K., Wang, J., Cheng, W., Wang, Y., Zhou, Y., & Altansukh, O. (2022). Deep learning empowers the Google Earth Engine for automated water extraction in the Lake Baikal Basin. International Journal of Applied Earth Observation and Geoinformation, 112, 102928 https://doi.org/10.1016/j.jag.2022.102928

Chatterjee, R., Chatterjee, A., & Islam, S. H. (2022). Deep learning techniques for observing the impact of the global warming from satellite images of water-bodies. Multimedia Tools and Applications, 81(5), 6115-6130. https://doi.org/10.1007/s11042-021-11811-1

Pai, M. M., Mehrotra, V., Verma, U., & Pai, R. M. (2020). Improved semantic segmentation of water bodies and land in SAR images using generative adversarial networks. International Journal of Semantic Computing, 14(01), 55-69. https://doi.org/10.1142/S1793351X20400036

Li, L., Yan, Z., Shen, Q., Cheng, G., Gao, L., & Zhang, B. (2019). Water body extraction from very high spatial resolution remote sensing data based on fully convolutional networks. Remote Sensing, 11(10), 1162. https://doi.org/10.3390/rs11101162

Ch, A., Ch, R., Gadamsetty, S., Iwendi, C., Gadekallu, T. R., & Dhaou, I. B. (2022). ECDSA-based water bodies prediction from satellite images with UNet. Water, 14(14), 2234. https://doi.org/10.3390/w14142234

Tambe, R. G., Talbar, S. N., & Chavan, S. S. (2021). Deep multi-feature learning architecture for water body segmentation from satellite images. Journal of Visual Communication and Image Representation, 77, 103141. https://doi.org/10.1016/j.jvcir.2021.103141

Dang, B., & Li, Y. (2021). MSResNet: Multiscale residual network via self-supervised learning for water-body detection in remote sensing imagery. Remote Sensing, 13(16), 3122. https://doi.org/10.3390/rs13163122

Li, Z., Wang, R., Zhang, W., Hu, F., & Meng, L. (2019). Multiscale features supported DeepLabV3+ optimization scheme for accurate water semantic segmentation. Ieee Access, 7, 155787-155804. https://doi.org/10.1109/ACCESS.2019.2949635

Maiti, A., Oude Elberink, S. J., & Vosselman, G. (2022). Effect of label noise in semantic segmentation of high resolution aerial images and height data. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2, 275-282. https://doi.org/10.5194/isprs-annals-V-2-2022-275-2022