A Study on Improving Water Body Detection Accuracy in CAS500-1 Satellite Imagery Using Deep Learning

SeoJin Kong[1], Won-Woo Seo[2], Sooahm Rhee[3*]

[1]*Image Eng. Research Center, 3DLabs Co. Ltd, Seoul, South Korea, ksj8973@3dlabs.co.kr*
[2]*Image Eng. Research Center, 3DLabs Co. Ltd, Seoul, South Korea, winter0233@3dlabs.co.kr*
[3*]*Image Eng. Research Center, 3DLabs Co. Ltd, Seoul, South Korea, ahmkun@3dlabs.co.kr*

Abstract: Efficient management and continuous monitoring of water resources are essential for agriculture, urban, disaster response and other sectors. Accordingly, the demand for automated water body detection techniques is steadily increasing. CAS500-1, a high-resolution satellite developed in Korea, provides Analysis Ready Data (ARD), including surface reflectance images and additional information such as water body and cloud masks. The water body mask is currently created through manual digitizing, which is time-consuming, costly, and limited in reflecting temporal changes. This study aims to automate and improve the accuracy of water body detection in CAS500-1 satellite imagery through deep learning models. To this end, the deep learning model U-Net, designed for semantic segmentation, was applied and its performance was evaluated. And the resulting water body masks were evaluated against existing labeled data. The results showed high performance: U-Net achieved an Accuracy of 0.92 and Precision of 0.94 and Recall of 0.89 and F1-score of 0.92. Notably, the models were able to distinguish small objects such as ships and bridges from water bodies, even when such details were absent in the label data. This study overcomes the limitations of the existing manual method and enables automated detection of water body using high-resolution satellite imagery. It facilitates continuous and precise monitoring of surface water areas and is expected to contribute meaningfully to decision-making processes related to water resource utilization.

Keywords: ARD, CAS500-1, Deep Learning, Semantic Segmentation, Water body detection

1. Introduction

Efficient management and continuous monitoring of water resources are essential in various fields such as agriculture, urban planning, and disaster response. Accordingly, the demand for automated water body detection technologies has been steadily increasing. Kim et al. (2024) applied PlanetScope imagery and the High-Resolution Network (HRNet) model to detect water bodies in domestic dams and rivers and evaluated the applicability of monitoring surface water areas. An and Rui (2022) addressed the accuracy assessment and improvement methods of high-precision water body extraction using high-resolution satellite data. These studies highlight the importance of water body detection and available water resource management utilizing satellite imagery

CAS500-1, a high-resolution satellite developed in Korea, provides Analysis Ready Data (ARD), which includes surface reflectance imagery along with additional information such as water body and cloud masks. The water body mask included in ARD is currently generated through manual digitizing, which requires significant time and cost. Moreover, since it is not updated, it has limitations in reflecting temporal changes.

A traditional technique for water body detection is the Normalized Difference Water Index (NDWI). However, since NDWI is a single index-based method, it has limitations in distinguishing objects that share spectral characteristics similar to water. Therefore, this study aims to automate and improve the accuracy of water body detection in CAS500-1 satellite imagery using deep learning models. To this end, the deep learning model U-Net, designed for semantic segmentation, was applied and its performance was evaluated.

2. Methodology

In this study, the Red, Green, Blue, and NIR single-band images and the water mask from the CAS500-1 satellite ARD data were used. The CAS500-1 satellite, operated by the National Geographic Information Institute (NGII) of Korea. It was launched in 2021 and provides a spatial resolution of 0.5 m for panchromatic imagery and 2 m for multispectral imagery

A total of 110 ARD datasets were used, and datasets containing a large proportion of irrelevant background pixels were excluded. The temporal range of the data was from 2022 to 2024, and the spatial range covered the entire Korean Peninsula, including coastal areas, islands, rivers, and lakes. The collected data were resized to 1024×1024 and normalized to a range of 0 to 255. To mitigate the influence of outlier values, normalization was performed using the lower 2% and upper 98% of the pixel value distribution. Subsequently, the images were tiled into 512×512 patches, which are suitable for AI model input, and then split into train, test, and validation datasets.

The model adopted in this study was U-Net, which was originally proposed for pixel-level classification of medical images through semantic segmentation. Semantic segmentation assigns class labels to each pixel. U-Net is based on the Fully Convolutional Networks (FCN) model, with key differences from conventional FCNs. During the image reconstruction process, it incorporates two 3×3 convolution operations and utilizes feature maps of the same dimensions when restoring spatial information. Through these characteristics, U-Net enables more accurate image segmentation and achieves faster network performance. Hyperparameters are parameters set by the user prior to training, and model performance can vary depending on these settings. In this study, hyperparameters were optimized using the grid search method, focusing on batch size, epochs, and learning rate. The values were set as batch size 2, epoch 100, and learning rate 0.001.

The resulting water body masks were quantitatively and qualitatively evaluated against existing labeled data. In addition, they were compared with the results of the traditional water body detection method, the Normalized Difference Water Index (NDWI). Otsu's algorithm was applied to automatically determine an optimal threshold for each image, effectively separating water and non-water pixels. Binary water masks were subsequently generated by classifying pixels with NDWI values above the computed threshold as water. In general, NDWI is calculated using the NIR and SWIR bands. However, since the CAS500-1 satellite does not provide a SWIR band, this study used the Green and NIR bands instead.

$$NDWI = \frac{Green - NIR}{Green + NIR}$$

3. Results/Findings

Model performance was evaluated using Accuracy, Precision, Recall, and F1-Score. Accuracy represents the overall proportion of correct predictions, while Precision and Recall assess the correctness and completeness of positive predictions, respectively. F1-Score, the harmonic mean of Precision and Recall, provides a balanced metric reflecting both aspects.

As a result of model training, the model achieved high performance with an Accuracy of 0.92, Precision of 0.94, Recall of 0.89, and an F1-score of 0.92. It was also confirmed that the model effectively detected water body areas that were not reflected in the True mask over time. Ships and small island areas were clearly distinguished from water bodies and even in cases where the original satellite imagery exhibited color distortions, as shown in Figure 1 (b), water bodies were reliably detected.

.

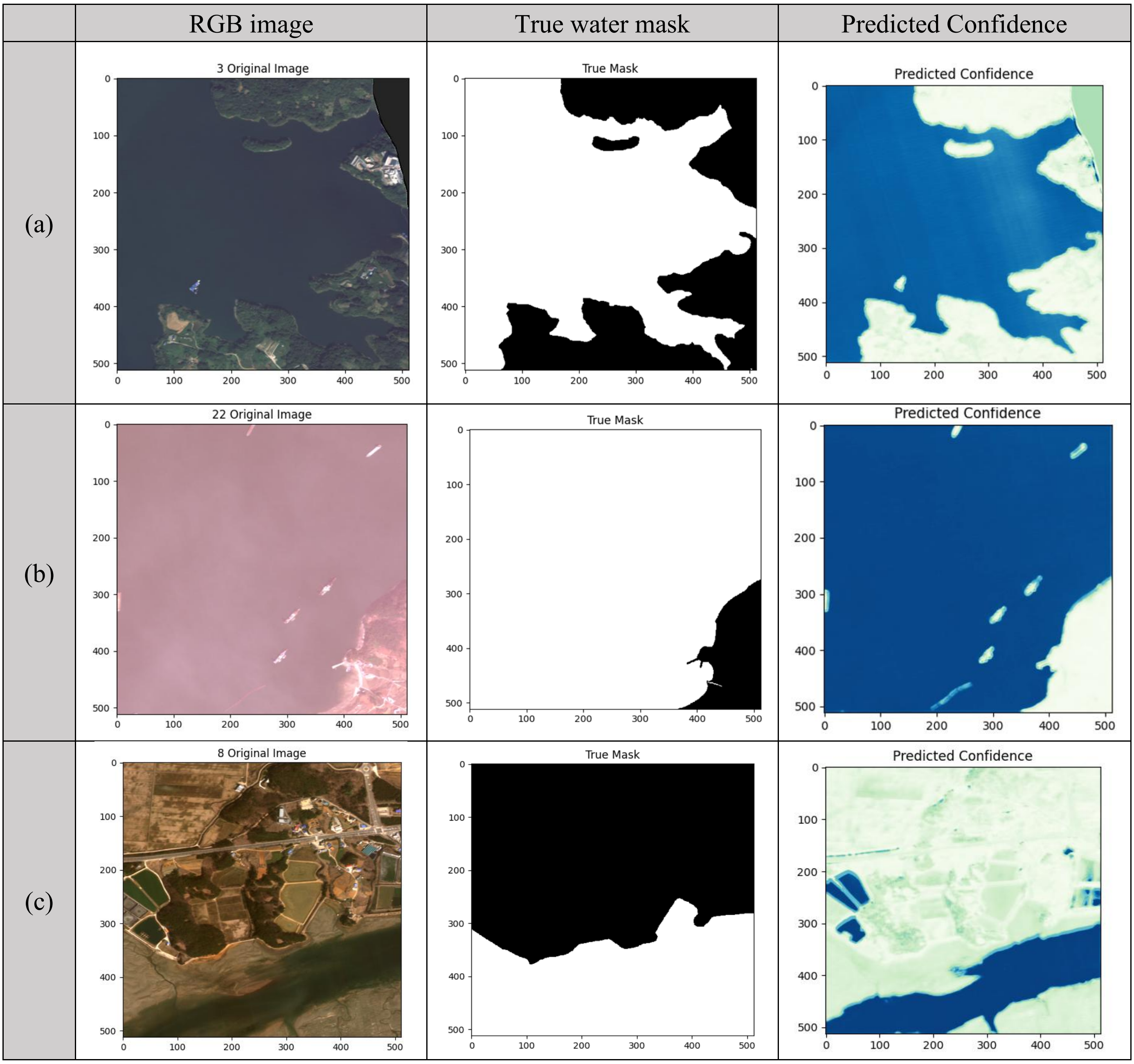

Figure 1. Result of prediction

A comparison of the true mask, U-Net-based detection results, and NDWI results demonstrated that U-Net achieved the highest accuracy in water body detection. The NDWI-based method achieved an Accuracy of 0.49, Precision of 0.83, Recall of 0.35, and F1-Score of 0.41. While the high Precision indicates that NDWI effectively identifies water pixels when predicted, the low Recall suggests that a large proportion of actual water bodies were missed. In contrast, the U-Net model demonstrated superior performance with an Accuracy of 0.90, Precision of 0.93, Recall of 0.91, and F1-Score of 0.91, indicating that it can accurately and comprehensively detect water bodies by leveraging spatial context and multi-band information.

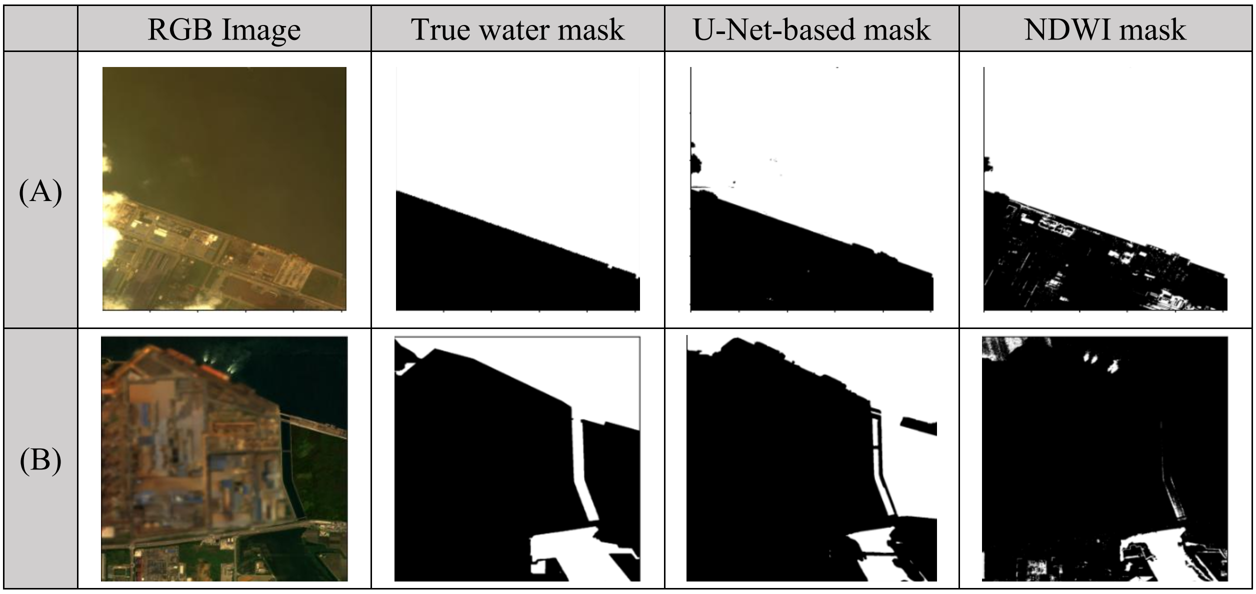

Figure 2. Comparison between U-Net-based Water Body Detection and NDWI

The application of the U-Net model to the same region demonstrated its ability to accurately detect water body by effectively capturing temporal color differences and subtle changes such as ship movements.

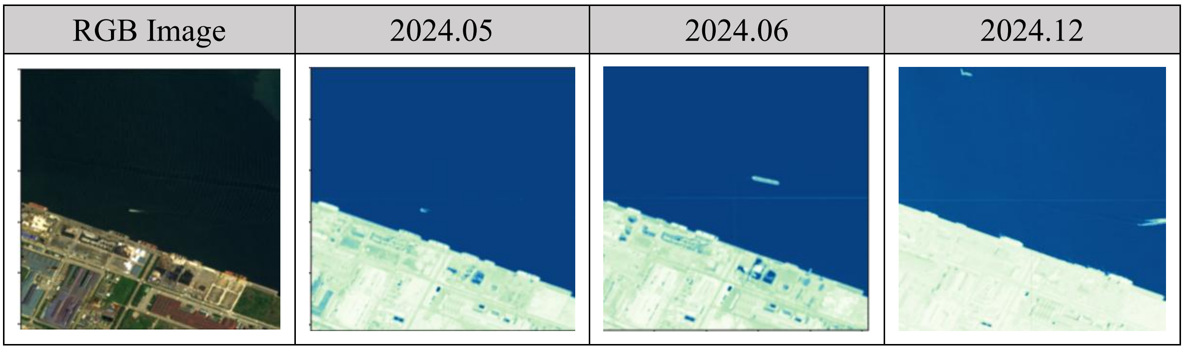

Figure 3. Time-series comparison

4. Conclusion

In this study, we proposed an automated approach for water body detection in CAS500-1 high-resolution satellite imagery using the U-Net deep learning model. For the methodology, RGB and NIR bands along with the water mask from ARD data were utilized. The images were preprocessed by resizing to 1024×1024, normalized, and then tiled into 512×512 patches for model training. The U-Net model, based on semantic segmentation, performed pixel-level classification of water body.

Evaluation results showed that U-Net outperformed the traditional NDWI-based method, achieving an Accuracy of 0.92, Precision of 0.94, Recall of 0.89, and F1-score of 0.92. The relatively lower Recall may result from the limited accuracy of the original water body mask, suggesting inadequacies in the ground truth labeling. NDWI has limitations in distinguishing objects with spectral characteristics similar to water, such as shadows, wetlands, and certain artificial structures. It cannot utilize spatial context or surrounding environmental information. In contrast, U-Net leverages multi-band information and learns spatial patterns within the image, enabling more accurate and comprehensive water body detection.

The model reliably detected water body even in cases of color distortion, the presence of ships, and small islands, and effectively captured subtle temporal variations, detecting water areas not represented in the static True mask. However, the accuracy of water body detection was found to be somewhat lower in areas surrounding mountainous terrain and regions covered by fog. To overcome these limitations, it is necessary to incorporate features that can better distinguish the intrinsic characteristics of water body, such as spatial patterns and spectral information. Additionally, it may be necessary to expand the multi-spectral range by generating synthetic bands or to employ more advanced deep learning models.

These findings demonstrate that deep learning-based semantic segmentation can significantly improve the automation and accuracy of water body detection in high-resolution satellite imagery. The proposed approach can support water resource management and monitoring for applications such as agriculture, urban planning, and disaster response.

Acknowledgements

This work is supported by the Korea Agency for Infrastructure Technology Advancement (KAIA) grant funded by the Ministry of Land, Infrastructure and Transport (Grant RS-2022-00155763) and Research Program for Agricultural Science & Technology Development

(Project No. PJ0162342025)", National Academy of Agricultural Science, Rural Development Administration, Republic of Korea.

References

Cho, N., Lee, J., & Kim, H. (2023). A study on the land change detection and monitoring using high-resolution satellite images and artificial intelligence: A case study of Jeongeup City. Journal of cadastre & land informatix, v.53 no.1, pp.107 – 121.

Kim, W., Cho, S., Jeong, J., Kim, Y., Kim, H. O., & Choi, M. (2024). Evaluation of deep learning-based water bodies and flooded area detection with nanosatellites: The PlanetScope satellite imageries and HRNet model. Korean Journal of Remote Sensing, 40, pp. 617-627.

Lee, H., Kim, K., Lee, K. (2024). Application of Geo-Segment Anything Model (SAM) Scheme to Water Body Segmentation: An Experiment Study Using CAS500-1 Images. *Korean Journal of Remote Sensing*, *40*(4), pp. 343-350