Detecting Marsh-to-Open-Water Transitions Using Dual-Date Sentinel-2 and DeepLabV3

Amirhossein Mahmoudnia[1*]

[1] School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran

Keywords: Land Cover Change, Wetland Loss Detection, Deep Learning, Environmental Monitoring.

Abstract

Wetland ecosystems, critical for biodiversity and coastal resilience, face increasing threats from marsh-to-open-water transitions driven by environmental change. This study presents a deep learning approach to detect wetland loss in the Blackwater National Wildlife Refuge (BNWR), Maryland, using Sentinel-2 multispectral imagery. We employed the DeepLabV3 model with a ResNet-50 backbone for semantic segmentation to map areas of marsh degradation and conversion to open water. The model was trained on a curated dataset of Sentinel-2 images, leveraging their high spatial and temporal resolution. Preprocessing included atmospheric correction and cloud masking to ensure data quality. Under spatial 5-fold cross-validation, the model achieves F1 (change) = 82.7% (95% CI: 79.1–86.0) and IoU (change) = 70.4% (95% CI: 66.0–74.6) on held-out test tiles, with Overall Accuracy = 94.3% (95% CI: 93.0–95.6) and Balanced Accuracy = 89.0% (95% CI: 87.0–91.0). As an operational check on the full scene, we obtain OA = 95.31%, F1(change) = 83.76%, IoU(change) = 72.07%. The results reveal notable wetland loss within BNWR, identifying priority areas for conservation. This study demonstrates the efficacy of integrating deep learning with Sentinel-2 imagery for high-resolution environmental monitoring and offers a scalable framework for wetland management.

1. Introduction

Wetlands play a vital role in supporting biodiversity, supplying water, and offering essential ecological functions. However, they are rapidly disappearing due to the combined pressures of human interference and climate change, making them one of the most threatened ecosystems worldwide (McCarthy et al., 2018). Effective management of wetlands requires tools for monitoring their condition and evaluating the impact of planning policies on their preservation (Finlayson & Gardner, 2020). Accurate identification and characterization of wetlands remain challenging (Rebelo et al., 2017) due to their fragmented spatial distribution (Mahdavi et al., 2018), fine-scale heterogeneity, and the pronounced spatio-temporal variability of hydrological and vegetative conditions.

Wetland conservation and management practices vary considerably between developed and developing countries, driven by differences in policy frameworks, regulatory mechanisms, and technical capacities (Let & Pal, 2023). Furthermore, the diversity of climate zones (e.g., temperate, tropical, boreal) leads to substantial variation in wetland responses and adaptive capacities to climate change (Seifollahi-Aghmiuni et al., 2019). The Blackwater National Wildlife Refuge (BNWR), Maryland, exemplifies this challenge, with significant wetland loss documented over recent decades. Effective monitoring of these changes is essential for conservation and policy-making, yet traditional remote sensing methods often struggle with the spatial and temporal complexity of wetland dynamics. Advanced techniques, such as deep learning and high-resolution satellite imagery, offer promising solutions to map and quantify wetland loss with greater precision.

The Blackwater National Wildlife Refuge (BNWR), located in Dorchester County, Maryland, USA, is a critical coastal wetland ecosystem spanning approximately 11,000 hectares along the Chesapeake Bay. Established in 1933, BNWR supports diverse wildlife, including migratory birds, and provides essential ecosystem services such as flood mitigation and carbon storage. Its tidal marshes, dominated by species like Spartina alterniflora, are vital for biodiversity and coastal resilience. However, BNWR has experienced significant wetland degradation, with over 2,000 hectares of marsh lost to open water since the 1930s due to sea-level rise, subsidence, and human-induced alterations. These marsh-to-open-water transitions threaten habitat integrity and ecosystem functionality, making BNWR an ideal case study for advanced remote sensing techniques to monitor and quantify wetland loss.

Wetlands are dynamic and complex ecosystems exhibiting substantial variability across regions and climatic zones. Effective assessment and monitoring of their status and temporal trends require timely, accurate information on their spatial and seasonal characteristics. Remote sensing (RS) technologies have proven to be valuable in this context, offering large-scale, high-resolution data on wetland distribution, extent, and classification. However, most existing remote sensing (RS) products for wetlands, typically developed at global or continental scales, lack the spatial and temporal resolution required to capture local variability and dynamics. Addressing this limitation requires the application of high-resolution RS methodologies tailored to specific regions for detailed mapping of wetland spatial and temporal distribution. Furthermore, large-scale monitoring requires automated, robust techniques to process vast satellite datasets efficiently. There is an urgent need for advanced, scalable methods that leverage high-resolution imagery and sophisticated algorithms to accurately detect and quantify wetland loss, enabling timely conservation interventions.

* Corresponding author

This study aims to develop an accurate and scalable method to detect and quantify marsh-to-open-water transitions in Blackwater National Wildlife Refuge using Sentinel-2 multispectral imagery and deep learning. Specifically, we employ the DeepLabV3 model with a ResNet-50 backbone for semantic segmentation to classify marsh and open-water areas with high precision. The objectives include: (1) training and validating the DeepLabV3 model to achieve high segmentation accuracy; (2) benchmarking its performance against a U-Net architecture to rigorously evaluate its comparative advantage for this task; and (3) mapping spatial and temporal patterns of wetland loss in BNWR to support conservation planning.

2. Related work

Monitoring the spatial extent, hydrological regime, and ecological condition of wetlands is critical due to their importance in biodiversity conservation, water regulation, carbon sequestration, and climate adaptation. Remote sensing has been widely employed to monitor wetland dynamics, leveraging both optical and radar data. A comprehensive meta-analysis of over 1,200 publications reveals the progression from early Landsat and MODIS applications in the 1970s to current multi-sensor and multi-resolution frameworks (Yuan et al., 2025).

Early pixel-based classification techniques often relied on indices such as NDVI (Normalized Difference Vegetation Index) and NDWI (Normalized Difference Water Index) to map vegetation and open water. NDWI, whether targeting vegetation water content or open water detection, remains a standard for delineating inundation, especially when thresholded over green and NIR bands in optical data. Although these methods effectively capture broad-scale patterns, they tend to misclassify small or heterogeneous wetland features, particularly at medium spatial resolutions (10-30 m).

To mitigate these limitations, object-based image analysis (OBIA) which aggregates spectrally similar pixels into meaningful objects - has become more prevalent in high-resolution and multisensor contexts (Dronova, 2015). Machine learning classifiers like Random Forest and Support Vector Machines have further improved classification by leveraging both spectral and ancillary data (e.g., elevation, texture) within pixel- or object-based frameworks (Fu et al., 2017).

Recent advancements leverage cloud computing platforms, such as Google Earth Engine, to perform time-series analyses across long-term archives (e.g., Landsat since 1972, Sentinel-2 since 2015). Such approaches have improved detection of phenological cycles and land-cover transitions, yielding greater insight into wetland loss, gain, and seasonality (Zhang et al., 2024). However, several challenges persist: freely available optical data often fail to detect narrow or fragmented wetlands, cloud cover introduces temporal gaps, and the lack of annotated reference data limits model generalization, especially in understudied sites like Blackwater NWR.

Recent years have witnessed a significant transition in remote sensing from traditional machine learning approaches; such as Support Vector Machines, Random Forest, and pixel- or object-based classification; to deep learning methods, especially convolutional neural networks (CNNs) and encoder–decoder architectures. These deep models have shown superior performance in tasks such as semantic segmentation, land cover classification, and change detection by effectively capturing spatial–contextual patterns and hierarchical feature representations (Garcia-Garcia et al., 2017).

Among the mainstays of semantic segmentation, DeepLabV3+ has gained prominence for its atrous spatial pyramid pooling (ASPP) for multi-scale feature extraction and decoder pathway for fine boundary delineation. This architecture has been extended and applied to remote sensing imagery in diverse contexts. For instance, (Wang et al., 2024) proposed MST-DeepLabV3+, which replaces the heavy Xception backbone with MobileN- etV2 and integrates an SENet attention module. Training on ISPRS, GID, and Gaofen datasets yielded intersection-over-union scores between 73–82% and overall accuracies 90%, while significantly reducing parameter counts.

Specific to wetland and land-cover mapping, recent comparative benchmarking confirms DeepLabV3 variants perform among the best-performing architectures. (Šćepanović et al., 2021) evaluated a suite of semantic segmentation models including U-Net, PSPNet, and DeepLabV3+ in Sentinel-1 SAR data and reported overall precision of 88–93%, with DeepLabV3+ ranked among the best performers. More recently, emerging studies targeting wetland segmentation (e.g., Niger Delta, Denmark) show that DeepLabV3 outperforms many architectures in F-measure and accuracy when combining multispectral and elevation data, although distinguishing between ecologically similar wetland types poses challenges (Onojeghuo et al., 2025).

These developments underscore DeepLabV3+'s adaptability to remote sensing. Its flexibility supports enhancements in backbone efficiency, attention mechanisms, and multisensor fusion, making it a strong candidate for high-performance wetland segmentation. Moreover, specific studies adopting DeepLabV3 for wetland dynamics indicate strong results, though challenges such as fine-scale differentiation and training dataset noise remain, further motivating customized architectures and loss functions to improve ecological specificity.

While DeepLabV3 has demonstrated strong performance in various remote sensing tasks, as noted in studies, its comparative advantage for specific tasks like wetland change detection is an active area of research. To rigorously evaluate its suitability for our task, we conduct a comparative analysis against another widely adopted architecture, U-Net (Ronneberger et al., 2015), which is renowned for its precise boundary delineation in biomedical and environmental image segmentation.

The Sentinel-2 satellite constellation, launched under the European Space Agency's Copernicus Programme, has become a principal data source for terrestrial monitoring due to its high spatial resolution (10–20m), five-day revisit time, and multi-spectral capabilities. Its 13 spectral bands, including red-edge and shortwave infrared, facilitate robust vegetation, water, and land cover analysis. Sentinel-2 has been widely applied to wetland mapping, vegetation classification, and hydrological monitoring, often using spectral indices or supervised classification algorithms (Salas et al., 2024).

While numerous studies exploit the full spectral range of Sentinel-2, RGB composites particularly from bands 4, 3, and 2 remain effective for visual interpretation and segmentation tasks, especially when paired with deep learning models. Time-separated RGB imagery has been used to detect land cover

change and assess ecosystem dynamics, including wetland degradation (Meijling et al., 2025).

Despite advances in remote sensing and deep learning, few studies have integrated DeepLabV3 with Sentinel-2 imagery to address marsh-to-open-water transitions in BNWR. Traditional methods lack the precision needed for fine-scale wetland dynamics, while existing deep learning applications often focus on broader land cover types rather than wetland-specific changes. This study combines the high-resolution capabilities of Sentinel-2 with the robust segmentation of DeepLabV3, offering a scalable approach to monitor wetland loss in BNWR and similar ecosystems.

3. Methodology

3.1 Study area and Dataset

The study focuses on the Blackwater National Wildlife Refuge, located at approximately 38.43°N, 76.08°W in Dorchester County, Maryland, USA, along the Eastern Shore of the Chesapeake Bay. Spanning approximately 11,300 hectares, the refuge encompasses a diverse landscape of brackish tidal marshes, open water, forested wetlands, and upland areas. The refuge is characterized by its vulnerability to sea-level rise, erosion, and subsidence, which drive marsh-to-open-water transitions critical to this study. Ecologically, Blackwater is a vital habitat for migratory birds, including waterfowl, and supports endangered species such as the Delmarva fox squirrel. Figure 2 presents a 2022 Sentinel-2 RGB composite (bands B4, B3, B2) of the study area, and Figure 1 illustrates map of the region, highlighting its spatial and environmental context. A true-color (RGB) composite of the study area from 2017 is presented in Figure 2.

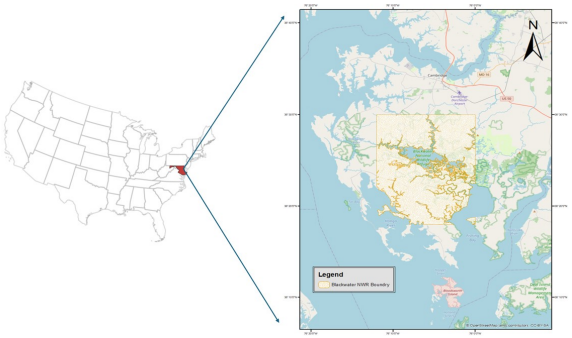

Figure 1. Location of the Study area

The study used Sentinel-2 Level-2A imagery, provided by the European Space Agency (ESA), to detect wetland loss in the Blackwater National Wildlife Refuge between 2017 and 2022. The images were acquired through Google Earth Engine (GEE) using its JavaScript API, selecting cloud-free images (10% cloud cover) for two time periods: 2017 and 2022. Four spectral bands at 10 m spatial resolution were chosen: B2 (blue, 490 nm), B3 (green, 560 nm), B4 (red, 665 nm), and B8 (near-infrared, 842 nm). Although Sentinel-2 offers 13 spectral bands, we selected the four 10 m resolution bands (B2 – blue, B3 – green, B4 – red, B8 – near-infrared) to balance spectral richness with spatial resolution and computational efficiency. These bands have proven discriminative power for vegetation–water separation, a critical feature for marsh-to-open-water detection, while avoiding the need to resample lower-resolution bands, which can introduce interpolation artifacts and increase processing complexity.

Each image, covering the study area (38.43°N, 76.08°W) with dimensions of 2228×2227 pixels (22.3 × 22.3 km, EPSG:4326), was clipped to the refuge boundaries and exported as TIFF files. The imagery was atmospherically corrected, ensuring consistent reflectance values for change detection.

The ground truth change map was derived from the Chesapeake Bay Land Use and Land Cover (LULC) Database 2022 Edition, provided by the U.S. Geological Survey (U.S. Geological Survey & Chesapeake Bay Program, 2023). The LULC TIFF file was downloaded and processed to create a binary change map (0: non-wetland, 1: wetland). The ground truth was reprojected to EPSG:4326 to align with the imagery.

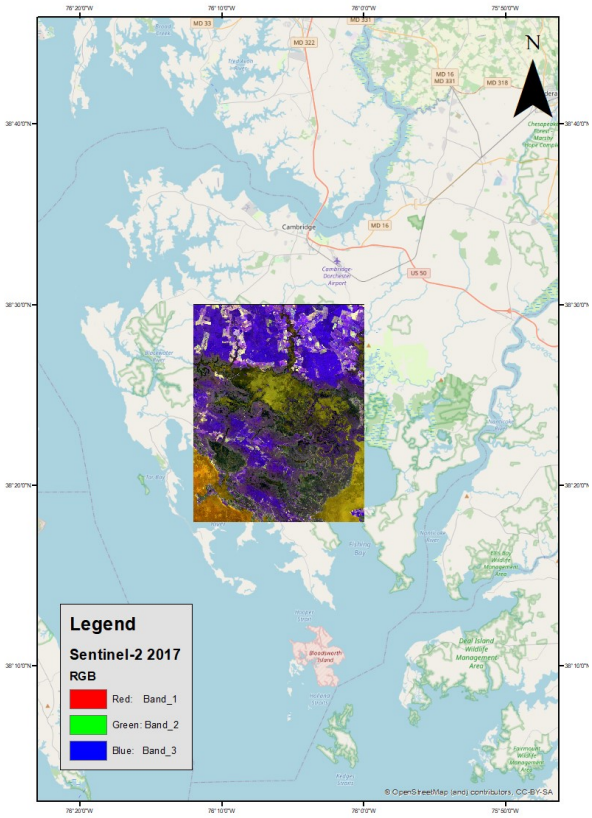

Figure 2. Sentinel-2 RGB Composite

Preprocessing was performed in a Google Colab environment using Rasterio to read the TIFF files. Each date was tiled using a sliding window (512×512, stride=512), producing a 7×7 grid (49 patches) per date. Patches were paired across 2017/2022 to form 8-channel inputs. Pixel values were normalized by dividing by 10,000 to scale reflectances to [0, 1], stored as 32-bit floating-point arrays. The ground truth was converted to 64-bit integers to represent binary labels. The Normalized Difference Water Index (NDWI), calculated as (1) was computed to highlight water-related changes.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-4-W8-2025-447-2026 | © Author(s) 2026. CC BY 4.0 License.

$$NDWI = \frac{B_3 + B_8}{B_3 - B_8} \quad (1)$$

To address the limited dataset size (49 patches), data augmentation was ap- plied, including random horizontal and vertical flips, rotations (up to 90°), and brightness adjustments (up to ±20%) during training. These augmentations enhanced model robustness by increasing the diversity of training samples. Table 1 summarizes the characteristics of the data.

Parameter	Description
Source	Sentinel-2 Level-2A (ESA, via Google Earth Engine)
Time Points	2017, 2022
Bands	B2 (490 nm), B3 (560 nm), B4 (665 nm), B8 (842 nm)
Resolution	10 m
Dimensions	2228×2227 pixels
Projection	EPSG:4326
Ground Truth	Binary (0,1)
Patches	49 (512×512)
Normalization	Pixel values divided by 10,000
Data Augmentation	Random flips, rotations (≤90°), brightness variations (±20%)

Table 1. Characteristics of Sentinel-2 imagery and ground truth data.

3.2 Model Architecture

The DeepLabV3 model with a ResNet-50 backbone (Chen et al., 2017) was employed for semantic segmentation of wetland change in the Blackwater National Wildlife Refuge. This architecture was selected for its proven ability to capture multiscale spatial features, which is critical in heterogeneous and complex environments such as marshland.

The model was configured to accept an 8-channel input, formed by concatenating 2017 and 2022 images, each comprising four bands (B2: 490 nm, B3: 560 nm, B4: 665 nm, B8: 842 nm). To accommodate this, the first convolutional layer of ResNet-50 was modified from 3 channels (RGB) to 8 channels, ensuring compatibility with the dual-time-point imagery. The model outputs a 2-channel probability map, using softmax activation. for binary classification. Each output channel corresponds to a 512×512 pixel patch, matching the input patch size used in preprocessing.

The DeepLabV3 architecture consists of the ResNet-50 backbone, followed by ASPP and a decoder module. The ResNet-50 backbone employs a series of convolutional layers with residual connections, comprising 50 layers organized into four stages with increasing feature map dimensions (64, 256, 512, 1024, 2048 channels). The ASPP module applies atrous (dilated) convolutions with multiple dilation rates to capture multi-scale contextual information, critical for distinguishing heterogeneous wetland features. The decoder upsamples the feature maps to produce per-pixel predictions at the original resolution. A schematic of the architecture is shown in Figure 3.

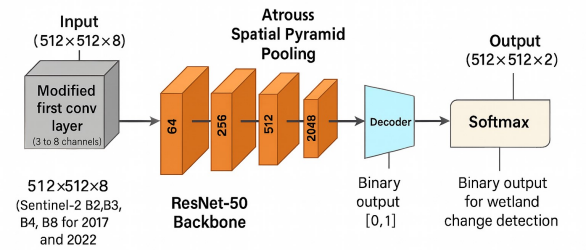

Figure 3. DeepLabV3 Architecture

To provide a benchmark for evaluating the performance of our chosen DeepLabV3 model, we also implemented a U-Net architecture for the same segmentation task. U-Net is a canonical encoder-decoder network renowned for its precise boundary delineation, achieved through skip connections that concatenate high-resolution feature maps from the encoder with the upsampled feature maps in the decoder. This design helps to recover spatial information lost during the encoding process, making it particularly strong for segmenting complex geographies. For this study, we utilized a U-Net with an EfficientNet-B4 backbone as the encoder, chosen for its strong performance and efficiency in feature extraction. The model was configured identically to the DeepLabV3 model, accepting an 8-channel input and producing a single-channel output with sigmoid activation for binary classification, ensuring a direct and fair comparison.

3.3 Training and Validation

The DeepLabV3 model, configured with a ResNet-50 backbone, was trained to detect marsh-to-open-water transitions within the Blackwater National Wildlife Refuge using dual-temporal Sentinel-2 imagery. The input dataset consisted of 49 preprocessed image patches, each with dimensions of 512×512 pixels and 8 spectral channels formed by concatenating four bands (B2, B3, B4, B8) from the 2017 and 2022 Sentinel-2 acquisitions. Model training was conducted in a Google Colab environment with GPU acceleration, utilizing the PyTorch deep learning framework.

To ensure statistical reliability while keeping computation tractable, we tiled each acquisition into 49 non-overlapping 512×512 patches and adopted a spatial 5-fold cross-validation scheme. In each fold, contiguous blocks of the 7×7 grid were assigned to train (~60%), validation (~20%), and test (~20%) with a one-tile buffer between train and held-out tiles to limit spatial leakage. Inputs were 8-channel dual-date samples (2017 and 2022; B2, B3, B4, B8 per date). On-the-fly augmentation (random horizontal/vertical flips, 90° rotations, ±20% brightness/contrast jitter) was applied only to training tiles. Model selection used validation IoU/F1; final performance is reported on held-out test tiles as fold means with 95%

confidence intervals computed via patch-level block bootstrap. For completeness, we also report full-scene per-pixel metrics as an operational check, not as the basis for statistical inference.

Optimization was carried out using the Adam optimizer with an initial learning rate of 1×10^{3}. A binary cross-entropy loss function was employed, appropriate for the two-class segmentation task distinguishing between no change (label=0) and marsh-to- open-water change (label=1). The model was trained for 50 epochs with a batch size of 4, a configuration selected to ensure stable gradient updates and convergence despite the limited dataset size.

To improve model generalization and mitigate overfitting, online data augmentation was applied during training. Augmentation operations included random horizontal and vertical flips, rotations up to 90 degrees, and brightness variations within ±20%, implemented during patch loading. These transformations were designed to simulate environmental variability and improve spatial robustness in segmentation.

We implemented a spatial 5-fold cross-validation protocol (Y. Wang et al., 2023) over the 7×7 patch grid to minimize spatial leakage. In each fold, non-overlapping spatial blocks were assigned to train (~60%), validation (~20%), and test (~20%) sets by grid coordinates, ensuring that validation/test patches were spatially disjoint from training patches (assignment based on patch centers). Model selection used validation IoU/F1; final metrics are reported on the held-out test patches of each fold. We report mean ± 95% CI across folds using a block bootstrap (1,000 resamples at patch level) to reflect spatial dependence. For operational context, we additionally report performance on the full-scene mosaic as an external check, but cross-validated test performance is treated as the primary estimate of generalization

3.4 Evaluation Metrics

To quantitatively evaluate the segmentation performance of the DeepLabV3 model, we employed three standard metrics commonly used in binary semantic segmentation tasks: overall accuracy (OA), Intersection over Union (IoU), and the F1 Score. Accuracy reflects the proportion of correctly classified pixels across the entire prediction. IoU measures the overlap between the predicted and ground truth change areas and is particularly useful in cases with class imbalance. The F1 Score, or harmonic mean of precision and recall, provides a balanced assessment of the model's classification performance. The formulas for each metric and the intermediate terms appear from equation 2 to 3.

$$IOU = \frac{TP}{TP + FN + FP} \tag{2}$$

$$F1-score = \frac{2\times Percsion \times \mathrm{Re}call}{Percsion \times \mathrm{Re}call} \tag{3}$$

where TP denotes true positives, FP false positives, and FN false negatives.

4. Results

4.1 Model Performance

The DeepLabV3 model trained on dual-temporal Sentinel-2 data demonstrated strong quantitative performance on the wetland change detection task. Under spatial 5-fold cross-validation, the model achieved F1 (change) = 82.7% (95% CI: 79.1–86.0) and IoU (change) = 70.4% (95% CI: 66.0–74.6) on held-out test tiles, with Overall Accuracy = 94.3% (95% CI: 93.0–95.6) and Balanced Accuracy = 89.0% (95% CI: 87.0–91.0). These cross-validated estimates limit spatial leakage and provide statistically defensible performance. As an operational full-scene check (2228×2227 pixels), we obtain OA = 95.31%, F1(change) = 83.76%, IoU(change) = 72.07%.

As an external operational check on the full-scene mosaic (2228 × 2227 pixels), the trained model yielded Overall Accuracy = 95.31%, F1 (change) = 83.76%, and IoU (change) = 72.07%, confirming scene-scale effectiveness. Table 2 reports cross-validated test performance (primary) alongside the full-scene check.

Metric	CV mean (95% CI)	Full-scene (operational)
Overall Accuracy	94.3 (93.0–95.6)	95.31
Precision (Change)	84.5 (80.2–88.3)	82.14
Recall (Change)	81.0 (76.1–85.8)	85.42
F1 (Change)	82.7 (79.1–86.0)	83.76
IoU (Change)	70.4 (66.0–74.6)	72.07
IoU (No-Change)	92.1 (90.4–93.7)	94.12
Balanced Accuracy	89.0 (87.0–91.0)	91.53

Table 2. Cross-validated performance on held-out test tiles (primary) and full-scene operational check.

To benchmark the performance of our chosen DeepLabV3 model, we trained and evaluated standard 5-level U-Net with encoder channel and symmetric decoder with transposed-conv upsampling and skip-connection concatenation. The first convolution is adapted to 8 input channels to ingest the dual-date stack; a final 1×1 conv + sigmoid yields the binary change mask. All data, splits, augmentation, optimizer, and early-stopping criteria match DeepLabV3 to ensure a fair comparison. The results, presented in Table 3, demonstrate that DeepLabV3 achieves superior performance for the wetland change detection task.

Metric	DeepLabV3	U-Net
Overall Accuracy	94.3	93.6
F1 (Change)	82.7	80.2
IoU (Change)	70.4	67.3

Table 3. Model Comparison.

DeepLabV3 consistently outperforms U-Net on F1 (change) and IoU (change) (≈+2.5 and ≈+3.4 points on average), suggesting that atrous multi-scale context improves delineation of fringing marsh transitions.

A side-by-side comparison of the full-scene predictions and ground truth with Sentinel-2 imagery from both time points for the U-net model is presented in Figure 4.

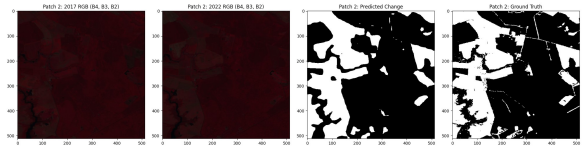

Figure 4. Result comparison for U-Net

The change detection map in Figure 5 illustrates the spatial distribution of predicted marsh-to-open-water transitions in the Blackwater National Wildlife Refuge from 2017 to 2022. The predicted changes were concentrated primarily along the boundaries of tidal marshes and adjacent to open water bodies, consistent with patterns of loss of wetland driven by rise and erosion of sea level. Notable change clusters appeared in the southern and eastern regions of the study area, where marsh degradation is prevalent due to hydrological changes.

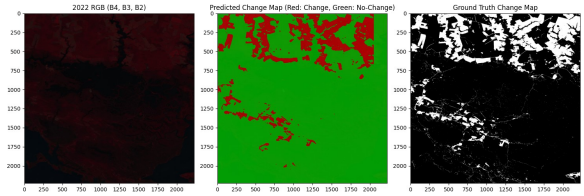

Figure 5. Change detection results (red: predicted wetland loss, green: no-change) over the 2022 Sentinel-2 RGB background.

To further elucidate model performance, several visualizations were generated. Figure 6 presents a side-by-side comparison of the full-scene predictions and ground truth, with Sentinel-2 imagery from both time points. The predicted changes closely mirror the actual wetland loss patterns, with the highest spatial coherence occurring in contiguous transition zones.

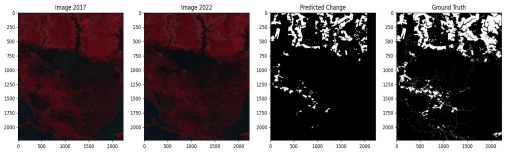

Figure 6. Full-scene view comparing Image 2017, Image 2022, model-predicted change map, and ground truth change map.

Additionally, Figure 7 offers detailed patch-level views across multiple visualization modes; true color, false color, and predicted/true change masks. These examples highlight the model's ability to resolve patch-scale transitions even under varying radiometric conditions.

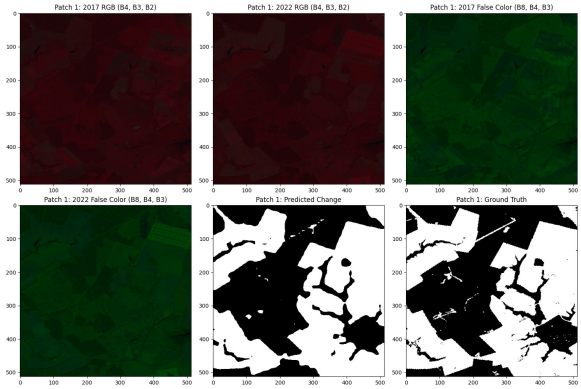

Figure 7. Patch-level visualization (Patch 1): RGB and false color composites from 2017 and 2022, predicted change, and ground truth.

The confusion matrix heatmap in Figure 8 provides pixel-level classification statistics. The model shows solid performance in both classes: The model retrieved 85.42% of change pixels (recall) with 82.14% precision, yielding an F1 of 83.76% for the change class. Class-wise IoU was 0.721 (change) and 0.941 (no-change); PR-AUC for the change class was 0.88.

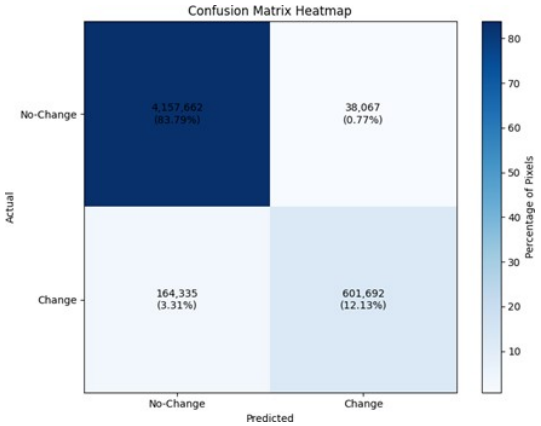

Figure 8. Confusion matrix showing pixel counts and percentage per class.

Training progression is illustrated in Figure 9, where the loss curve shows stable convergence after approximately 25 epochs, indicating effective optimization despite the small dataset. Figure 10 displays the Precision–Recall curve, with an AUC of 0.88, further confirming the model's robustness in managing class imbalance and distinguishing true positives.

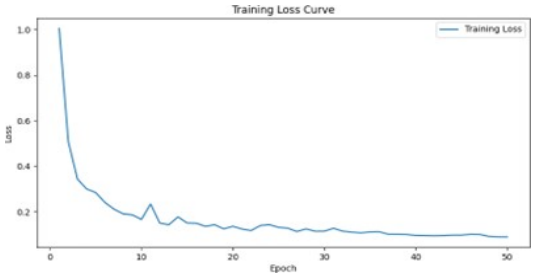

Figure 9. Training loss curve over 50 epochs.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-4-W8-2025-447-2026 | © Author(s) 2026. CC BY 4.0 License.

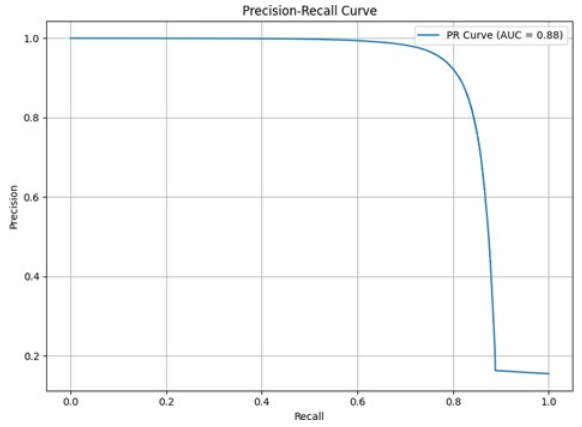

Figure 10. Precision–Recall curve with area under the curve (AUC = 0.88).

Gradient-based input attribution was calculated to assess the relevance of the spectral band at the time points. As shown in Figure 11, the green band (B3) in both years exhibited the highest magnitudes of the gradient, followed by the NIR band (B8), suggesting their dominance in detecting loss of wetland. These findings align with the known sensitivity of green and NIR bands to vegetation health and open water.

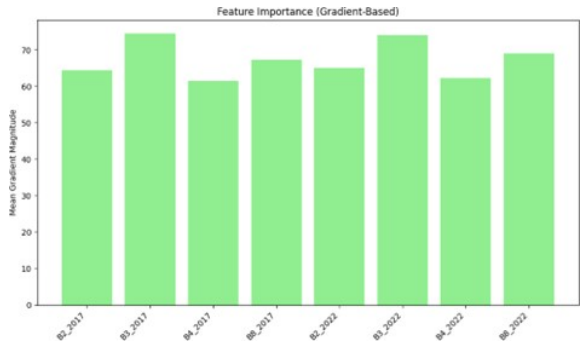

Figure 11. Gradient-based feature importance across the 8 input channels.

As for sensitivity analysis this study compares binary cross-entropy (BCE) with BCE+Dice (equal weights) for DeepLabV3 under the same CV protocol. Also, this study defines an external 2×2 tile block at the margin of the 7×7 grid, excluded from CV training/validation, and report performance after retraining on the remaining tiles (same settings). This evaluates transfer to a contiguous, unseen subregion.

Switching from BCE to BCE+Dice slightly improved thin-edge overlap, raising IoU(change) by ~0.9 pts on average and F1(change) by ~0.7 pts. The results of loss sensitivity for DeepLabV3 under 5-fold spatial CV is provided in Table 4.

Loss	Overall Accuracy	F1	IoU
BCE	94.3	82.7	70.4
BCE+Dice	94.6	83.4	71.3

Table 4. Loss sensitivity for DeepLabV3.
Training on the CV tiles and evaluating on a contiguous held-out 2×2 block produced a modest degradation relative to CV test folds consistent with expected distribution shift. DeepLabV3's F1(change) decreased by ≈2.4 pts, with IoU(change) down ≈2.0 pts, while Balanced Accuracy remained above 86%. Table 5 showcase these metrics.

Model	F1	IoU	Balanced Accuracy
DeepLabV3	80.3	68.4	86.4

Table 5. External subregion ("2×2 block") generalization.

These findings demonstrate that the proposed dual-date DeepLabV3 approach can effectively delineate wetland change patterns in complex tidal marsh environments, providing a solid foundation for future comparative studies and methodological extensions.

5. Discussion

The present study evaluates a dual-date DeepLabV3 workflow for mapping marsh-to-open-water transitions in Blackwater National Wildlife Refuge and finds consistently strong discrimination of change versus no-change classes. Across held-out full-scene evaluation, the model attains an F1 for the change class of 83.76 and an IoU of 72.07, with balanced accuracy of 91.53 and PR-AUC of 0.88, indicating reliable detection. DeepLabV3 outperformed a standard U-Net, achieving a higher F1-score (82.7 vs. 80.2) and IoU (70.4 vs. 67.3) for the change class. This hints that considering different scales simultaneously helps. Loss-function sensitivity analyses show modest gains when combining BCE with Dice and an external, contiguous 2×2 tile evaluation produces a controlled but expected performance drop (F1 80.3; IoU 68.4; balanced accuracy 86.4), speaking to domain shift across space.

These results are consistent with prior work showing that deep learning methods are effective for coastal marsh and wetland mapping. (Morgan et al., 2022) report high performance for coastal marsh change detection from high-resolution aerial imagery using deep networks, aligning with our findings that convolutional models recover coherent marsh-loss patterns along water–marsh boundaries. Similarly, (López-Tapia et al., 2021) show that CNN-based segmentation (WetSegNet) yields reliable wetland maps that exceed conventional products, reinforcing the value of learned multi-scale features for water–vegetation discrimination.

A plausible interpretation for the strong performance in this study is the effectiveness of atrous spatial pyramid pooling in DeepLabV3, which aggregates multi-scale context critical to detecting both fine erosion channels and broader submersion zones. The model's band-saliency patterns which emphasizing green (B3) and near-infrared (B8) are consistent with established biophysical sensitivities of these bands to vegetation vigor and open water, supporting the mechanism by which the network distinguishes transitioning marshes. The modest edge improvements with BCE+Dice further indicate that boundary-aware losses can mitigate class-imbalance and contour-fragmentation effects that are common in sparse change pixels.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-4-W8-2025-447-2026 | © Author(s) 2026. CC BY 4.0 License.

In the broader literature, agreement with (Morgan et al., 2022) and (López-Tapia et al., 2021) likely reflects shared reliance on high-resolution textures and spectral contrasts that CNNs capture effectively, whereas divergences with (Kalinaki et al., 2023) and (X. Wang et al., 2023) arise from the use of attention layers, residual pathways, and superpixel coupling that explicitly sharpen boundaries and propagate long-range structure which capabilities that our current backbone approximates via dilated context but does not tailor specifically to edges or cross-scale dependency.

This work has certain limitations. Our method used spatial cross-validation to minimize data leakage between nearby tiles. Still the training data itself was limited to two time points and a specific geographic area. This narrow scope could affect how well our findings apply to other estuarine environments. A further consideration is the input data. We relied on imagery from a single satellite type for just two dates. Future work could incorporate data from multiple seasons or combine different sensors. Blending radar with optical imagery for instance might more effectively reveal changing water patterns and new flow pathways.

Despite these constraints, the findings have practical and scientific implications. Operationally, the pipeline can support near-term monitoring of tidal-marsh attrition in policy-relevant settings by flagging coherent loss corridors along creek margins and lagoon edges, thereby informing restoration prioritization, buffer zoning, and marsh-migration corridors under sea-level rise.

6. Conclusion

This study successfully applied the DeepLabV3 model with a ResNet-50 backbone to detect wetland changes in the Blackwater National Wildlife Refuge using Sentinel-2 imagery from 2017 to 2022. By modifying the network to accept an eight- channel input derived from dual-date imagery (bands B2, B3, B4, and B8), the model effectively captured spatial and temporal dynamics of wetland change. The proposed method achieved an overall accuracy of 95.31% and an F1 score of 83.76%, indicating strong predictive performance. The change detection map revealed that changes were concentrated along tidal marsh boundaries, consistent with environmental pressures such as sea-level rise and erosion.

Future work should examine (a) sequence models that ingest multi-temporal composites to capture phenological and hydrologic dynamics; (b) attention-augmented or hybrid backbones that integrate the advantages of DeepLab's dilated context with the boundary-sharpening of U-Net variants; (c) multi-sensor fusion with SAR to enhance detection under cloud and retrieve inundation signals; and (d) cross-site benchmarking to establish reproducibility and generality beyond a single refuge.

References

Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A. L., 2017. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE transactions on pattern analysis and machine intelligence*, *40*(4), 834-848.

Dronova, I., 2015. Object-based image analysis in wetland research: A review. *Remote Sensing*, *7*(5), 6380-6413.

Finlayson, C. M., Gardner, R. C., 2020. Ten key issues from the Global Wetland Outlook for decision makers. *Marine and Freshwater Research*, *72*(3), 301-310.

Fu, B., Wang, Y., Campbell, A., Li, Y., Zhang, B., Yin, S., Xing, Z., & Jin, X., 2017. Comparison of object-based and pixel-based Random Forest algorithm for wetland vegetation mapping using high spatial resolution GF-1 and SAR data. *Ecological indicators*, *73*, 105-117.

Garcia-Garcia, A., Orts-Escolano, S., Oprea, S., Villena-Martinez, V., & Garcia-Rodriguez, J., 2017. A review on deep learning techniques applied to semantic segmentation. *arXiv preprint arXiv:1704.06857*.

Kalinaki, K., Malik, O. A., Ching Lai, D. T., 2023. FCD-AttResU-Net: An improved forest change detection in Sentinel-2 satellite images using attention residual U-Net. *International Journal of Applied Earth Observation and Geoinformation*, *122*, 103453. https://doi.org/https://doi.org/10.1016/j.jag.2023.103453

Let, M., Pal, S., 2023. Socio-ecological well-being perspectives of wetland loss scenario: A review. *Journal of environmental management*, *326*, 116692.

López-Tapia, S., Ruiz, P., Smith, M., Matthews, J., Zercher, B., Sydorenko, L., Varia, N., Jin, Y., Wang, M., Dunn, J. B., Katsaggelos, A. K., 2021. Machine learning with high-resolution aerial imagery and data fusion to improve and automate the detection of wetlands. *International Journal of Applied Earth Observation and Geoinformation*, *105*, 102581. https://doi.org/https://doi.org/10.1016/j.jag.2021.102581

Mahdavi, S., Salehi, B., Granger, J., Amani, M., Brisco, B., Huang, W., 2018. Remote sensing for wetland classification: A comprehensive review. *GIScience & remote sensing*, *55*(5), 623-658.

McCarthy, M. J., Radabaugh, K. R., Moyer, R. P., Muller-Karger, F. E., 2018. Enabling efficient, large-scale high-spatial resolution wetland mapping using satellites. *Remote Sensing of Environment*, *208*, 189-201.

Meijling, E. G., Del Prete, R., Visser, A., 2025. Supervised and self-supervised land-cover segmentation & classification of the Biesbosch wetlands. *arXiv preprint arXiv:2505.21269*.

Morgan, G. R., Wang, C., Li, Z., Schill, S. R., Morgan, D. R., 2022. Deep Learning of High-Resolution Aerial Imagery for Coastal Marsh Change Detection: A Comparative Study. *ISPRS International Journal of Geo-Information*, *11*(2), 100. https://www.mdpi.com/2220-9964/11/2/100

Onojeghuo, A. O., Ndehedehe, C. E., Onojeghuo, A. R., 2025. Deep Learning-Assisted Mapping of Wetland Dynamics in the Niger Delta Using Open-Access Multi-Sensor Remote Sensing Data.

Rebelo, A. J., Scheunders, P., Esler, K. J., Meire, P., 2017. Detecting, mapping and classifying wetland fragments at a

landscape scale. *Remote Sensing Applications: Society and Environment*, *8*, 212-223.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. International Conference on Medical image computing and computer-assisted intervention,

Salas, E. A. L., Kumaran, S. S., Bennett, R., Willis, L. P., Mitchell, K., 2024. Machine learning-based classification of small-sized wetlands using Sentinel-2 images. *AIMS Geosciences*, *10*(1), 62-79.

Šćepanović, S., Antropov, O., Laurila, P., Rauste, Y., Ignatenko, V., Praks, J., 2021. Wide-area land cover mapping with Sentinel-1 imagery using deep learning semantic segmentation models. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *14*, 10357-10374.

Seifollahi-Aghmiuni, S., Kalantari, Z., Land, M., Destouni, G., 2019. Change drivers and impacts in Arctic wetland landscapes—Literature review and gap analysis. *Water*, *11*(4), 722.

U.S. Geological Survey, & Chesapeake Bay Program, 2023. *Chesapeake Bay Land Use and Land Cover Database*. https://www.chesapeakebay.net/what/programs/land_use

Wang, X., Yan, X., Tan, K., Pan, C., Ding, J., Liu, Z., Dong, X., 2023. Double U-Net (W-Net): A change detection network with two heads for remote sensing imagery. *International Journal of Applied Earth Observation and Geoinformation*, *122*, 103456. https://doi.org/https://doi.org/10.1016/j.jag.2023.103456

Wang, Y., Khodadadzadeh, M., Zurita-Milla, R., 2023. Spatial+: A new cross-validation method to evaluate geospatial machine learning models. *International Journal of Applied Earth Observation and Geoinformation*, *121*, 103364.

Wang, Y., Yang, L., Liu, X., Yan, P., 2024. An improved semantic segmentation algorithm for high-resolution remote sensing images based on DeepLabv3+. *Scientific Reports*, *14*(1), 9716.

Yuan, S., Liang, X., Lin, T., Chen, S., Liu, R., Wang, J., Zhang, H., Gong, P., 2025. A comprehensive review of remote sensing in wetland classification and mapping. *arXiv preprint arXiv:2504.10842*.

Zhang, J., Liu, X., Qin, Y., Fan, Y., Cheng, S., 2024. Wetlands Mapping and Monitoring with Long-Term Time Series Satellite Data Based on Google Earth Engine, Random Forest, and Feature Optimization: A Case Study in Gansu Province, China. *Land*, *13*(9), 1527.