geomatics

MDPI

Article

A Quantitative Assessment of the Inconsistency Between Waterbody Segmentation and Shoreline Positioning in Deep Learning Models

Wei Wang [1,2,†], Boyuan Lu [1,2,†], Yihan Li [1,2] and Fujiang Ji [3,*]

[1] Department of Civil and Environmental Engineering, University of Wisconsin–Madison, Madison, WI 53706, USA; wwang487@wisc.edu (W.W.); blu38@wisc.edu (B.L.); yihan.li@wisc.edu (Y.L.)
[2] Department of Computer Sciences, University of Wisconsin–Madison, Madison, WI 53706, USA
[3] Department of Forest & Wildlife Ecology, University of Wisconsin–Madison, Madison, WI 53706, USA
* Correspondence: fujiang.ji@wisc.edu
† These authors contributed equally to this work.

Abstract

Accurate shoreline positioning is critical for coastal monitoring and management, yet deep learning shoreline products are often evaluated using conventional waterbody segmentation metrics that do not explicitly measure boundary alignment. Using 20,689 NAIP aerial images covering the Great Lakes shoreline from the Coastal Aerial Imagery Dataset (CAID), we benchmark five semantic segmentation models and quantify the inconsistency between image-level segmentation accuracy (pixel accuracy, IoU) and shoreline positioning accuracy measured by the Shoreline Intersection Ratio (SIR) and Average Eulerian Distance (AED). Although segmentation performance is consistently high (pixel accuracy typically >98% and IoU often >90%), shoreline agreement is substantially lower and strongly landscape-dependent, with the poorest results in wetlands and urban scenes. Correlation analyses across coastal types and water-surface conditions show that the correspondence between segmentation metrics and SIR varies with shoreline morphology. Multivariate regressions confirm the shoreline-to-water ratio (SWR) as the dominant predictor of both SIR and AED, while shoreline complexity (SCI) and mean water hue (MWH) have weaker, context-dependent effects. These results demonstrate that high segmentation accuracy does not guarantee precise shoreline delineation and motivate shoreline-aware evaluation protocols.

Keywords: shoreline positioning; waterbody segmentation; model evaluation; deep learning; aerial imagery

Check for updates

Academic Editor: Alessandro Vitale

Received: 22 December 2025
Revised: 2 February 2026
Accepted: 14 February 2026
Published: 16 February 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Coastal zones are among the most developed regions on Earth, with more than 2.8 billion population living around [1]. These regions are home to dense populations, infrastructure, and economic activities, and thus play a central role in regional and global development [2,3]. At the same time, coastal zones are highly dynamic systems that are increasingly exposed to multiple natural and anthropogenic pressures, including sea-level rise [4–6], erosion [7,8], extreme storm events [9], and changing wave climates [10]. Within this context, the shoreline—the physical boundary between land and water—holds significant importance as a spatial reference for monitoring coastal change, managing risks, and informing environmental and planning policies [11,12]. Accurate shoreline positioning is crucial not only for scientific analysis but also for operational tasks such as shoreline

erosion risk assessment, coastal defense design, and habitat conservation [13,14]. Given the pace of environmental change in coastal areas, timely and precise shoreline mapping has become increasingly important.

Remote sensing technologies provide a practical and scalable means of observing shoreline dynamics across broad spatial extents and frequent time intervals [15–18]. Traditionally, waterbody extraction from satellite and aerial imagery has relied on pixel-level methods using indices or statistical, or machine learning methods, including spectral index (e.g., NDWI, MNDWI, DDWI) [19–21] , Otsu thresholding [22], and supervised pixel classification [23]. However, these approaches often underperform in coastal environments where spectral ambiguity, atmospheric conditions, and land–water mixing introduce noise and confusion [3,24]. In addition, their reliance on predefined spectral thresholds or hand-crafted features limits their flexibility and generalizability, especially in dynamic or heterogeneous coastal zones.

In contrast, deep learning–based semantic segmentation has emerged as a powerful alternative, offering substantial improvements in both accuracy and adaptability. These models learn to classify each pixel in an image by optimizing feature hierarchies through large volumes of labeled data [25,26], enabling a more effective extraction of spatial and contextual information than traditional methods. [27,28]. Fully Convolutional Networks (FCNs) laid the foundation for dense prediction tasks by replacing fully connected layers with convolutional ones, allowing input images of arbitrary size and producing per-pixel outputs [29]. Beyond this foundation, several architectural strategies have been proposed to improve boundary delineation and contextual awareness, including global context encoding, long-range dependency modeling through non-local interactions, multi-scale spatial aggregation, and attention-based feature reweighting [30,31]. Such deep learning–based models are well suited to coastal environments, where shoreline geometry, adjacent land cover, and water-surface conditions vary across spatial scales and exhibit strong spatial heterogeneity.

Despite the rapid adoption of new deep learning architectures for waterbody segmentation, model evaluation practices have remained largely unchanged, relying heavily on image-based metrics such as precision, recall, F1-score, and Intersection over Union (IoU) [32–35]. Consequently, most existing shoreline and waterbody mapping studies evaluate performance primarily through area-based agreement, without explicitly assessing shoreline boundary alignment or positional accuracy. While these metrics capture overall classification performance, they provide limited insight into boundary placement, and a high IoU score may still mask substantial mismatches at the land–water edge [36,37]. This discrepancy becomes particularly problematic in settings with narrow, fragmented, or highly variable shorelines [38,39], where small geometric errors can significantly affect downstream analyses. Furthermore, the spatial inconsistency between predicted and reference shorelines is rarely examined directly, leading to a gap between reported segmentation performance and the actual reliability of shoreline outputs. This gap motivates a systematic analysis of when high segmentation performance translates into accurate shoreline positioning, and which factors contribute to their divergence.

In this study, we quantify the mismatch between waterbody segmentation performance and shoreline positioning accuracy for deep learning models. Using Coastal Aerial Imagery Dataset (CAID) as a benchmark, we compare conventional segmentation metrics (pixel accuracy and IoU) with shoreline-specific accuracy, and we test how their relationships change across coastal landscapes and water-surface conditions. We then relate the observed mismatch to interpretable image-level descriptors that capture waterbody geometry, shoreline shape, and water appearance. Together, these analyses clarify when high

segmentation scores fail to indicate reliable shoreline placement and motivate shoreline-aware evaluation protocols for practical shoreline mapping.

2. Materials and Methodology

2.1. Dataset

This study utilized the Coastal Aerial Imagery Dataset (CAID) [36], a publicly available high-resolution dataset developed for shoreline segmentation in coastal environments. CAID contains 20,689 manually labeled aerial images (0.6–1.0 m spatial resolution) collected from National Agriculture Imagery Program (NAIP) scenes across the Great Lakes region. Each image represents a distinct coastal segment that includes both land and water components, manually annotated to delineate the shoreline boundary between the two classes.

The dataset encompasses diverse shoreline landscapes—including urban, vegetated, beach, rural, rocky, and wetland coasts—as well as variable water quality and surface conditions ranging from calm to wavy. These characteristics make CAID an ideal foundation for evaluating shoreline detection performance and for conducting a detailed assessment of the mismatch between waterbody segmentation accuracy and shoreline positioning under diverse environmental settings.

Based on the CAID benchmark, five deep learning models that demonstrated the highest shoreline positioning accuracy were selected for further analysis. These models include the Asymmetric Non-local Network (ANN) [40], Criss-Cross Network (CCNet) [31], Context Encoding Network (EncNet) [30], Fully Convolutional Network (FCN) [29], and Pyramid Scene Parsing Network (PSPNet) [41]. These five models were trained with 70% of CAID images and validated with 15% of CAID images. A comparative analysis was performed using the 3107 images in the CAID test set to investigate how shoreline positioning deviates from waterbody segmentation across models, to identify potential factors causing these discrepancies, and to assess whether such differences are consistent among architectures or model-specific.

2.2. Coastal Type and Water Surface Condition Identification

To provide contextual interpretation of shoreline detection performance, all images containing a visible land–water interface were manually classified according to their coastal type and water surface condition. Images consisting entirely of land or entirely of water, which occasionally appear in the CAID dataset, were excluded from this analysis to ensure that only valid shoreline scenes were evaluated.

Each qualified image was visually inspected and assigned two categorical attributes describing both the surrounding landscape and the state of the adjacent water surface:

1. **Coastal landscape type.** The coastal zone adjacent to the shoreline was classified into six landscape types representing the dominant geomorphic and land-cover settings that influence shoreline texture and detectability. Examples of each category are shown in Figure 1a–f:
    - *Beach* (Figure 1a), consisting of exposed sand or gravel bars with minimal vegetation;
    - *Rural* (Figure 1b), generally natural shorelines that may contain small human-made features such as local armoring, piers, or minor shoreline structures;
    - *Urban* (Figure 1c), characterized by extensive impervious surfaces and engineered shoreline modifications, including seawalls, harbors, groins, and other large coastal infrastructure;
    - *Rocky* (Figure 1d), defined by shorelines dominated by large rocky cliffs or smaller exposed rock formations rather than sand or vegetation;

- *Vegetated* (Figure 1e), consisting primarily of grasses, shrubs, or trees along the shoreline; and
- *Wetland* (Figure 1f), dominated by emergent aquatic vegetation or saturated soils that transition gradually into open water.

2. **Water surface condition.** The visible water surface was also categorized into two conditions reflecting differences in surface hydrodynamic activity, as illustrated in Figure 2a,b:
   - *Non-wavy* (Figure 2a), where the water surface appears calm, smooth, or reflective with minimal texture variation.
   - *Wavy* (Figure 2b), where wind-induced roughness, wave crests, foam streaks, or other textural patterns indicate more dynamic hydrodynamic conditions.

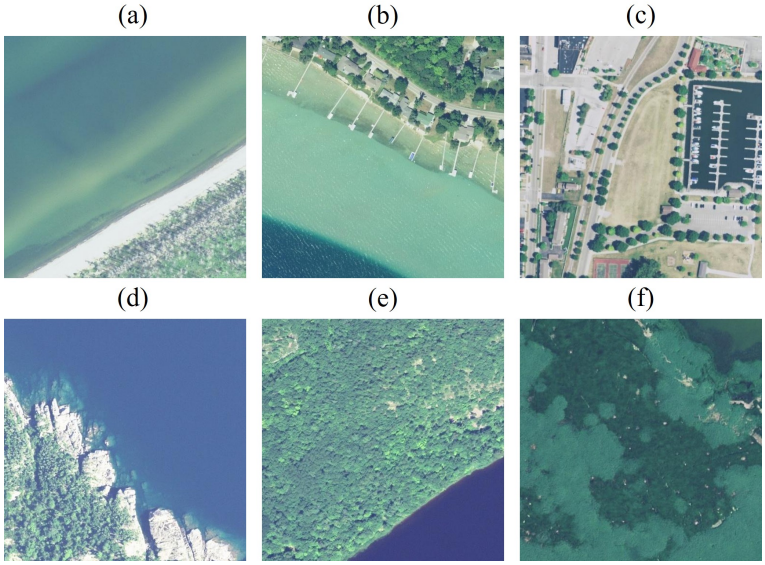


Figure 1. Examples of different coastal landscapes: (**a**) beach, (**b**) rural area, (**c**) urban, (**d**) rocky coast, (**e**) vegetated coast, (**f**) wetland.

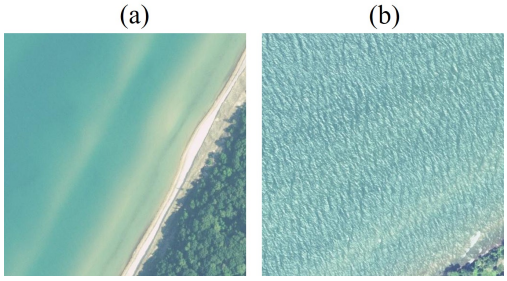


Figure 2. Examples of different water surface conditions: (**a**) non-wavy, (**b**) wavy.

The manual classification was independently performed by four authors to ensure consistent interpretation across the dataset. To evaluate inter-annotator reliability, a subset of 200 images was randomly selected and cross-validated among all annotators. The resulting classifications achieved a 99.5% agreement rate, demonstrating excellent consistency and reproducibility of the labeling process across different observers.

2.3. Image-Level Descriptors

To better understand how shoreline complexity and water appearance influence segmentation performance, we derived three image-level descriptors: mean water hue (MWH), shoreline complexity index (SCI), and shoreline-to-water ratio (SWR). Together, these metrics captured the spectral and geometric attributes of each coastal scene and were used as variables in our downstream analysis.

2.3.1. Mean Water Hue (MWH)

The spectral variation of water surfaces—caused by depth, turbidity, and biological content—was expected to influence model predictions. To quantify this, we computed the average hue of all water-labeled pixels. Each RGB image was converted into HSV color space, and the hue channel was isolated. The MWH for each image was calculated as:

$$\text{MWH} = \frac{1}{N_w} \sum_{i=1}^{N_w} H_i \tag{1}$$

where N_w is the number of water pixels, and H_i is the hue value at pixel i. This measurement captured water color variability, with lower values associated with sediment-rich or algae-affected water, and higher values typical of clear, deeper water.

2.3.2. Shoreline Complexity Index (SCI)

To assess the morphological complexity of shoreline boundaries, we introduced a measure of shoreline length relative to the image's diagonal extent. Binary segmentation masks were processed to extract edge pixels corresponding to the land–water interface, with edges arising solely from image borders explicitly excluded to ensure that only true shoreline boundaries were considered. Let L_s be the count of shoreline pixels, and h and w be the image height and width, respectively. The SCI is defined as:

$$\text{SCI} = \frac{L_s}{\sqrt{h^2 + w^2}} \tag{2}$$

This index increases with shoreline fragmentation and curvature. It is low for smooth or linear coasts and higher for irregular forms such as wetlands, inlets, and estuaries.

2.3.3. Shoreline-to-Water Ratio (SWR)

To capture the density of shoreline pixels relative to the extent of water, we computed the ratio between shoreline pixels and total water pixels:

$$\text{SWR} = \frac{L_s}{N_w} \tag{3}$$

This ratio helps distinguish broad, open water bodies from narrower or more intricately bounded ones. Higher values suggest smaller or more elongated water areas, where shoreline delineation is more error-prone.

These three descriptors provided interpretable summaries of image-level variability and were used throughout our analysis to relate scene characteristics to inconsistencies between waterbody segmentation and shoreline positioning.

2.4. Waterbody Segmentation Metrics

The performance of waterbody segmentation models was evaluated using standard semantic segmentation metrics, including mean Intersection over Union (mIoU) and mean pixel accuracy. These metrics summarize how well the predicted segmentation aligns with ground-truth labels across all classes.

For a given class c, the Intersection over Union (IoU) is defined as:

$$\text{IoU}_c = \frac{|P_c \cap G_c|}{|P_c \cup G_c|}, \tag{4}$$

where P_c and G_c denote the sets of pixels predicted as class c and labeled as class c in the ground truth, respectively. The mean Intersection over Union was computed by averaging over all classes:

$$\text{mIoU} = \frac{1}{C}\sum_{c=1}^{C}\text{IoU}_c, \tag{5}$$

where C is the total number of classes.

Mean pixel accuracy measures the proportion of correctly classified pixels across the entire image and is defined as:

$$\text{Pixel Accuracy} = \frac{\sum_{c=1}^{C}|P_c \cap G_c|}{\sum_{c=1}^{C}|G_c|}. \tag{6}$$

2.5. Shoreline Positioning Metrics

Two different metrics, the mean shoreline intersection ratio and the average Eulerian distance, were calculated to determine the shoreline positioning performance of different deep learning models.

2.5.1. Mean Shoreline Intersection Ratio

To quantitatively evaluate the correspondence between detected and reference shorelines, we introduced the Shoreline Intersection Ratio (SIR). This boundary-based metric was expected to measure the spatial agreement between two shorelines within a local neighborhood and capture positional offsets that are not reflected by conventional pixel-based metrics.

For a given image pair (I_1, I_2), the SIR was calculated as:

$$\text{SIR}_I = \frac{1}{2}\left(\frac{|M(S_1) \cap S_2|}{|S_2|} + \frac{|M(S_2) \cap S_1|}{|S_1|}\right), \tag{7}$$

where S_1 is the detected shoreline and S_2 is the reference shoreline. The operator $M(\cdot)$ denotes a morphological buffering process using an N-neighborhood configuration; in this study, we adopted a 1-neighborhood buffer. This operation expanded each shoreline by one pixel (corresponds to 1 m for pre-2011 images, and 0.6 m for post-2011 images) to accommodate sub-pixel misalignments in high-resolution imagery. In addition, this buffer size was consistent with typical shoreline change magnitudes in the Great Lakes region (approximately 0.48–0.78 m over a two-year NAIP sampling interval [12]). The intersection terms quantified the proportion of shoreline pixels that overlap after buffering, normalized by the total length (pixel count) of each shoreline.

The mean Shore Intersection Ratio (mSIR) was then computed across all image pairs to provide a comprehensive measure of shoreline detection accuracy:

$$\text{mSIR} = \frac{1}{N}\sum_{i=1}^{N}\text{SIR}_{I_i}, \tag{8}$$

where N is the total number of evaluated shoreline pairs. Image pairs lacking valid shoreline pixels in either S_1 or S_2 were excluded from the calculation. Higher mSIR values indicate stronger geometric consistency and closer agreement between detected and reference shorelines.

2.5.2. Average Eulerian Distance

To evaluate the geometric accuracy of the detected shoreline, we computed the Average Eulerian Distance (AED), which measures the mean Euclidean distance between each

detected shoreline pixel and its nearest ground-truth (GT) shoreline pixel. For each image I, the AED is defined as:

$$\mathrm{AED}_I = \frac{1}{|S_1|} \sum_{p \in S_1} \min_{q \in S_2} \|p - q\|_2, \tag{9}$$

where S_1 and S_2 represent the sets of pixels belonging to the detected and reference shorelines, respectively; p and q denote pixel coordinates in the two sets; and $\|p - q\|_2$ is the Euclidean distance between them.

Because all CAID images share the same spatial resolution and dimensions, smaller AED values indicate higher shoreline-detection accuracy. The overall performance across the dataset was assessed using the mean AED:

$$\mathrm{mAED} = \frac{1}{N} \sum_{i=1}^{N} \mathrm{AED}_{I_i}. \tag{10}$$

Following the same criteria used for mSIR, any images lacking valid shoreline pixels in either the detected or reference mask were excluded from AED and mAED computation.

2.6. Correlation Analysis Between Segmentation Metrics and Shoreline Positioning Accuracy

To examine how well conventional waterbody segmentation metrics reflect shoreline positioning accuracy, we evaluated the relationship between pixel accuracy or IoU and the shoreline intersection ratio (SIR) across different environmental settings. Analyses were carried out within each coastal landscape and water-surface category—wavy beach, non-wavy beach, rocky, rural, urban, vegetated, and wetland coasts. To further consider the role of shoreline morphology, additional stratification was applied using the SWR and SCI descriptors, producing separate groups for SWR > 0.01 versus SWR ≤ 0.01 and for SCI > 0.25 versus SCI ≤ 0.25.

Within each of these groups, Spearman's rank correlation, Pearson's correlation, and the coefficient of determination (R^2) were computed to quantify monotonic, linear, and variance-explained relationships between segmentation quality and shoreline alignment. This grouping-based analysis enabled a targeted assessment of how segmentation–shoreline correspondence varies across coastal environments and under differing shoreline morphological conditions.

All analyses used only images with valid detected and reference shorelines to ensure consistency with SIR computation.

2.7. Multivariate Regression Analysis

To identify which shoreline morphology and waterbody characteristics most strongly influence shoreline positioning accuracy, we performed a multivariate regression analysis using the image-level descriptors MWH, SCI, and SWR. This analysis quantifies how variation in water appearance, shoreline curvature, and shoreline-to-water geometry contributes to differences in shoreline detection error. Only images containing valid detected and reference shorelines were included in the analysis.

For each coastal landscape and surface condition (wavy beach, non-wavy beach, rocky, rural, urban, vegetated, and wetland), separate regressions were fitted for SIR_1 and AED. For each deep learning model, we used a linear formulation:

$$\mathrm{Metric} = \beta_0 + \beta_1 \mathrm{SWR} + \beta_2 \mathrm{SCI} + \beta_3 \mathrm{MWH} + \varepsilon, \tag{11}$$

where the response variable is either SIR_1 or AED. Estimated coefficients and associated p-values indicate the relative importance and statistical significance of each descriptor within a given environmental setting.

By fitting the regression models separately for each landscape–surface category, this analysis allows the influence of shoreline geometry, waterbody configuration, and water appearance on shoreline positioning accuracy to be examined under comparable environmental conditions. This stratified design reduces confounding between landscape types and highlights how the effects of SWR, SCI, and MWH vary across different coastal settings.

3. Results

3.1. Inconsistency Between Water Body Segmentation and Shoreline Positioning

3.1.1. Waterbody Segmentation Performance

Across all coastal landscapes, the deep learning models achieved consistently high waterbody segmentation accuracy (Figure 3). Pixel accuracy exceeded 98% in nearly all environments, reaching 99.27–99.32% on beaches, 99.22–99.32% on rocky coasts, and 99.18–99.37% on vegetated coasts. Rural landscapes also showed strong performance (98.99–99.13%), while urban areas exhibited slightly lower accuracy (97.62–98.26%). Wetlands produced the lowest pixel accuracy, ranging from 94.77% (FCN) to 96.42% (ENCNet).

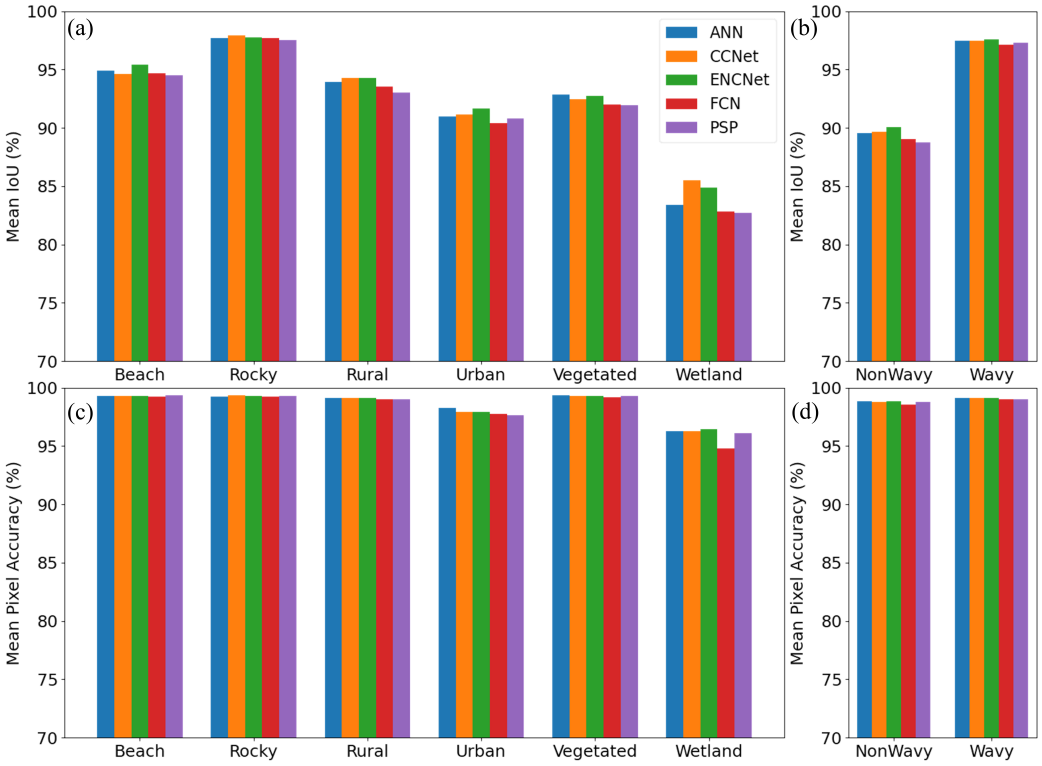


Figure 3. Benchmark performance of five deep learning models on waterbody segmentation metrics across diverse coastal landscapes and water surface conditions. Panels (**a**,**b**) present the mean IoU stratified by landscape type and surface condition. Panels (**c**,**d**) show the corresponding mean pixel accuracy for the same environmental categories.

IoU values reflected similar spatial patterns. Rocky coasts yielded the highest IoU (97.50–97.92%), followed by beaches (94.49–95.40%) and rural (93.02–94.25%) and vegetated settings (91.93–92.84%). Urban coasts showed lower IoU (90.39–91.63%), whereas wetlands were the most challenging, with IoU values between 82.70% and 85.50%.

Differences by water-surface condition were also evident. Wavy scenes produced higher segmentation performance, with IoU values of 97.15–97.61% compared to 88.75–90.07% for non-wavy conditions. Pixel accuracy displayed smaller variation but remained slightly higher for wavy images (99.00–99.14%) than non-wavy images (98.52–98.84%). These results confirm that the models distinguish land and water with high precision, with performance decreasing primarily in structurally complex environments such as wetlands and urban shorelines.

3.1.2. Shoreline Positioning Performance

In contrast to the strong segmentation results, shoreline positioning performance was substantially lower and varied markedly across landscapes (Figure 4). For SIR, beaches yielded the highest correspondence between detected and reference shorelines (47.89–49.63%), followed by rocky (40.98–44.93%), urban (40.28–43.97%), vegetated (40.54–42.51%), and rural coasts (38.75–41.64%). Wetlands exhibited the poorest shoreline agreement, with SIR values of only 27.13–30.41%.

AED results further highlighted spatial discrepancies. Rocky coasts showed the smallest positional offsets (2.71–4.95 pixels), while beaches and vegetated coasts exhibited intermediate errors (6.04–9.48 pixels and 7.16–9.78 pixels, respectively). Rural and urban coasts showed substantially larger deviations, with AED values of 8.91–11.15 pixels (rural) and 15.91–21.24 pixels (urban). Wetlands again presented the greatest difficulty, with AED ranging from 19.60 to 29.55 pixels.

Water-surface condition also influenced shoreline positioning accuracy. Non-wavy scenes produced higher SIR values (42.46–44.85%) than wavy scenes (38.19–40.14%). AED displayed the opposite trend: wavy conditions resulted in smaller positional errors (5.67–7.24 pixels) compared to non-wavy conditions (12.43–16.04 pixels).

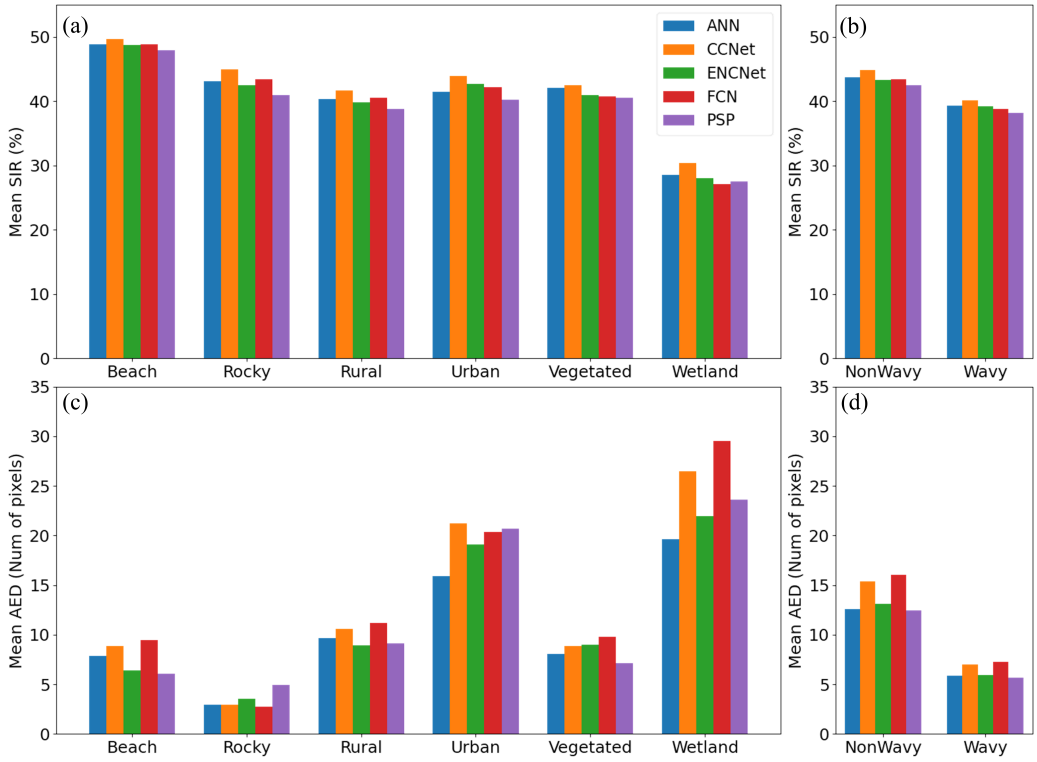


Figure 4. Benchmark performance of five deep learning models on shoreline-specific metrics across diverse coastal landscapes and water surface conditions. Panels (**a**,**b**) present the mean Shoreline Intersection Ratio (SIR) stratified by landscape type and surface condition. Panels (**c**,**d**) show the corresponding mean Average Eulerian Distance (AED) for the same environmental categories.

The divergence between waterbody segmentation accuracy and shoreline positioning accuracy was further illustrated in Figure 5. In the second column, green denoted pixels correctly classified by the model, red indicated ground-truth water misclassified as land, yellow represented ground-truth land misclassified as water, and black corresponded to pixels classified as land in both the prediction and the ground truth. Across the beach, urban, and wetland examples, the ANN model achieved consistently high segmentation performance, with IoU values ranging from 0.949 to 0.995 and pixel accuracy from 0.984 to 0.997.

For the beach example, characterized by a relatively regular shoreline geometry, high segmentation accuracy corresponded to good shoreline positioning. ANN achieved an SIR of 70.3% with an AED of 0.979 pixels, while the best-performing model, PSP, reached an SIR of 82.8% and an AED of 0.655 pixels. In contrast, for the urban shoreline—particularly within the inner-harbor region—localized misclassification of water as land had little influence on segmentation metrics but produced substantial shoreline misalignment. As a result, ANN attained only 57.7% SIR with an AED of 14.488 pixels, and even the best-performing model (CCNet) reached only 60.7% SIR with an AED of 6.271 pixels. A similar pattern was observed in wetlands, where diffuse land–water transitions led to poor shoreline agreement despite high segmentation accuracy: ANN achieved 15.2% SIR with an AED of 7.905 pixels, and the best-performing model (PSP) reached only 20.0% SIR with an AED of 6.128 pixels.

These examples provided a qualitative illustration of the patterns reported in Figrues 3 and 4, showing that high waterbody segmentation accuracy did not necessarily translate into accurate shoreline positioning, particularly in geometrically complex environments such as urban coasts and wetlands.

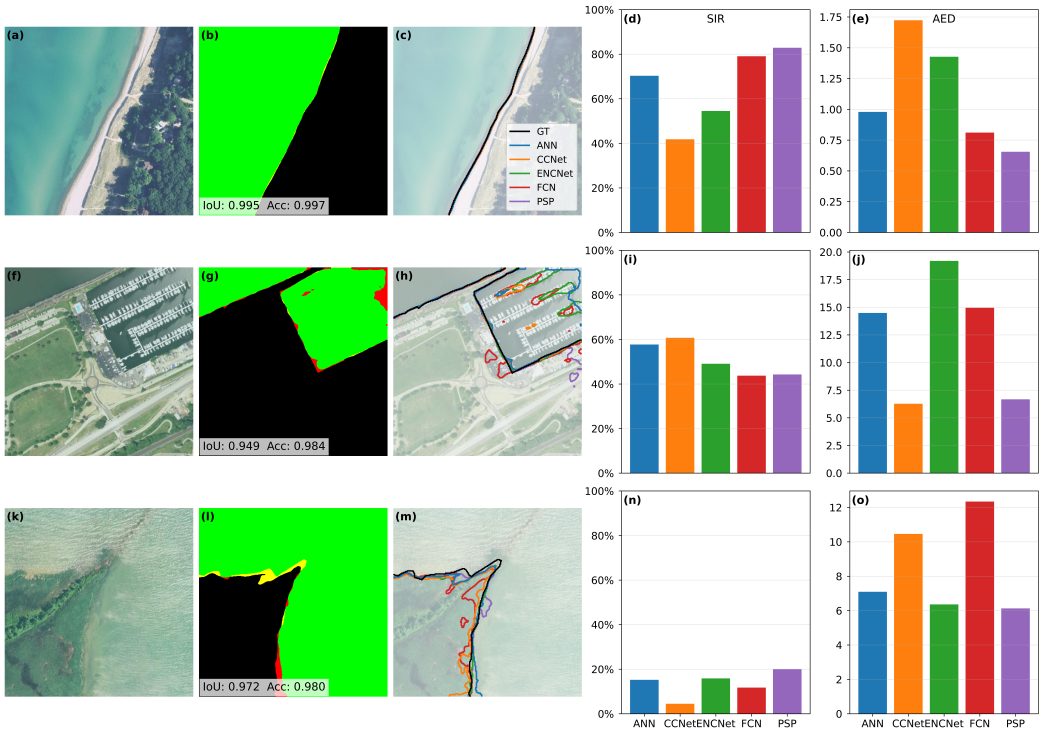


Figure 5. Examples illustrating the inconsistency between waterbody segmentation accuracy and shoreline positioning accuracy. The first row (**a**–**e**) shows a beach shoreline example, the second row (**f**–**j**) an urban shoreline example, and the third row (**k**–**o**) a wetland shoreline example. The first column presents the original RGB image, the second column shows a comparison between ground-truth labels and ANN segmentation results, and the third column displays shoreline positioning results from the five evaluated models relative to the ground truth. The fourth and fifth columns report the Shoreline Intersection Ratio (SIR) and Average Eulerian Distance (AED), respectively, for each model on the corresponding image.

3.2. Correlation Between Waterbody Segmentation Metrics and Shoreline Positioning

3.2.1. Wavy Beach

On wavy beach shorelines (Figure 6), pixel accuracy shows moderate and relatively stable monotonic relationships with SIR. Spearman correlations are similar between SWR > 0.01 and SWR ≤ 0.01 for all models (approximately 0.58–0.75 in both groups), indicating that the rank-order association between pixel accuracy and SIR does not vary strongly with waterbody size under wavy beach conditions. In contrast, Pearson correlations and R^2 remain generally low for all models, with only ENCNet and PSP achieving noticeably higher linear correspondence. SCI stratification produces clearer trends: for most models, SCI > 0.25 yields higher Spearman correlations than SCI ≤ 0.25, although Pearson and R^2 vary considerably across models and SCI groups. Again, ENCNet and PSP show stronger linearity, whereas the remaining models maintain weak linear relationships even when monotonicity improves.

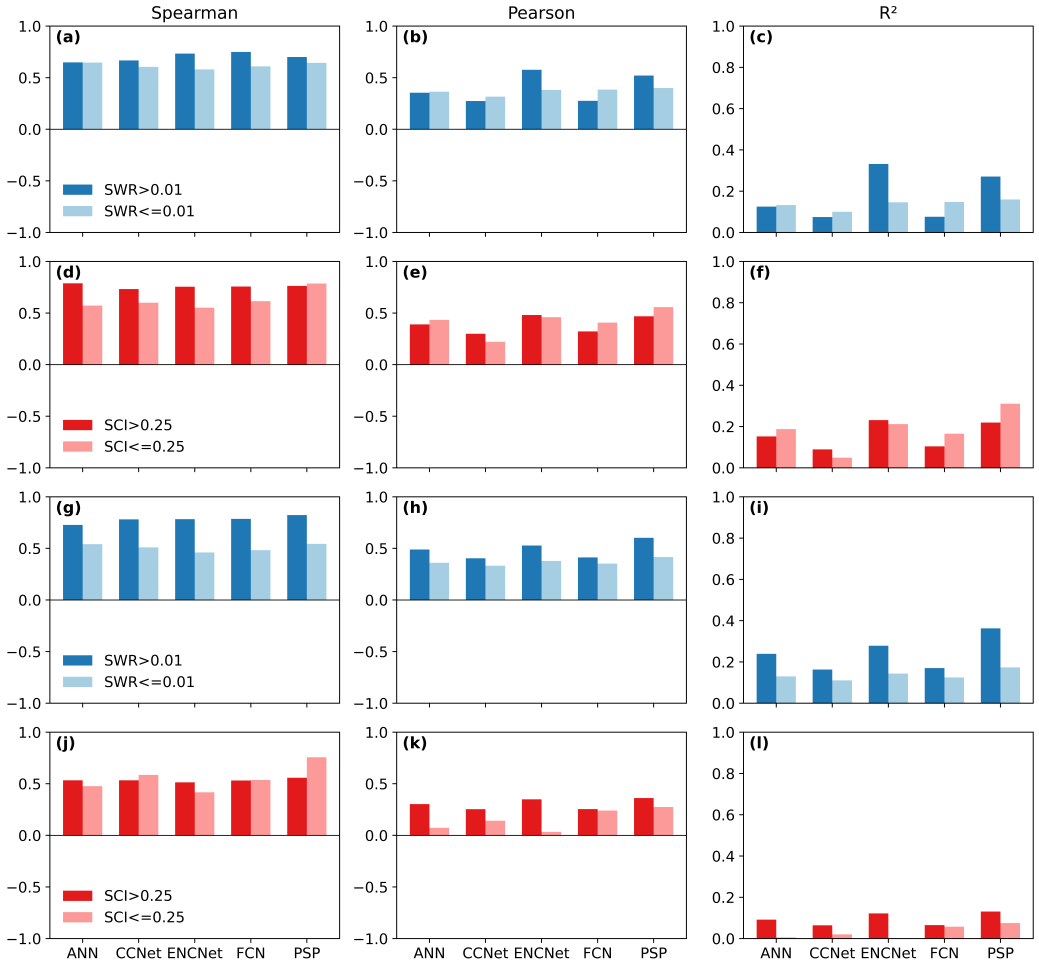


Figure 6. Correlation analysis between waterbody segmentation metrics and shoreline positioning accuracy for wavy beach images. Panels (**a**–**f**) show the relationships between pixel accuracy and the SIR, with points colored by SWR (red) and SCI (blue). Panels (**g**–**l**) show the corresponding relationships using IoU instead of pixel accuracy.

For IoU, SWR exerts a much stronger influence. Under SWR > 0.01, both Spearman and Pearson correlations are consistently higher than those from pixel accuracy (Spearman around 0.73–0.82 across models), and R^2 values increase substantially, especially for ENCNet and PSP. When SWR ≤ 0.01, both monotonicity and linearity deteriorate across models, showing that IoU provides a more faithful reflection of shoreline alignment only when associated waterbodies are relatively small or narrow. For SCI stratification, Spearman values differ only slightly between the two SCI groups and vary by models, Pearson correlations are particularly poor for SCI ≤ 0.25, and R^2 remains low for both SCI ranges.

Overall, Figure 6 indicates that SWR is a primary factor affecting how segmentation metrics relate to SIR on wavy beaches. IoU performs better under SWR > 0.01, showing stronger rank and linear associations than pixel accuracy, while pixel accuracy exhibits similar Spearman correlations across SWR groups. SCI has a clearer effect on pixel accuracy—higher SCI generally produces stronger monotonicity—whereas its influence on IoU varies by model and does not yield strong linearity. Model differences are small for Spearman correlations but higher for Pearson and R^2: ENCNet and PSP show comparatively higher linear correspondence with SIR, whereas ANN, CCNet, and FCN perform more weakly.

3.2.2. Non-Wavy Beach

For non-wavy beaches (Figure 7), pixel accuracy shows moderate monotonic correspondence with SIR across all models. Spearman correlations are slightly higher for SWR ≤ 0.01 than for SWR > 0.01 (increasing from roughly 0.54–0.64 to 0.68–0.72), indicating that pixel accuracy ranks shoreline performance more consistently when waterbodies are large. Pearson correlations and R^2 remain low across all SWR groups, but ENCNet and PSP exhibit noticeably stronger linearity for SWR ≤ 0.01, where R^2 reaches 0.37 and 0.41, respectively. SCI stratification yields comparable Spearman correlations between the two SCI groups (generally 0.73–0.78), with weak Pearson and R^2 values for all models, indicating limited linear predictive value regardless of shoreline complexity.

IoU behaves differently. Under SWR > 0.01, IoU–SIR relationships are much stronger than under SWR ≤ 0.01 for all models: Spearman correlations fall in the 0.74–0.82 range, Pearson correlations increase substantially (0.49–0.60), and R^2 reaches 0.27–0.38. When SWR ≤ 0.01, however, both Spearman and Pearson correlations weaken, reflecting reduced association between IoU and shoreline alignment in large waterbodies. SCI stratification also shows a clearer trend: SCI ≤ 0.25 consistently produces stronger correlations—both Spearman and Pearson—across models, with CCNet, ENCNet, FCN, and PSP showing R^2 values between 0.27 and 0.46. In contrast, SCI > 0.25 yields weaker performance for IoU, particularly in terms of linearity.

Across models, differences in Spearman correlations are modest, but linearity varies substantially: ENCNet, PSP, and FCN generally yield higher Pearson correlations and R^2 than ANN and CCNet. Collectively, Figure 7 shows that pixel accuracy aligns more strongly with SIR when SWR ≤ 0.01, whereas IoU aligns more strongly when SWR > 0.01 and when SCI ≤ 0.25. IoU outperforms pixel accuracy in all linear comparisons, while monotonic performance depends on the stratification conditions.

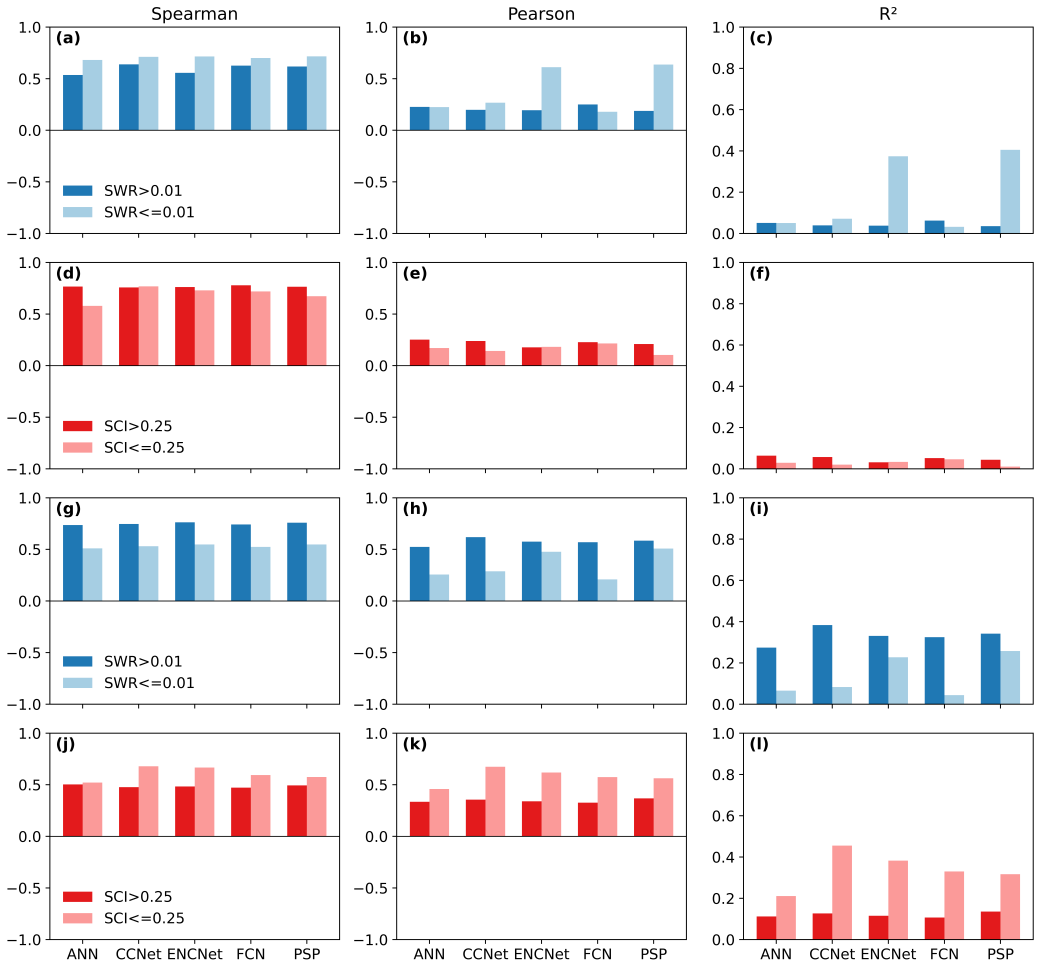


Figure 7. Correlation patterns between segmentation quality and shoreline positioning metrics (SIR) for non-wavy beach images. The top two rows (**a**–**f**) show correlations using pixel accuracy, while the bottom two rows (**g**–**l**) show correlations using IoU. Points are colored by SWR (red) and SCI (blue) to highlight differences in shoreline morphology.

3.2.3. Rocky Coasts

For rocky coasts (Figure 8), pixel accuracy shows a clear dependence on SWR. When SWR > 0.01, all models achieve higher Spearman, Pearson, and R^2 values (Spearman = 0.55–0.66; R^2 up to 0.22), whereas performance deteriorates sharply for SWR ≤ 0.01, where correlations are consistently weak. SCI stratification yields more mixed results: most models maintain modest Spearman correspondence for both SCI groups, but ENCNet drops notably for SCI ≤ 0.25, showing the weakest performance among the architectures.

IoU exhibits the similar SWR trend to SIR. For SWR > 0.01, all models show substantially higher monotonicity and linearity (Spearman = 0.71–0.81; R^2 up to 0.45), whereas correlations become very weak for SWR ≤ 0.01 across all models. Under SCI stratification, models generally perform better for SCI ≤ 0.25, with FCN showing the strongest and most consistent Spearman, Pearson, and R^2 values in this group, while ENCNet remains the exception and does not improve for low SCI.

Overall, Figure 8 illustrates that rocky coasts follow a coherent pattern: (1) both pixel accuracy and IoU correlate well with SIR only when SWR > 0.01, (2) SWR ≤ 0.01 produces uniformly weak associations, and (3) for IoU, SCI ≤ 0.25 yields the stronger relationships than SCI > 0.25, with FCN performing best and ENCNet the most inconsistent.

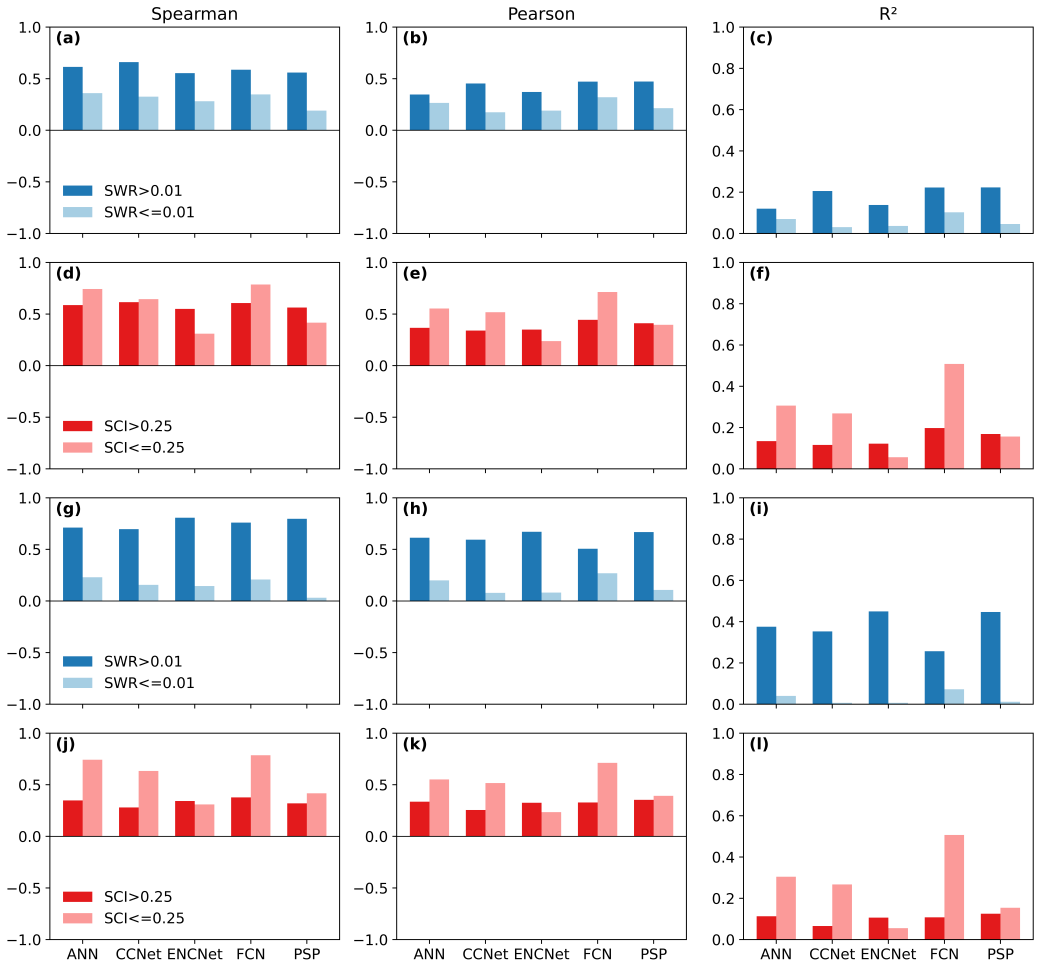


Figure 8. Correlation analysis between segmentation performance and shoreline positioning accuracy (SIR) for the rocky coast category. Panels (**a**–**f**) report correlations with pixel accuracy, and panels (**g**–**l**) report correlations with IoU. Color coding indicates SWR (red) and SCI (blue).

3.2.4. Rural Coasts

For rural coastlines, the correspondence between segmentation metrics and SIR is moderate and relatively stable between the SWR and SCI stratifications. For pixel accuracy, Spearman correlations fall within similar ranges for both SWR groups (approximately 0.54–0.63 for SWR $>$ 0.01 and 0.64–0.71 for SWR $\leq$ 0.01), indicating broadly comparable rank-order relationships regardless of waterbody size. Pearson correlations and R^2 values remain low in all cases, though models such as ENCNet and PSP show noticeably larger linear correlations for SWR $\leq$ 0.01. Under SCI stratification, SCI $>$ 0.25 in general produces higher Spearman values for most models (0.69–0.73), whereas Pearson correlations and R^2 remain small and vary considerably between models and SCI groups, with no consistent pattern favoring either side of the SCI threshold.

IoU shows clearer contrasts across the SWR groups. When SWR $>$ 0.01, IoU exhibits stronger monotonic and linear associations with SIR (Spearman around 0.68–0.73 and R^2 around 0.27–0.33 across models). For SWR $\leq$ 0.01, both Spearman and Pearson correlations are lower across all models, indicating weaker association between IoU and SIR for large rural waterbodies. SCI stratification results in model-dependent behavior: some models (e.g., CCNet, FCN, PSP) show higher Pearson correlations and R^2 for SCI $\leq$ 0.25, while others do not exhibit this pattern. Spearman correlations likewise show mixed behavior across SCI groups, with no uniform trend.

Overall, Figure 9 shows that segmentation–shoreline associations in rural settings vary across models, and the only consistent pattern is that IoU exhibits stronger correspondence with SIR when SWR > 0.01 than when SWR ≤ 0.01. In contrast, SCI stratification does not produce a uniform trend, as correlations differ across models without a consistent direction.

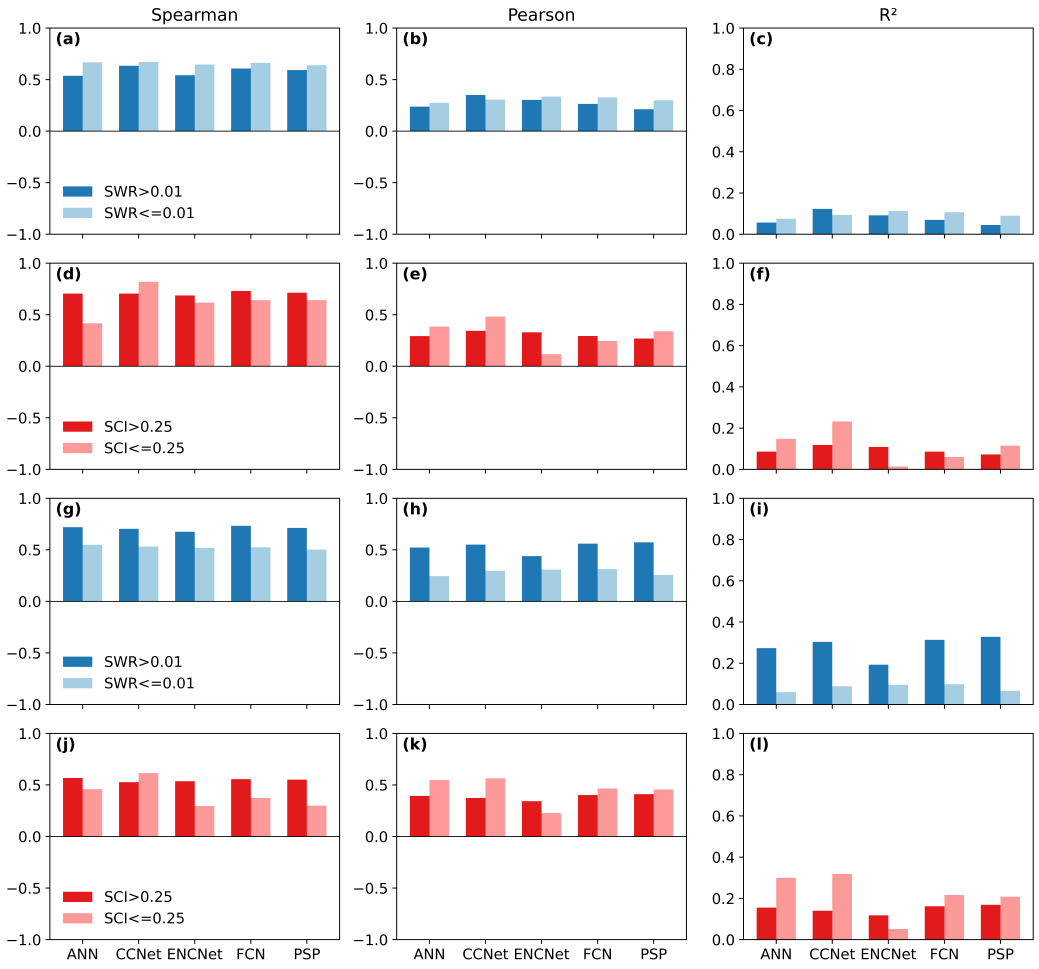


Figure 9. Correlation between segmentation metrics (SIR) and shoreline positioning accuracy for rural coast images. Panels (**a**–**f**) correspond to pixel accuracy; panels (**g**–**l**) correspond to IoU. Points are colored by SWR (red) and SCI (blue).

3.2.5. Urban Shoreline

For urban shorelines, the correspondence between segmentation metrics and SIR displays clear differences across SWR and SCI groupings.

Pixel accuracy shows weaker monotonic associations with SIR when SWR > 0.01 (Spearman: 0.39–0.47) but becomes substantially stronger for SWR ≤ 0.01 (Spearman: 0.55–0.64) across all models, indicating that pixel accuracy reflects shoreline alignment more reliably when waterbodies are large and geometrically simple. Linear relationships remain limited overall, though ENCNet exhibits noticeably higher Pearson correlations and R^2 under SWR ≤ 0.01. SCI stratification produces the strongest contrast in the pixel-accuracy results: the SCI ≤ 0.25 group yields much higher Spearman, Pearson, and R^2 values for all models (e.g., Spearman up to 0.93; R^2 up to 0.58), demonstrating that pixel accuracy corresponds more closely with SIR along short or relatively simple shorelines. In contrast, IoU displays the opposite behavior with respect to SWR. Across all models, SWR > 0.01 produces stronger association (Spearman: 0.66–0.71, R^2: 0.32–0.38) than SWR ≤ 0.01, where correlations decline in both monotonic and linear terms. SCI stratification, however, yields no uniform pattern: some models (e.g., ANN, FCN) show higher

Pearson correlations and R^2 for SCI > 0.25, whereas others (e.g., CCNet) show stronger relationships under SCI ≤ 0.25. This indicates that SCI is a model-dependent indicator for IoU–SIR correspondence in urban environments.

Collectively, Figure 10 shows that pixel accuracy aligns more closely with SIR when waterbodies are large or shorelines are short, whereas IoU provides a stronger indication of shoreline agreement when waterbodies are narrow. No consistent SCI-based pattern emerges for IoU, and model-to-model differences are more pronounced for both linear correlations and rank-based measures.

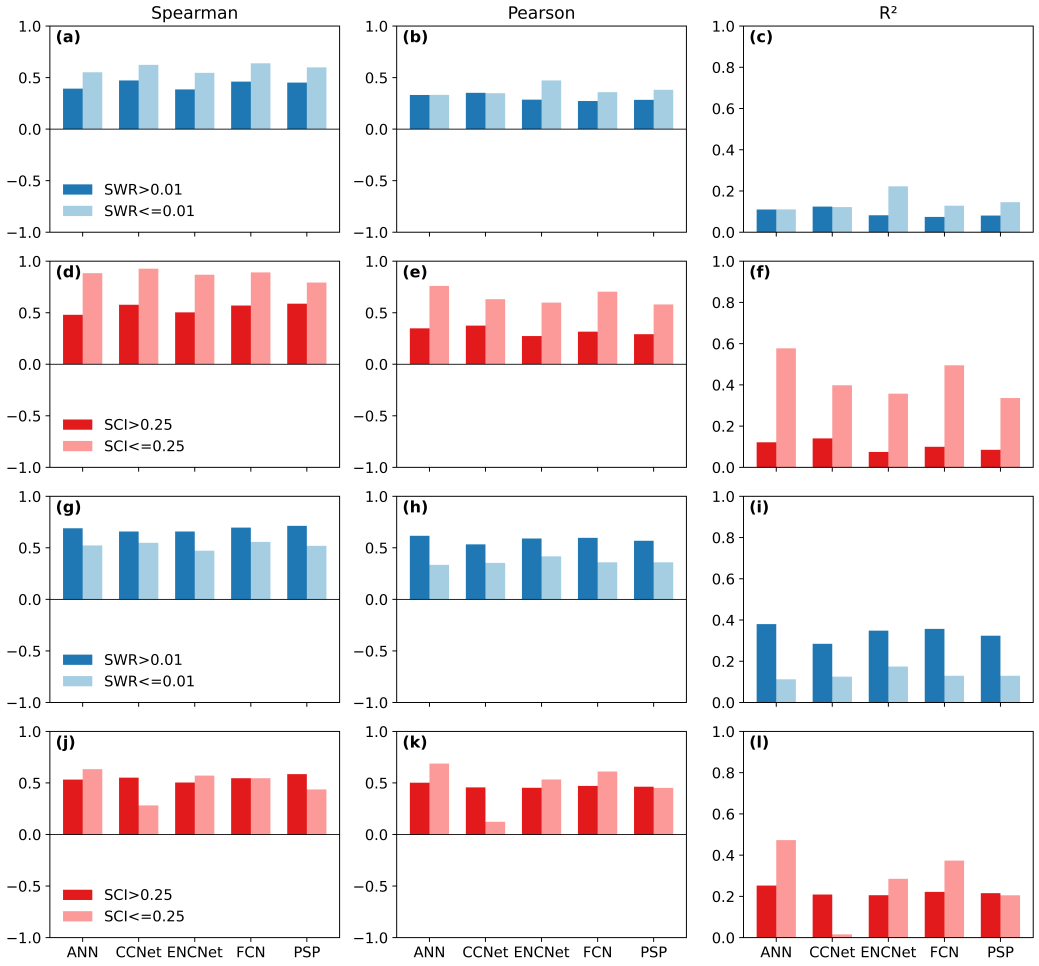


Figure 10. Correlation results for urban environments. The first two rows (**a**–**f**) show how pixel accuracy relates to shoreline positioning metrics, with SWR (red) and SCI (blue) serving as morphological indicators. The last two rows (**g**–**l**) show the same analysis using IoU.

3.2.6. Vegetated Shoreline

For vegetated coasts, the correspondence between segmentation metrics and SIR depends strongly on SWR grouping and varies across models for SCI grouping. For pixel accuracy, all models show consistently weaker associations with SIR when SWR > 0.01 than when SWR ≤ 0.01. This pattern holds for Spearman, Pearson, and R^2, with the SWR ≤ 0.01 group reaching Spearman values around 0.64–0.65 and noticeably higher linearity. SCI stratification shows that Spearman correlations are generally higher for SCI > 0.25, but Pearson correlations and R^2 vary widely among models, indicating no consistent direction for the linear relationship.

IoU exhibits the opposite SWR pattern. Across all models, SWR > 0.01 yields substantially stronger monotonic and linear associations with SIR (Spearman: 0.67–0.77; R^2: 0.31–0.38) than SWR ≤ 0.01, where correlations weaken for every model. SCI grouping

is also consistent among models: SCI $\leq$ 0.25 produces stronger Pearson correlations and higher R^2 across nearly all models, while Spearman values also tend to be higher for the SCI $\leq$ 0.25 group.

Overall, Figure 11 shows clear and opposite SWR effects for pixel accuracy and IoU in vegetated landscapes: pixel accuracy aligns more closely with SIR when waterbodies are large (SWR $\leq$ 0.01), whereas IoU aligns more closely when waterbodies are small or narrow (SWR $>$ 0.01). SCI grouping influences both metrics, but the direction is model-dependent for pixel accuracy and more consistent for IoU, where shorter shorelines (SCI $\leq$ 0.25) show stronger association.

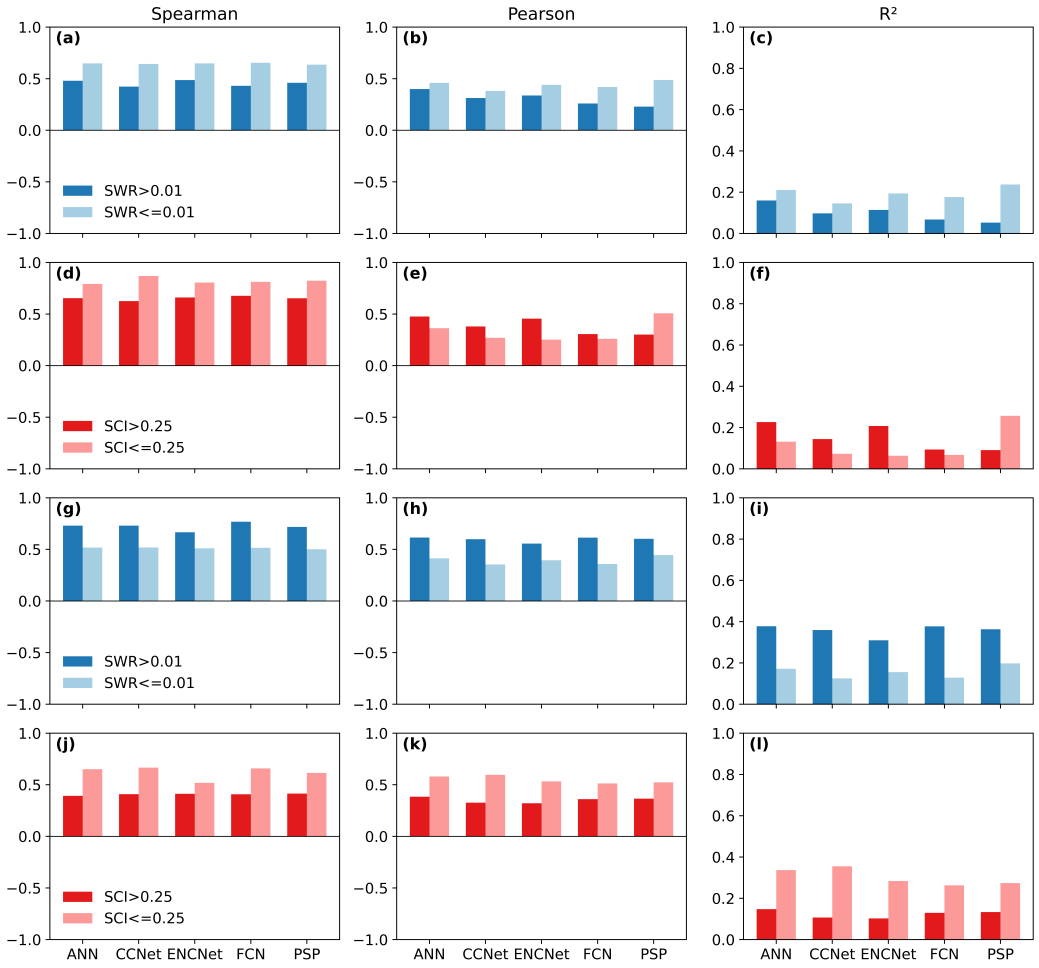


Figure 11. Correlation relationships between segmentation quality and shoreline positioning metrics for vegetated coasts. Panels (**a**–**f**) examine pixel accuracy, while panels (**g**–**l**) examine IoU. SWR (red) and SCI (blue) highlight the influence of shoreline morphology on these relationships.

3.2.7. Wetlands

Wetlands exhibit the weakest performance for both waterbody segmentation (Figure 3) and shoreline positioning (Figure 4), and have highly inconsistent associations between segmentation metrics and SIR. For pixel accuracy, SWR $\geq$ 0.01 produces slightly higher Spearman correlations (0.52–0.68) than SWR $<$ 0.01 (0.32–0.53), although Pearson correlations and R^2 show no consistent pattern and vary considerably across models. SCI stratification introduces even larger variability: several models show moderate Spearman values when SCI $>$ 0.25, but relationships for SCI $\leq$ 0.25 differ strongly by architecture. Notably, ENCNet produces negative Spearman and Pearson correlations within the SCI $\leq$ 0.25 group, indicating that higher pixel accuracy may coincide with poorer shoreline alignment

in wetlands. Other models maintain weak but positive associations, though R^2 remains low across both SCI groups.

IoU shows clearer differences under SWR grouping. For all architectures, SWR ≥ 0.01 yields substantially stronger associations with SIR—higher Spearman correlations (0.68–0.75), higher Pearson correlations, and R^2 values around 0.31–0.38—whereas correlations decrease uniformly when SWR < 0.01. SCI stratification again produces model-dependent patterns: some architectures (e.g., CCNet) obtain higher Pearson correlations and R^2 in the SCI ≤ 0.25 group, while others (especially ENCNet) show near-zero or negative values. Overall, SCI does not provide a uniform directional effect for either monotonic or linear relationships.

Overall, Figure 12 shows that wetland settings present the most challenging shoreline conditions, with weak and highly variable segmentation–shoreline associations for both pixel accuracy and IoU. SWR stratification produces a clear directional pattern only for IoU—stronger correspondence when SWR ≥ 0.01—whereas SCI stratification yields model-dependent responses for both metrics, including negative correlations for ENCNet under SCI ≤ 0.25.

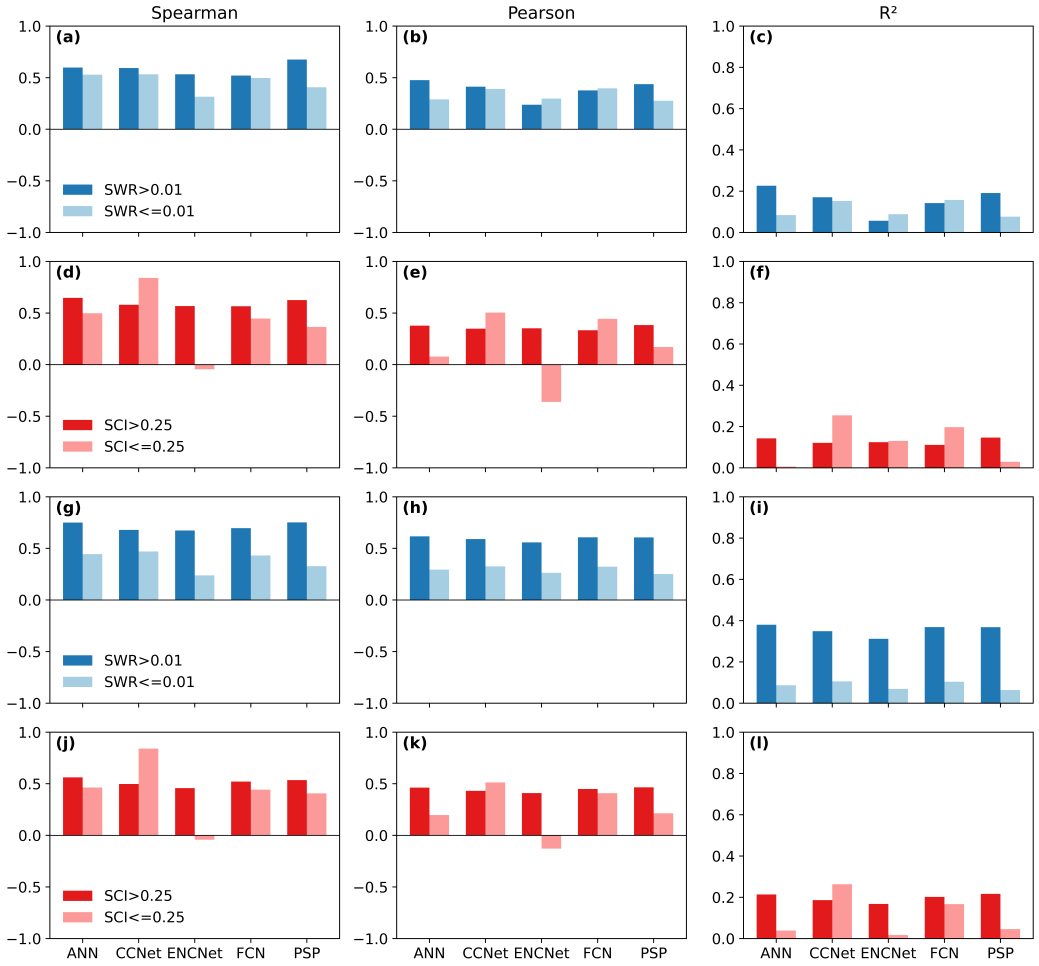


Figure 12. Correlation analysis between waterbody segmentation metrics and shoreline positioning accuracy for wetland scenes. Panels (**a**–**f**) illustrate the relationships between pixel accuracy and the SIR, with points colored by SWR (red) and SCI (blue). Panels (**g**–**l**) present the corresponding relationships using IoU.

3.3. Regression Analysis of Shoreline Positioning Accuracy

3.3.1. Regression Analysis for SIR

The multivariate regressions reveal clear and systematic differences in how shoreline morphology and water appearance influence SIR across landscapes and models (Table 1). Among the three predictors, SWR consistently exhibits the strongest and most statistically significant effects. For all non-wavy beach, rural, urban, vegetated, and wetland landscapes, SWR shows a robust negative association with SIR (coefficients typically between −0.6 and −2.5, $p < 0.0001$), indicating that images containing smaller or more elongated waterbodies (higher SWR) lead to substantially reduced shoreline alignment accuracy across all models. This pattern is weakest in rocky coasts, where SWR coefficients are small and not statistically significant, and in wavy beach settings where only PSP shows a significant SWR effect.

Table 1. Regression coefficients of SIR with SWR, SCI, and MWH across coastal landscapes.

Landscape + Surface	Model	SWR Coef	SWR *p*-Value	SCI Coef	SCI *p*-Value	MWH Coef	MWH *p*-Value
Wavy Beach	ANN	0.1271	0.9280	0.0129	0.6988	0.0010	0.2780
	CCNet	−1.5352	0.2743	0.0338	0.3045	0.0013	0.1523
	ENCNet	0.1466	0.9182	0.0393	0.2486	0.0011	0.2505
	FCN	−1.4974	0.2961	0.0540	0.1144	0.0005	0.6122
	PSP	−2.8901	0.0437	0.0364	0.2847	0.0009	0.3526
Non-wavy Beach	ANN	−1.7019	0.0000	0.0286	0.2193	0.0005	0.3648
	CCNet	−1.7957	0.0000	0.0058	0.8052	0.0005	0.3490
	ENCNet	−2.0830	0.0000	0.0074	0.7469	0.0007	0.2258
	FCN	−1.5719	0.0000	−0.0009	0.9687	0.0006	0.2549
	PSP	−2.4767	0.0000	0.0160	0.4952	0.0005	0.3943
Rocky	ANN	0.4950	0.8219	−0.0004	0.9921	−0.0004	0.8661
	CCNet	1.2746	0.5129	−0.0042	0.9015	−0.0000	0.9942
	ENCNet	0.2083	0.9195	0.0262	0.4684	0.0001	0.9502
	FCN	0.8543	0.6767	−0.0195	0.5865	0.0010	0.6378
	PSP	−0.3185	0.8701	0.0166	0.6215	−0.0005	0.8131
Rural	ANN	−2.2789	0.0000	−0.0087	0.5582	0.0020	0.0000
	CCNet	−1.7839	0.0000	−0.0213	0.1596	0.0021	0.0000
	ENCNet	−1.4881	0.0000	−0.0050	0.7424	0.0022	0.0000
	FCN	−2.3653	0.0000	−0.0142	0.3566	0.0021	0.0000
	PSP	−2.1908	0.0000	−0.0036	0.8072	0.0019	0.0000
Urban	ANN	−2.3530	0.0000	0.0191	0.0890	0.0008	0.1524
	CCNet	−1.4197	0.0019	0.0033	0.7810	0.0012	0.0487
	ENCNet	−1.9673	0.0000	0.0084	0.4464	0.0010	0.0829
	FCN	−2.2735	0.0000	0.0044	0.7150	0.0014	0.0303
	PSP	−2.2596	0.0000	0.0012	0.9226	0.0012	0.0649
Vegetated	ANN	−1.2591	0.0000	0.0177	0.2265	0.0005	0.0858
	CCNet	−1.5349	0.0000	0.0153	0.2939	0.0005	0.1026
	ENCNet	−0.9379	0.0001	0.0132	0.3663	0.0009	0.0033
	FCN	−1.5392	0.0000	0.0196	0.1814	0.0005	0.0831
	PSP	−1.3147	0.0000	0.0173	0.2400	0.0009	0.0057
Wetland	ANN	−0.7929	0.0201	0.0087	0.5245	0.0012	0.0391
	CCNet	−0.9135	0.0004	−0.0002	0.9853	0.0017	0.0012
	ENCNet	−0.5620	0.0245	0.0242	0.0516	0.0012	0.0213
	FCN	−0.9280	0.0001	0.0125	0.3004	0.0017	0.0008
	PSP	−0.9067	0.0024	0.0124	0.3490	0.0013	0.0236

In contrast, SCI exerts only weak and highly variable influences across landscapes. Its correlations are typically small and lack a consistent trend, suggesting that shoreline curvature is not a reliable determinant of SIR. Accordingly, shoreline complexity plays a comparatively minor role relative to the geometric constraint imposed by the shoreline-to-water ratio.

MWH shows weak but statistically detectable associations with SIR, and its influence is highly landscape-dependent. In rural, wetland, and several vegetated models, MWH exhibits small positive coefficients with significant p-values, indicating that water-appearance variation contributes modestly to shoreline error in settings where color or tonal heterogeneity is common. In contrast, coefficients in wavy beach, non-wavy beach, urban, and especially rocky environments are close to zero and often insignificant, and some models even yield negative estimates. This pattern indicates that MWH does not exert a uniform directional effect on SIR across landscapes; instead, its contribution is minor and inconsistent, varying with the environmental context and model architecture rather than reflecting a generalizable physical relationship.

Taken together, these regression results indicate that the primary factor governing shoreline-positioning performance is the relative density of shoreline pixels (SWR), which reflects underlying waterbody geometry. Shoreline curvature and water hue exert secondary and context-dependent effects. The consistency of these patterns across all five deep learning models further demonstrates that these relationships arise from scene geometry itself rather than from model-specific behavior.

3.3.2. Regression Analysis for AED

Regression results for AED exhibit a markedly different pattern from those of SIR, with substantially stronger and more consistent effects across predictors (Table 2). SWR is again the dominant factor: nearly all landscapes and models show strong positive and statistically significant SWR coefficients, often ranging from 300 to 900 in non-wavy beach, rural, urban, vegetated, and wetland settings ($p < 0.0001$). These large positive values indicate that AED increases substantially when waterbodies become smaller or more elongated (higher SWR), meaning that shorelines are positioned farther from their ground-truth locations in scenes dominated by narrow channels, ponds, or fragmented water features. Only rocky coasts show weak and non-significant SWR effects, reflecting their reduced sensitivity to shoreline-to-water ratios. SCI plays a more prominent role for AED than for SIR, particularly in urban and wetland environments. Across all five models, SCI coefficients are strongly negative and highly significant in urban landscapes ($p < 0.001$), indicating that greater shoreline complexity—associated with hardened structures, curved embayments, or harbor outlines—is systematically linked to larger AED values. Wetlands show a similar pattern, with significant negative SCI coefficients for most models ($p < 0.05$). These negative effects reflect that AED, as a distance-based metric, is more sensitive to errors along tortuous or highly curved boundaries, where small segmentation deviations can cause large geometric offsets.

MWH exhibits generally weak and inconsistent effects across landscapes, with occasional significance in wavy beach and non-wavy beach settings for CCNet, ENCNet, and PSP, but no coherent directional trend. This suggests that water appearance (color, turbidity, or hue variability) plays little role in governing distance-based shoreline detection errors compared to the dominant geometric constraints imposed by SWR and SCI.

Taken together, the AED regressions show that distance-based shoreline positioning errors are strongly governed by waterbody geometry (SWR) and shoreline curvature (SCI), with consistently larger effects than those observed for SIR. This highlights AED's sensitivity to both the spatial configuration of the waterbody and the complexity of the shoreline boundary, particularly in urban and wetland environments where shoreline patterns could be highly irregular.

Table 2. Regression coefficients of AED with SWR, SCI, and MWH across coastal landscapes.

Landscape + Surface	Model	SWR Coef	SWR *p*-Value	SCI Coef	SCI *p*-Value	MWH Coef	MWH *p*-Value
Wavy Beach	ANN	109.8028	0.0400	−0.5488	0.6638	0.0286	0.4106
	CCNet	323.6822	0.0110	0.9878	0.7395	0.1978	0.0167
	ENCNet	250.9076	0.0070	−1.1840	0.5913	0.1261	0.0381
	FCN	462.9885	0.0027	−0.1056	0.9770	0.3041	0.0027
	PSP	9.3617	0.8808	−0.0824	0.9558	0.0287	0.4830
Non-wavy Beach	ANN	739.8679	0.0000	−1.2813	0.7350	0.1492	0.1067
	CCNet	369.4981	0.0000	−6.5047	0.1152	0.1584	0.1162
	ENCNet	362.8047	0.0000	−5.8459	0.0820	0.1963	0.0166
	FCN	426.6709	0.0000	−6.8011	0.1446	0.1240	0.2794
	PSP	538.0175	0.0000	−3.4280	0.3076	0.2261	0.0061
Rocky	ANN	−7.0038	0.8391	0.8371	0.1683	−0.0103	0.7632
	CCNet	−13.9612	0.7519	0.7887	0.3016	−0.0035	0.9369
	ENCNet	141.5239	0.1902	−1.3275	0.4813	0.0384	0.7196
	FCN	2.7467	0.9131	0.9777	0.0288	−0.0045	0.8576
	PSP	316.7443	0.1262	−2.8802	0.4174	0.1231	0.5455
Rural	ANN	628.7102	0.0000	−4.5661	0.1079	−0.0143	0.8360
	CCNet	655.7743	0.0000	−5.0009	0.1177	−0.0364	0.6388
	ENCNet	699.4000	0.0000	−5.4797	0.0636	−0.0630	0.3808
	FCN	698.0021	0.0000	−5.1226	0.1417	−0.0702	0.4068
	PSP	619.5448	0.0000	−5.2130	0.0746	−0.0469	0.5130
Urban	ANN	1017.5658	0.0000	−9.3052	0.0004	−0.0544	0.6832
	CCNet	765.9771	0.0000	−12.6424	0.0001	0.2049	0.2249
	ENCNet	787.9670	0.0000	−11.6164	0.0003	0.1411	0.3901
	FCN	955.7987	0.0000	−12.1290	0.0001	0.1299	0.4206
	PSP	944.6132	0.0000	−13.4535	0.0001	0.1523	0.3744
Vegetated	ANN	494.0418	0.0000	−2.6220	0.3591	−0.0826	0.1806
	CCNet	358.3424	0.0000	−3.2167	0.2423	0.0387	0.5140
	ENCNet	572.3588	0.0000	−2.9353	0.3006	−0.1165	0.0574
	FCN	343.8342	0.0000	−4.6628	0.1106	−0.0283	0.6516
	PSP	235.1188	0.0000	−3.0592	0.1676	−0.0457	0.3378
Wetland	ANN	178.2994	0.0215	−6.1915	0.0476	−0.0765	0.5482
	CCNet	569.1739	0.0000	−13.5736	0.0011	0.0665	0.6941
	ENCNet	599.0224	0.0000	−9.8605	0.0077	0.0033	0.9827
	FCN	549.4707	0.0000	−13.8346	0.0008	−0.0165	0.9215
	PSP	761.6591	0.0000	−12.1245	0.0004	−0.0406	0.7752

4. Discussion

4.1. Summary and Interpretation of Key Findings

This study demonstrates a systematic mismatch between waterbody segmentation accuracy and shoreline positioning accuracy in deep learning–based shoreline mapping. Although pixel accuracy and IoU are consistently high across landscapes, shoreline-specific metrics indicate substantially lower agreement with reference shorelines, particularly in

wetlands, urban environments, and scenes dominated by large or homogeneous waterbodies. These findings confirm that strong image-level segmentation performance does not necessarily translate into reliable shoreline delineation.

The dominance of the shoreline-to-water ratio (SWR) reflects a fundamental geometric constraint in evaluating shorelines derived from classified land–water masks, as SWR controls how strongly area-based segmentation metrics respond to shoreline misalignment. In high-SWR settings—such as small water bodies, narrow rivers, fragmented wetlands, or highly indented urban shorelines—the land–water boundary constitutes a relatively large fraction of the waterbody extent, so small shoreline shifts convert a non-negligible number of boundary pixels and produce noticeable changes in IoU. In contrast, in low-SWR environments characterized by large or compact waterbodies, comparable boundary misalignments affect only a thin boundary band relative to the interior area, allowing pixel accuracy and IoU to remain high even when the shoreline is substantially displaced. Figure 13 illustrates this geometric alignment mechanism by considering a controlled shoreline displacement envelope: for the same range of landward and waterward shifts (±10 px), low-SWR configurations exhibit only minor changes in area-based metrics, whereas high-SWR configurations show pronounced IoU degradation under moderate displacement, while pixel accuracy remains comparatively stable. High SWR can also coincide with intrinsically more meandered or morphologically complex shorelines, which may be inherently more difficult to delineate and can contribute to lower SIR and higher AED in some settings. However, this effect is secondary to the geometric mechanism emphasized here and is not the primary focus of this study. Rather, the key implication is that SWR governs whether area-based segmentation metrics meaningfully reflect shoreline positioning accuracy. Consistent with this interpretation, IoU is generally more responsive to shoreline misalignment than pixel accuracy when boundary pixels represent a non-negligible fraction of the segmented waterbody, whereas both metrics become weak proxies for shoreline accuracy in low-SWR scenes where interior pixels dominate error statistics. Shoreline complexity (SCI) exerts a weaker and more context-dependent influence, contributing more clearly to distance-based errors in urban and wetland settings but showing no consistent effect on boundary overlap.

Furthermore, although the five deep learning architectures exhibit different segmentation and shoreline positioning performance across sites, no consistent architecture-specific pattern emerges. Performance differences among models vary with shoreline geometry and landscape setting, and the relative ranking of models changes across metrics and environments. These observations indicate that, within the tested frameworks, shoreline positioning accuracy is not governed primarily by network architecture, but is instead strongly influenced by scene characteristics and by the inherent decoupling between area-based segmentation accuracy and boundary localization.

Overall, these findings show that geometric characteristics of coastal scenes strongly condition the relationship between segmentation metrics and shoreline positioning accuracy. This underscores the limitations of conventional segmentation metrics for shoreline-focused applications and highlights the importance of incorporating shoreline-specific evaluation when deep learning models are used for coastal analysis.

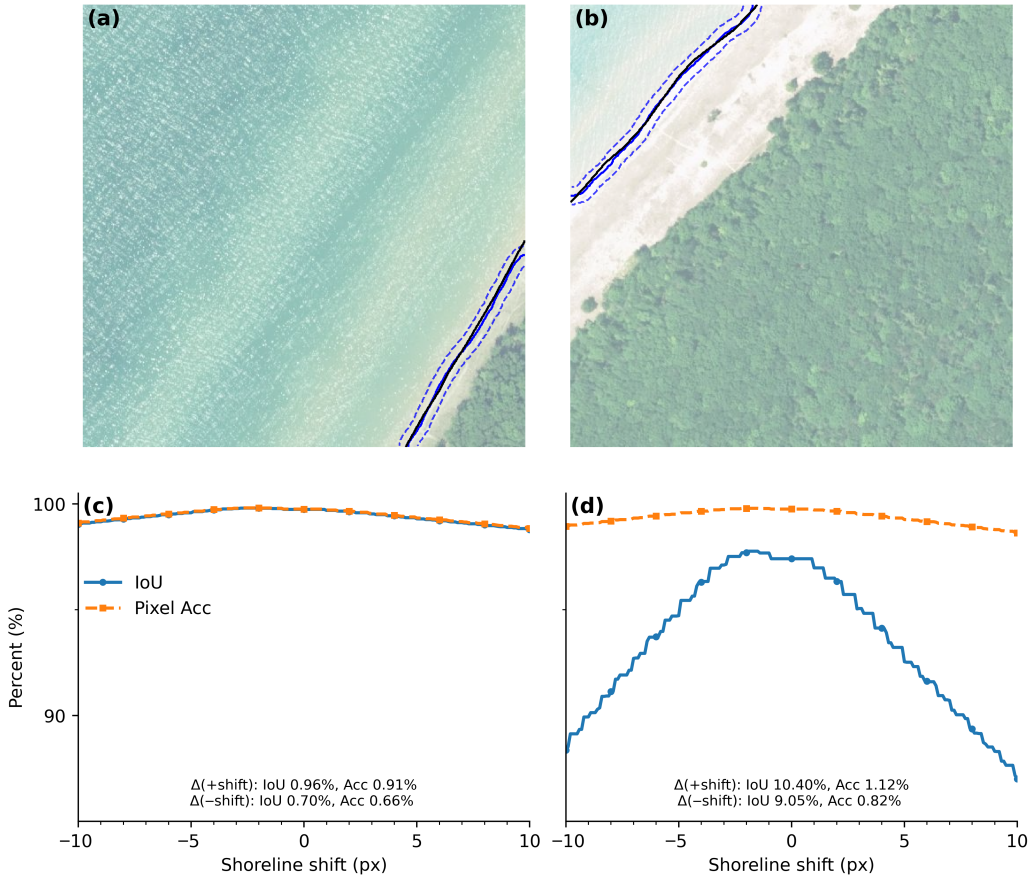


Figure 13. Illustration of the shoreline-to-water ratio (SWR) effect. Panels (**a**,**b**) show representative shoreline scenes with low and high SWR, respectively, with the reference shoreline (black), predicted shoreline (blue), and a displacement range of ±10 px (blue dashed). Panels (**c**,**d**) show how Intersection over Union (IoU) and pixel accuracy (%) change with shoreline shift distance (−10 to +10 px) for the low- and high-SWR cases.

4.2. Comparison with Existing Shoreline Mapping Approaches

Although this study evaluates shoreline products derived from deep learning–based image segmentation, the observed inconsistency between segmentation accuracy and shoreline positioning accuracy reflects a broader limitation in how shoreline mapping approaches are commonly evaluated. Spectral index–based methods (e.g., NDWI [42], MNDWI [43], AWEI [44], DDWI [20]), machine-learning classifiers [45,46], and deep learning segmentation models [47,48] typically derive shorelines from classified land–water representations and assess performance primarily using area-based metrics such as pixel accuracy, confusion-matrix statistics, or IoU. Consequently, high segmentation scores do not necessarily imply accurate shoreline placement. For example, Erdem et al. [47] reported IoU values exceeding 99% using an ensemble U-Net framework, and Dang et al. [48] achieved segmentation accuracies near 98%, yet neither study quantified shoreline positioning accuracy.

This evaluation pattern also appears in studies that directly compare index-based methods with machine-learning and deep learning approaches. Several studies [23,49] compared NDWI and MNDWI thresholding with supervised classification models (e.g., SVM, U-Net, pixel-based CNN, and pixel-based DNN) for water extraction from medium-resolution satellite imagery, such as Landsat and WorldView-2, and reported overall classification accuracy or training and validation performance. While these results are useful for assessing water–land separability across methods, they do not indicate whether the derived land–water boundary is spatially aligned with the true shoreline.

A smaller set of studies has moved beyond purely area-based scores by incorporating distance- or edge-based evaluations, for example by extracting explicit shorelines and reporting positional offsets or shoreline-change statistics, or by emphasizing boundary quality via edge-aware losses and boundary-focused architectures. Kumar et al. [45] evaluated shoreline change using machine-learning classifiers with explicit edge delineation. Similarly, recent deep learning studies have introduced boundary-aware losses or architectures to enhance edge sharpness, including Sobel-edge loss functions [50], boundary-aware convolution [51], and bi-directional cascade networks [52]. However, these approaches generally treat boundaries in a generic sense and rarely distinguish true shorelines (i.e., the land–open-water interface) from other water–land boundaries, such as those associated with narrow streams or small ponds. Moreover, boundary errors are often reported as standalone outcomes and interpreted primarily as technical refinements, without quantitatively analyzing when and why segmentation accuracy and shoreline positioning accuracy diverge, or how this divergence depends on shoreline geometry, water-surface condition, or landscape context.

To clarify these differences in evaluation practice, Table 3 summarizes common shoreline mapping paradigms in terms of how shorelines are derived, which metrics are reported, and whether the association between area-based segmentation accuracy and shoreline positioning accuracy is explicitly examined. Across index-based, machine-learning, and deep learning segmentation approaches, evaluation remains dominated by area-based metrics, which can be weakly sensitive to shoreline displacement. Boundary-aware optimization improves local edge quality, but still does not explain when segmentation scores are unreliable surrogates for shoreline positioning.

The geometric interpretation developed in this study provides a unifying explanation for these observations. The shoreline-to-water ratio (SWR) governs how sensitive shoreline position is to boundary errors relative to area-based segmentation metrics. When SWR is small, the shoreline represents only a minor fraction of the waterbody area, so even large shoreline offsets contribute negligibly to pixel accuracy or IoU. This geometric effect is independent of the extraction method, whether based on spectral thresholding [42], machine-learning classification [45], or deep learning segmentation [47]. By jointly evaluating segmentation metrics (IoU) and shoreline-specific measures (SIR and AED), and explicitly linking their divergence to shoreline geometry, water-surface condition, and landscape type, this study identifies the conditions under which segmentation metrics fail to predict shoreline accuracy and helps explain why this limitation persists across multiple image segmentation model architectures.

Table 3. Comparison of shoreline evaluation paradigms.

Evaluation Paradigm	Shoreline Derivation	Reported Metrics	What Is Measured	Displacement Consideration	Area–Shoreline Accuracy Association Deep-Dived
Index-based	Thresholded spectral index mask	Pixel accuracy, IoU	Area agreement	Low–moderate	No
Machine-learning	Classified land–water mask	Pixel accuracy, IoU	Area agreement	Low–moderate	No
Deep-learning	Segmentation mask (CNN-based)	Pixel accuracy, IoU	Area agreement	Low–moderate	No

Table 3. *Cont.*

Evaluation Paradigm	Shoreline Derivation	Reported Metrics	What Is Measured	Displacement Consideration	Area–Shoreline Accuracy Association Deep-Dived
Boundary-aware optimization	Boundary-emphasized training or post-processing	Pixel accuracy, IoU, Edge error	Area agreement + boundary alignment	Moderate–high	No
This study	Segmentation-derived shoreline vs. reference	Pixel accuracy, IoU, SIR, AED	Area agreement + shoreline alignment	High	Yes

4.3. Limitations

While this study provides a systematic assessment of the mismatch between waterbody segmentation metrics and shoreline positioning accuracy, several limitations should be acknowledged. First, the analysis is restricted to a set of established deep learning architectures trained for semantic segmentation and does not include more recent foundation models or vision–language models that leverage broader contextual information, multi-scale attention, or multi-modal inputs [53–57], nor models specifically designed to enhance image contrast in shadowed or low-light regions [58], such as vegetated or cliff-dominated land–water interfaces. These approaches may enhance boundary delineation and exhibit different relationships between segmentation performance and shoreline positioning accuracy, particularly in visually complex coastal environments.

Second, all experiments are based on NAIP RGB aerial imagery from the Great Lakes region. While NAIP provides high and relatively uniform spatial resolution, its spectral bands, acquisition conditions, and regional focus limit the generalizability of the results. As large glacially formed freshwater lakes with deep basins and energetic wave climates, the Great Lakes share several coastal processes, geomorphic characteristics, and management protocols with oceanic and marine coasts [59–61]. However, some important coastal environments—such as ice-covered shores, mangrove forests, tidal flats, deltas, coral islands, and macrotidal estuaries—are either uncommon in the Great Lakes or not captured during NAIP acquisition periods, and are therefore not represented in the dataset [62–64]. Likewise, other sensor types, like multispectral or hyperspectral ones mounted on satellites or UAVs , are also widely used in waterbody segmentation, shoreline mapping, riparian and coastal management, but not included in our analysis [65,66]. They may exhibit different relationships between segmentation performance and shoreline positioning, due to changes in contrast, water clarity, or vegetation structure.

Third, although NAIP imagery revisits the same locations at approximately two-year intervals, the present analysis considers only a single observation per site and does not explicitly address temporal variability. Consequently, the temporal stability of the relationships between segmentation metrics and shoreline positioning—such as their sensitivity to water-level fluctuations, seasonal vegetation dynamics, or ice cover—was not directly evaluated. This limitation is less critical in the Great Lakes region, where relevant processes evolve on seasonal to interannual timescales and NAIP imagery is collected during similar periods of the year. In contrast, in tidally dominated coastal environments, shoreline position and apparent geomorphology can change substantially over daily tidal cycles [67,68], meaning that acquisition timing (e.g., near high tide versus low tide) may strongly influence shoreline delineation and segmentation performance even at the same site.

Despite these constraints, the present work provides a systematic, large-sample demonstration that high waterbody segmentation performance does not imply accurate shoreline delineation, and that geometric descriptors such as SWR and SCI are essential for understanding when and why these discrepancies arise. As such, it offers a starting point and a transferable framework for future studies extending to additional models, sensors, and coastal environments.

4.4. Practical Implications for Model Evaluation and Coastal Applications

The findings have several practical implications for both the evaluation of deep learning models and their use in coastal management. First, they show that pixel accuracy and IoU, while useful for summarizing overall waterbody segmentation performance, are unreliable as universal surrogates for shoreline positioning. IoU generally provides a closer approximation to SIR than pixel accuracy, especially in high-SWR scenes where shoreline pixels represent a large fraction of the waterbody and boundary misalignment is more strongly penalized. However, even IoU exhibits landscape- and morphology-dependent behavior, failing to predict shoreline accuracy in rocky coasts with low SWR values. Practitioners should therefore avoid inferring shoreline quality solely from high segmentation scores, particularly in these challenging settings.

Second, the strong and systematic role of SWR—and the more context-dependent role of SCI—suggests that geometric descriptors can be explicitly integrated into model development and evaluation. In training, SWR and SCI could be used to design stratified sampling schemes, adjust loss weights, or define learning strategies that emphasize small, elongated, or highly curved waterbodies where shoreline sensitivity is greatest. Furthermore, shoreline-aware metrics such as SIR and AED could be incorporated into training objectives through boundary-aware or distance-based loss functions, and their effectiveness could be compared against conventional IoU- or pixel-accuracy–based losses to assess gains in shoreline positioning performance. In evaluation, reporting segmentation metrics separately for different SWR and SCI strata can reveal failure modes that would be obscured by aggregate scores, and help identify models that retain shoreline accuracy across diverse morphologies rather than only in broad, simple waterbodies.

Third, the results argue for incorporating shoreline-specific metrics such as SIR and AED into standard benchmarking protocols whenever models are used for applications that depend on accurate boundary placement, including erosion and accretion monitoring, hazard and floodplain mapping, navigation channel maintenance, habitat and wetland delineation, and infrastructure risk assessment. Evaluating models jointly with segmentation and shoreline metrics provides a more realistic basis for selecting architectures for operational deployments, and helps ensure that improvements in IoU or pixel accuracy translate into meaningful gains for management-relevant shoreline products.

Finally, by highlighting where and why segmentation metrics fail as shoreline surrogates, this study points toward more comprehensive, shoreline-aware evaluation frameworks and model designs. Future systems could, for example, couple conventional segmentation heads with explicit boundary-refinement modules, or jointly optimize for area-based and boundary-based losses, guided by SWR and SCI. Such approaches would move coastal mapping tools closer to the needs of real-world decision making, particularly in coastal environments where SWR-dependent errors can bias erosion risk assessment, infrastructure exposure, and habitat monitoring, and where the precise location of the shoreline is often more critical than the overall classification accuracy of land and water.

5. Conclusions

This study shows that accurate waterbody segmentation does not necessarily translate into accurate shoreline positioning, and that this gap has important implications for how deep learning models are evaluated and applied in coastal environments. Using diverse shoreline settings along the Great Lakes, we find that across a wide range of coastal landscapes, all tested models achieved high pixel accuracy and IoU, yet shoreline alignment often remained poor, particularly in wetlands and urban coasts. These results demonstrate that commonly reported segmentation metrics can substantially overstate the reliability of shoreline products used in coastal monitoring, risk assessment, and management. Our analysis further highlights the central role of shoreline–water geometry: the shoreline-to-water ratio (SWR) consistently governs when segmentation outputs meaningfully reflect shoreline position, whereas shoreline complexity (SCI) and water appearance exert weaker and highly context-dependent effects. Importantly, these patterns are largely consistent across different model architectures, indicating that the observed limitations are not specific to individual networks but arise from the geometric structure of coastal scenes themselves. From a practical perspective, this finding cautions against using pixel accuracy or IoU as stand-alone indicators of shoreline quality, especially in environments where narrow channels, fragmented waterbodies, or complex land–water transitions dominate. Together, these findings underscore the need for shoreline-aware evaluation strategies when shoreline location is a primary product. Metrics such as SIR and AED should be incorporated into routine model assessment, and future shoreline mapping systems should explicitly account for scene geometry during both training and validation, for example by stratifying performance by SWR or integrating boundary-aware objectives. In addition, extending this framework to oceanic and marine coasts using time-series imagery that accounts for tidal and water-level variability represents a natural next step, and will help clarify how temporal dynamics further influence the relationship between segmentation performance and shoreline positioning. By clarifying when and why segmentation metrics fail as shoreline surrogates, this work provides guidance for more reliable model selection, benchmarking, and deployment in real-world coastal applications, where the precise position of the shoreline is often more consequential than overall land–water classification accuracy.

Author Contributions: W.W., B.L. and F.J.: Data curation, Formal analysis, Validation, Methodology, Investigation, Conceptualization, Software, Visualization, Writing—original draft, Writing—review & editing. Y.L.: Conceptualization, Methodology, Writing—review & editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: CAID dataset used in this research is publicly available at https://zenodo.org/records/16461280 (accessed on 13 February 2026).

Acknowledgments: We acknowledge the United States Department of Agriculture (USDA) for providing NAIP aerial imagery and the U.S. Geological Survey (USGS) for offering the Earth Explorer API used for automated data acquisition.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

NAIP	National Agriculture Imagery Program
CAID	Coastal Aerial Imagery Dataset
IoU	Intersection over Union
MWH	Mean Water Hue

SCI	Shoreline Complexity Index
SWR	Shoreline-to-Water Ratio
SIR	Shoreline Intersection Ratio
AED	Average Eulerian Distance

References

1. Liu, H.; Jezek, K.C. Automated Extraction of Coastline from Satellite Imagery by Integrating Canny Edge Detection and Locally Adaptive Thresholding Methods. *Int. J. Remote Sens.* **2004**, *25*, 937–958. [CrossRef]
2. Cicin-Sain, B. Sustainable Development and Integrated Coastal Management. *Ocean Coast. Manag.* **1993**, *21*, 11–43. [CrossRef]
3. Boak, E.H.; Turner, I.L. Shoreline Definition and Detection: A Review. *J. Coast. Res.* **2005**, *21*, 688–703. [CrossRef]
4. Malone, T.; Davidson, M.; DiGiacomo, P.; Gonçalves, E.; Knap, T.; Muelbert, J.; Parslow, J.; Sweijd, N.; Yanagai, T.; Yap, H. Climate Change, Sustainable Development and Coastal Ocean Information Needs. *Procedia Environ. Sci.* **2010**, *1*, 324–341. [CrossRef]
5. Dean, R.G.; Houston, J.R. Determining Shoreline Response to Sea Level Rise. *Coast. Eng.* **2016**, *114*, 1–8. [CrossRef]
6. Griggs, G.; Reguero, B.G. Coastal Adaptation to Climate Change and Sea-Level Rise. *Water* **2021**, *13*, 2151. [CrossRef]
7. Pang, T.; Wang, X.; Nawaz, R.A.; Keefe, G.; Adekanmbi, T. Coastal Erosion and Climate Change: A Review on Coastal-Change Process and Modeling. *Ambio* **2023**, *52*, 2034–2052. [CrossRef]
8. Wang, W.; Lu, B.; Xu, Y.; Tekle, S.L.; Toni, A.T.; Zhang, X.; Li, Y.; Ajuwon, O.S.; Dahoro, D.A.; Martínez, E.; et al. Bibliometric Exploration of Infrastructure and Natural Hazards Research in Low and Middle Income Countries toward Sustainable Development Goals. *Discov. Sustain.* **2025**, *6*, 1407. [CrossRef]
9. Fruergaard, M.; Andersen, T.J.; Johannessen, P.N.; Nielsen, L.H.; Pejrup, M. Major Coastal Impact Induced by a 1000-Year Storm Event. *Sci. Rep.* **2013**, *3*, 1051. [CrossRef]
10. Lu, B.; Wang, W.; Wu, C.; Liu, Y. Wave Climate on the Southwestern Coast of Lake Michigan: Perspectives from Wave Directionality. *Ocean Eng.* **2026**, *343*, 123306. [CrossRef]
11. Li, X.; Damen, M.C.J. Coastline Change Detection with Satellite Remote Sensing for Environmental Management of the Pearl River Estuary, China. *J. Mar. Syst.* **2010**, *82*, S54–S61. [CrossRef]
12. Lu, B.; Wang, W.; Jordan, N.; Wright, D.; Bechle, A.; Zoet, L.; Wu, C. A Multi-Scale Assessment for Managing Coastal Geomorphic Changes in Southwestern Lake Michigan. *J. Environ. Manag.* **2025**, *395*, 127878. [CrossRef] [PubMed]
13. Ranasinghe, R.; Turner, I.L. Shoreline Response to Submerged Structures: A Review. *Coast. Eng.* **2006**, *53*, 65–79. [CrossRef]
14. Kuleli, T.; Guneroglu, A.; Karsli, F.; Dihkan, M. Automatic Detection of Shoreline Change on Coastal Ramsar Wetlands of Turkey. *Ocean Eng.* **2011**, *38*, 1141–1149. [CrossRef]
15. Ribas, F.; Simarro, G.; Arriaga, J.; Luque, P. Automatic Shoreline Detection from Video Images by Combining Information from Different Methods. *Remote Sens.* **2020**, *12*, 3717. [CrossRef]
16. Tsiakos, C.-A.D.; Chalkias, C. Use of Machine Learning and Remote Sensing Techniques for Shoreline Monitoring: A Review of Recent Literature. *Appl. Sci.* **2023**, *13*, 3268. [CrossRef]
17. Liu, Y.; Yang, Y.; Li, X.; Yang, F.; Xie, H.; Wang, W.; Dong, C. A Deep Learning-Based Pipeline for Detecting Rip Currents from Satellite Imagery. *Remote Sens.* **2026**, *18*, 368. [CrossRef]
18. Christofi, D.; Mettas, C.; Evagorou, E.; Stylianou, N.; Eliades, M.; Theocharidis, C.; Chatzipavlis, A.; Hasiotis, T.; Hadjimitsis, D. A Review of Open Remote Sensing Data with GIS, AI, and UAV Support for Shoreline Detection and Coastal Erosion Monitoring. *Appl. Sci.* **2025**, *15*, 4771. [CrossRef]
19. Yang, X.; Zhao, S.; Qin, X.; Zhao, N.; Liang, L. Mapping of Urban Surface Water Bodies from Sentinel-2 MSI Imagery at 10 m Resolution via NDWI-Based Image Sharpening. *Remote Sens.* **2017**, *9*, 596. [CrossRef]
20. Abdelhady, H.U.; Troy, C.D.; Habib, A.; Manish, R. A Simple, Fully Automated Shoreline Detection Algorithm for High-Resolution Multi-Spectral Imagery. *Remote Sens.* **2022**, *14*, 557. [CrossRef]
21. Chang, L.; Cheng, L.; Huang, C.; Qin, S.; Fu, C.; Li, S. Extracting Urban Water Bodies from Landsat Imagery Based on mNDWI and HSV Transformation. *Remote Sens.* **2022**, *14*, 5785. [CrossRef]
22. Tang, W.; Zhao, C.; Lin, J.; Jiao, C.; Zheng, G.; Zhu, J.; Pan, X.; Han, X. Improved Spectral Water Index Combined with Otsu Algorithm to Extract Muddy Coastline Data. *Water* **2022**, *14*, 855. [CrossRef]
23. Li, K.; Wang, J.; Yao, J. Effectiveness of Machine Learning Methods for Water Segmentation with ROI as the Label: A Case Study of the Tuul River in Mongolia. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *103*, 102497. [CrossRef]
24. Pawana, I.G.N.A.; Widyantara, I.M.O.; Sudarma, M.; Linawati; Wirastuti, N. Image Enhancement Using CLAHE and Noise Removal for Shoreline Detection Framework. In *Proceedings of the 2023 3rd International Conference on Smart Cities, Automation & Intelligent Computing Systems (ICON-SONICS), Bali, Indonesia, 6–8 December 2023*; IEEE: New York, NY, USA, 2023; pp. 183–188.
25. Attya, M.; Abo-Seida, O.M.; Abdulkader, H.M.; Mohammed, A.M. A Hybrid Deep Learning Approach for Accurate Water Body Segmentation in Satellite Imagery. *Earth Sci. Inform.* **2025**, *18*, 418. [CrossRef]

26. Minaee, S.; Boykov, Y.Y.; Porikli, F.; Plaza, A.J.; Kehtarnavaz, N.; Terzopoulos, D. Image Segmentation Using Deep Learning: A Survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *44*, 3523–3542. [CrossRef]
27. Miao, Z.; Fu, K.; Sun, H.; Sun, X.; Yan, M. Automatic Water-Body Segmentation From High-Resolution Satellite Images via Deep Networks. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 602–606. [CrossRef]
28. Gautam, S.; Singhai, J. Critical Review on Deep Learning Methodologies Employed for Water-Body Segmentation through Remote Sensing Images. *Multimed. Tools Appl.* **2024**, *83*, 1869–1889. [CrossRef]
29. Shelhamer, E.; Long, J.; Darrell, T. Fully Convolutional Networks for Semantic Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 640–651. [CrossRef]
30. Zhang, H.; Dana, K.; Shi, J.; Zhang, Z.; Wang, X.; Tyagi, A.; Agrawal, A. Context Encoding for Semantic Segmentation. In *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018*; IEEE: New York, NY, USA, 2018; pp. 7151–7160.
31. Huang, Z.; Wang, X.; Huang, L.; Huang, C.; Wei, Y.; Liu, W. CCNet: Criss-Cross Attention for Semantic Segmentation. In Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Republic of Korea, 27 October–2 November 2019; pp. 603–612.
32. Guo, Z.; Wu, L.; Huang, Y.; Guo, Z.; Zhao, J.; Li, N. Water-Body Segmentation for SAR Images: Past, Current, and Future. *Remote Sens.* **2022**, *14*, 1752. [CrossRef]
33. Shi, T.; Guo, Z.; Li, C.; Lan, X.; Gao, X.; Yan, X. Improvement of Deep Learning Method for Water Body Segmentation of Remote Sensing Images Based on Attention Modules. *Earth Sci. Inform.* **2023**, *16*, 2865–2876. [CrossRef]
34. Sun, D.; Gao, G.; Huang, L.; Liu, Y.; Liu, D. Extraction of Water Bodies from High-Resolution Remote Sensing Imagery Based on a Deep Semantic Segmentation Network. *Sci. Rep.* **2024**, *14*, 14604. [CrossRef]
35. Erfani, S.M.H.; Wu, Z.; Wu, X.; Wang, S.; Goharian, E. ATLANTIS: A Benchmark for Semantic Segmentation of Waterbody Images. *Environ. Model. Softw.* **2022**, *149*, 105333. [CrossRef]
36. Wang, W.; Lu, B.; Li, Y.; Shi, W. Descriptor: Coastal Aerial Imagery Dataset for Shoreline Segmentation (CAID). *IEEE Data Descr.* **2025**, *2*, 286–295. [CrossRef]
37. Bernhard, M.; Amoroso, R.; Kindermann, Y.; Baraldi, L.; Cucchiara, R.; Tresp, V.; Schubert, M. What's Outside the Intersection? Fine-Grained Error Analysis for Semantic Segmentation Beyond IoU. In *Proceedings of the 2024 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), Waikoloa, HI, USA, 3–8 January 2024*; IEEE: New York, NY, USA, 2024; pp. 957–966.
38. Khurram, S.; Pour, A.B.; Bagheri, M.; Ariffin, E.H.; Akhir, M.F.; Hamzah, S.B. Developments in Deep Learning Algorithms for Coastline Extraction from Remote Sensing Imagery: A Systematic Review. *Earth Sci. Inform.* **2025**, *18*, 292. [CrossRef]
39. Bengoufa, S.; Niculescu, S.; Mihoubi, M.K.; Belkessa, R.; Abbad, K. Rocky Shoreline Extraction Using a Deep Learning Model and Object-based Image Analysis. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* **2021**, *XLIII-B3-2021*, 23–29. [CrossRef]
40. Zhu, Z.; Xu, M.; Bai, S.; Huang, T.; Bai, X. Asymmetric Non-Local Neural Networks for Semantic Segmentation. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Republic of Korea, 27 October–2 November 2019; pp. 593–602.
41. Zhao, H.; Shi, J.; Qi, X.; Wang, X.; Jia, J. Pyramid Scene Parsing Network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 2881–2890.
42. McFeeters, S.K. The Use of the Normalized Difference Water Index (NDWI) in the Delineation of Open Water Features. *Int. J. Remote Sens.* **1996**, *17*, 1425–1432. [CrossRef]
43. Xu, H. Modification of Normalised Difference Water Index (NDWI) to Enhance Open Water Features in Remotely Sensed Imagery. *Int. J. Remote Sens.* **2006**, *27*, 3025–3033. [CrossRef]
44. Feyisa, G.L.; Meilby, H.; Fensholt, R.; Proud, S.R. Automated Water Extraction Index: A New Technique for Surface Water Mapping Using Landsat Imagery. *Remote Sens. Environ.* **2014**, *140*, 23–35. [CrossRef]
45. Kumar, L.; Afzal, M.S.; Afzal, M.M. Mapping Shoreline Change Using Machine Learning: A Case Study from the Eastern Indian Coast. *Acta Geophys.* **2020**, *68*, 1127–1143. [CrossRef]
46. Sheykhmousa, M.; Mahdianpari, M.; Ghanbari, H.; Mohammadimanesh, F.; Ghamisi, P.; Homayouni, S. Support Vector Machine Versus Random Forest for Remote Sensing Image Classification: A Meta-Analysis and Systematic Review. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 6308–6325. [CrossRef]
47. Erdem, F.; Bayram, B.; Bakirman, T.; Bayrak, O.C.; Akpinar, B. An Ensemble Deep Learning Based Shoreline Segmentation Approach (WaterNet) from Landsat 8 OLI Images. *Adv. Space Res.* **2021**, *67*, 964–974. [CrossRef]
48. Dang, K.B.; Dang, V.B.; Ngo, V.L.; Vu, K.C.; Nguyen, H.; Nguyen, D.A.; Nguyen, T.D.L.; Pham, T.P.N.; Giang, T.L.; Nguyen, H.D.; et al. Application of Deep Learning Models to Detect Coastlines and Shorelines. *J. Environ. Manag.* **2022**, *320*, 115732. [CrossRef] [PubMed]
49. Choung, Y.-J.; Jo, M.-H. Comparison between a Machine-Learning-Based Method and a Water-Index-Based Method for Shoreline Mapping Using a High-Resolution Satellite Image Acquired in Hwado Island, South Korea. *J. Sens.* **2017**, *2017*, 8245204. [CrossRef]

50. Seale, C.; Redfern, T.; Chatfield, P.; Luo, C.; Dempsey, K. Coastline Detection in Satellite Imagery: A Deep Learning Approach on New Benchmark Data. *Remote Sens. Environ.* **2022**, *278*, 113044. [CrossRef]
51. Zou, N.; Xiang, Z.; Chen, Y.; Chen, S.; Qiao, C. Boundary-Aware CNN for Semantic Segmentation. *IEEE Access* **2019**, *7*, 114520–114528. [CrossRef]
52. Mahmoud, A.S.; Mohamed, S.A.; Helmy, A.K.; Nasr, A.H. BDCN_UNet: Advanced Shoreline Extraction Techniques Integrating Deep Learning. *Earth Sci. Inform.* **2025**, *18*, 187. [CrossRef]
53. Zhang, Z.; Chen, J.; Shi, W.; Yi, L.; Wang, C.; Yu, Q. Contrastive Learning for Knowledge-Based Question Generation in Large Language Models. In *Proceedings of the 2024 5th International Conference on Intelligent Computing and Human-Computer Interaction (ICHCI), Nanchang, China, 27–29 September 2024*; IEEE: New York, NY, USA, 2024; pp. 583–587.
54. Ozdemir, S.; Akbulut, Z.; Karsli, F.; Kavzoglu, T. Extraction of Water Bodies from High-Resolution Aerial and Satellite Images Using Visual Foundation Models. *Sustainability* **2024**, *16*, 2995. [CrossRef]
55. Zhao, J.; Yabuki, N.; Fukuda, T. VAGen: Waterbody Segmentation with Prompting for Visual in-Context Learning. *AI Civ. Eng.* **2024**, *3*, 24. [CrossRef]
56. Zhou, T.; Huang, Z.; Lin, H.; Zhou, Z.; Hu, J. MACityChat: Integrating Remote Sensing Professional Large Model with General-Purpose Large Model for Multi-Domain Urban Land Use Analysis. *Appl. Soft Comput.* **2025**, *185*, 113929. [CrossRef]
57. Zhang, W.; Cai, M.; Zhang, T.; Zhuang, Y.; Li, J.; Mao, X. EarthMarker: A Visual Prompting Multimodal Large Language Model for Remote Sensing. *IEEE Trans. Geosci. Remote Sens.* **2025**, *63*, 5604219. [CrossRef]
58. Wang, J.; He, Y.; Li, K.; Li, S.; Zhao, L.; Yin, J.; Zhang, M.; Shi, T.; Wang, X. MDANet: A Multi-Stage Domain Adaptation Framework for Generalizable Low-Light Image Enhancement. *Neurocomputing* **2025**, *627*, 129572. [CrossRef]
59. Lawrence, P.L. Development of Great Lakes Shoreline Management Plans by Ontario Conservation Authorities. *Ocean Coast. Manag.* **1995**, *26*, 205–223. [CrossRef]
60. Rao, Y.R.; Schwab, D.J. Transport and Mixing Between the Coastal and Offshore Waters in the Great Lakes: A Review. *J. Great Lakes Res.* **2007**, *33*, 202–218. [CrossRef]
61. Gronewold, A.D.; Fortin, V.; Lofgren, B.; Clites, A.; Stow, C.A.; Quinn, F. Coasts, Water Levels, and Climate Change: A Great Lakes Perspective. *Clim. Change* **2013**, *120*, 697–711. [CrossRef]
62. Farquharson, L.M.; Mann, D.H.; Swanson, D.K.; Jones, B.M.; Buzard, R.M.; Jordan, J.W. Temporal and Spatial Variability in Coastline Response to Declining Sea-Ice in Northwest Alaska. *Mar. Geol.* **2018**, *404*, 71–83. [CrossRef]
63. Thakur, S.; Mondal, I.; Bar, S.; Nandi, S.; Ghosh, P.B.; Das, P.; De, T.K. Shoreline Changes and Its Impact on the Mangrove Ecosystems of Some Islands of Indian Sundarbans, North-East Coast of India. *J. Clean. Prod.* **2021**, *284*, 124764. [CrossRef]
64. Mao, Y.; Harris, D.L.; Xie, Z.; Phinn, S. Efficient Measurement of Large-Scale Decadal Shoreline Change with Increased Accuracy in Tide-Dominated Coastal Environments with Google Earth Engine. *ISPRS J. Photogramm. Remote Sens.* **2021**, *181*, 385–399. [CrossRef]
65. Shenbagaraj, N.; Senthil Kumar, K.; Rasheed, A.M.; Leostalin, J.; Kumar, M.N. Mapping and Electronic Publishing of Shoreline Changes Using UAV Remote Sensing and GIS. *J. Indian Soc. Remote Sens.* **2021**, *49*, 1769–1777. [CrossRef]
66. Wang, W.; Lu, B.; Wu, C.H. Cost-Effective Drone Monitoring and Evaluating Toolkits for Stream Habitat Health: Development and Application. *Environ. Monit Assess* **2025**, *198*, 10. [CrossRef]
67. Young, D.R.; Brantley, S.T.; Zinnert, J.C.; Vick, J.K. Landscape Position and Habitat Polygons in a Dynamic Coastal Environment. *Ecosphere* **2011**, *2*, art71. [CrossRef]
68. Cooper, J.A.G. Geomorphological Variability among Microtidal Estuaries from the Wave-Dominated South African Coast. *Geomorphology* **2001**, *40*, 99–122. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.