Evaluating the Impact of Super-Resolution for Coastal Boundary Segmentation Using Deep Learning for High-Resolution Imagery

Marc-André Blais[1], Moulay A. Akhloufi[1]

[1]Department of Computer Science, Université de Moncton
PRIME Lab (Perception, Robotics and Intelligent Machines)
Moncton, NB, Canada
{emb9357, moulay.akhloufi}@umoncton.ca

Keywords: Super-Resolution, Deep Learning, Segmentation, Coastal Boundary, Remote Sensing

Abstract

Coastal areas play an important role economically, socially and environmentally due to their many functions. However, these regions are at risk of erosion, which is further exacerbated by human-driven climate change. Tracking and monitoring coastal boundaries enable efficient allocation of conservation and protection efforts. Due to the vast size and complexity of coastal areas, on-site monitoring to track erosion is inefficient. Artificial intelligence has shown impressive results in segmenting and extracting these boundaries from remote sensing imagery. Historical remote sensing data make it possible to track long-term erosion but remain challenging due to the coarse resolution of older data. Our work proposes studying the impact of super-resolution on coastal boundary segmentation using high-resolution imagery. ESRGAN and SRCNN have proven highly beneficial in improving the quality of coarse-resolution samples, achieving superior performance compared to bicubic interpolation across scaling factors ranging from ×2 to ×12. ESRGAN super-resolved samples achieved F1-scores ranging from 97.75% to 89.92% for scaling factors ×2 to ×12, while bicubic interpolation achieved between 97.34% and 65.27%. These improvements demonstrate that SR enhances boundary delineation and robustness across scales. Our work also explores the applicability of tracking erosion through historical data. Results demonstrate a coastal boundary change of 0.23 m per year over seven years, which is on par with expected values.

1. Introduction

Coastal areas are crucial due to their economic and environmental importance (Seitz et al., 2014). In fact, as much as 80% of goods traded worldwide are sea-based (United Nations Conference on Trade and Development, 2025) and more than two billion people live within 50 kilometres of the coast (Cosby et al., 2024), making coastal cities vital to economies. Birds, fish and other animals use these regions as feeding grounds (McLachlan et al., 1980), mating areas and habitats (Silvano et al., 2006). However, these regions are at risk of destruction due to erosion. As the shoreline retreats, infrastructure is damaged and population centres are forced to adapt (Li et al., 2021, Phillips and Jones, 2006). Erosion, although a natural phenomenon, is exacerbated by human-driven climate change (Pang et al., 2023), pollution (Vikas and Dwarakish, 2015) and urban development (Pasquali and Marucci, 2021). Coastal erosion monitoring is thus crucial to prevent and slow down further damage to these areas. In-situ techniques, although accurate and viable for limited areas, are not suitable for large-scale monitoring.

Remote sensing has seen an increase in popularity for diverse applications in recent decades due to its accessibility (Li et al., 2024). It consists of scanning the Earth’s surface using satellites, aircraft or drones equipped with passive or active sensors. Passive sensors include optical, infrared and thermal instruments, while active sensors consist of systems such as Synthetic Aperture Radar (SAR) and Light Detection and Ranging (LiDAR). Passive sensors have shorter revisit times, higher resolutions and extensive historical archives compared to active sensors. For example, multispectral and hyperspectral passive sensors provide valuable historical datasets dating back to the 1970s (Smedes et al., 1970). These characteristics make passive sensor data better suited for long-term erosion monitoring.

Although manually tracking boundaries in remote sensing data is possible, it becomes inefficient for large-scale applications. Artificial intelligence, particularly Deep Learning (DL), has been proposed as a solution to automatically extract and track coastal boundaries (Khurram et al., 2025, Blais and Akhloufi, 2025). DL has proven highly proficient and robust to noise and data variation in segmenting coastal areas. It has also demonstrated strong generalization capabilities across datasets and regions. DL combined with historical remote sensing data offers an interesting approach to track erosion over decades. However, older remote sensing data often have much coarser resolutions, which limits the applicability of DL. Super-resolution (SR), an enhancement technique based on DL, has shown great potential for increasing the resolution of remote sensing data (Fernandez-Beltran et al., 2017). However, the impact and performance of SR for coastal boundary extraction in high-resolution imagery have not been widely studied. In this study, we aim to fill this gap by evaluating the impact of SR on coastal boundary segmentation. High-resolution data is used to train and test SR algorithms against bicubic interpolation at various scales for segmentation. The applicability of the algorithms is also tested on historical data to track erosion over a seven-year period. Four contributions are presented in this paper:

- Compare SR with bicubic interpolation for remote sensing data.
- Evaluate SR for coastal boundary segmentation on high-resolution imagery.
- Test the generalization of SR-enhanced segmentation on two new data sources.
- Demonstrate its use for temporal erosion monitoring.

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-XI-3-2026-35-2026

2. Related Works

Various studies have explored the use of remote sensing data for coastline extraction, such as Zhou et al. (Zhou et al., 2023), who presented an overview of this field. The authors discussed recent progress related to data sources and extraction methods. Various types of coastlines were identified and defined, namely rocky, sandy, silty, biological and artificial, with nearly 75% of coastlines consisting of the rocky type. The authors also explored coastline indicators, noting that most previous studies used the instantaneous waterline as an indicator. In total, seven extraction methods were identified: remote sensing indices, thresholding, edge detection, active contour models, polarization methods, knowledge graphs and machine-learning-based techniques. Remote sensing indices have shown good results and rely on extracting the coastal boundary based on the waterline. However, they require manual tuning when the region, terrain or lighting conditions vary. Similarly, other techniques perform well in specific conditions but require frequent adjustment of parameters. The authors noted that machine-learning-based techniques achieved overall superior performance. They also highlighted that DL-based techniques are currently limited by dataset size, resolution and variety.

Blais and Akhloufi (Blais and Akhloufi, 2025) presented an in-depth literature review, while Khurram et al. (Khurram et al., 2025) presented a systematic review. In the work of Blais and Akhloufi, the authors reviewed various data sources, datasets and DL-based solutions for coastal boundary extraction. They described multiple remote sensing indices and available platforms for data acquisition, such as optical and SAR data. In terms of datasets, seven publicly available datasets were presented, varying in resolution, size and geographic region. The majority of publicly available datasets consist of optical imagery and have coarse resolutions, ranging from 10 m per pixel to more than 60 m. The authors focused on segmentation and direct extraction solutions, as these were the most widely used. Various DL-based algorithms were explored, ranging from simple encoder-decoder architectures to more complex transformer-based techniques. Overall, most DL-based methods achieved strong segmentation results, with transformers offering slightly better performance than traditional encoder-decoders. The authors also identified several limitations, namely the lack of high-resolution datasets, real-time monitoring, large and diverse data sources, historical data and temporal information. Khurram et al. (Khurram et al., 2025) applied the PRISMA paradigm to identify relevant studies for their systematic review, resulting in 24 selected works. They identified UNet as the most relevant algorithm for this task, mainly due to its wide adoption and strong performance in segmentation tasks. They also found that ensemble-learning methods can improve the overall performance of DL-based solutions. The authors further identified the lack of high-resolution data, coastal complexity and real-time monitoring as major limitations in this field.

Various works have been proposed in recent years for coastal boundary extraction. We highlight the work of Blais and Akhloufi (Blais and Akhloufi, 2021) due to its use of high-resolution imagery. In their study, the authors used orthoimagery with a submetre resolution per pixel. The images, collected in New Brunswick, Canada, were manually labelled by experts. The authors compared various DL models and encoders, with FPN (Lin et al., 2017) combined with VGG16 (Simonyan and Zisserman, 2014) achieving slightly better results. Their work focused on traditional encoder-decoder architectures and did not include transformer-based algorithms. As most historical remote sensing data have much coarser resolutions, often above 10 m per pixel, erosion monitoring remains challenging. Various methods have been proposed to enhance the resolution of historical data, such as SR techniques based on DL. Lei et al. (Lei et al., 2025) proposed a multi-scale residual super-resolution generative adversarial network for this purpose. The authors used Sentinel-2 and GF-2 satellite data from nine typical land cover types found in coastal zones. Their approach consisted of three key modules: fusion attention-enhanced residual modules, multi-scale attention fusion, and multi-scale feature extraction. The goal of these modules was to mitigate gradient vanishing and extract features at multiple scales to improve overall performance. At a ×4 scale, their approach improved the Learned Perceptual Image Patch Similarity (LPIPS) with a reduction of 14.34% while improving the Structural Similarity Index Measure (SSIM) by 11.85%. Visually, their approach produced strong results, although some methods achieved slightly sharper outputs.

In similar work to this study, Pala and Alganci (Pala and Algancı, 2025) investigated the performance of coastline segmentation using super-resolved data. The authors used Landsat-8 and Sentinel-2 imagery, with respective native resolutions of 30 m and 10 m. These images were downsampled using factors of 2 and 4, resulting in resolutions of 120 m and 60 m for Landsat-8 and 20 m and 40 m for Sentinel-2. The authors used these newly formed datasets to train SRGAN (Ledig et al., 2017). SRGAN uses two models, a generator based on a deep ResNet (He et al., 2016) and a discriminator that follows a structure similar to VGG19 (Simonyan and Zisserman, 2014). As remote sensing data often contain 32-bit values, SRGAN was modified from 8-bit to handle 32-bit data. Three segmentation models were compared: UNet (Ronneberger et al., 2015), LinkNet (Chaurasia and Culurciello, 2017) and DeepLabV3+ (Chen et al., 2017), each trained separately on the four datasets. Additionally, GT-1 images, with resolutions between 0.5 m and 2 m, were used as references for segmentation evaluation. Overall, the authors observed improved boundary delineation with SR images compared to the downsampled datasets. For the ×2 scale, the SSIM improved from 0.93 to 0.99 and the Peak Signal-to-Noise Ratio (PSNR) increased from 31 to 41, while the ×4 scale showed a lower increase in performance. DeepLabV3+ achieved the highest segmentation accuracy for low-resolution data, LinkNet performed best on SR samples and LinkNet and UNet achieved the best performance for GT-1. Overall, the authors demonstrated that super-resolution at ×2 and ×4 scales enabled slightly better segmentation performance. However, the use of coarse-resolution data, namely 10 m and 60 m, limits the applicability of this study for erosion monitoring, which often requires high-resolution imagery.

3. Proposed Approach

We propose a two-phase approach consisting of super-resolution and segmentation phases. The super-resolution phase analyzes and compares the impact of SR on remote sensing images at various scales. The segmentation phase, on the other hand, evaluates the influence of SR on coastal boundary segmentation. For both tasks, the dataset from Blais and Akhloufi (Blais and Akhloufi, 2021) was utilized due to its high resolution. Originally, the dataset consisted of 183, 7 and 4 large training, validation and testing samples with matching ground truth (GT). Each file was processed by cropping samples of size 480×480, each containing at least 1,000 pixels of each class. The final

generated dataset consisted of 4 483, 175 and 100 training, validation and test samples, respectively. For SR, the images were first downsampled using pixel-area relation resampling. Various scales were evaluated (×2, ×3, ×4, ×5, ×6, ×8, ×10, ×12), resulting in low-resolution (LR) images ranging in size from 240×240 pixels to 40×40 pixels. The matching LR and high-resolution (HR) samples were used to train two DL-based SR algorithms, namely the Super-Resolution Convolutional Neural Network (SRCNN) (Dong et al., 2015) and the Enhanced Super-Resolution GAN (ESRGAN) (Wang et al., 2018). SRCNN utilizes a simple CNN to learn and transform LR samples into their HR counterparts. ESRGAN, on the other hand, employs a more complex architecture consisting of a discriminator and a generator. The generator, based on RRDBNet, is responsible for upsampling the images, while the discriminator predicts whether a real sample looks more realistic than a generated one. SRCNN uses a loss function based on the Mean Squared Error (MSE), whereas ESRGAN combines adversarial, perceptual and content-based losses. These algorithms were chosen due to their lightweight and easily modifiable structures. Both algorithms were trained for 50 epochs, with the Peak Signal-to-Noise Ratio (PSNR), derived from the MSE, used to compare their performance.

For segmentation, the cropped dataset at full resolution was used to train a SegFormer with the MiT-B5 backbone (Xie et al., 2021). SegFormer was chosen due to its ability to handle higher resolutions and its superior performance in benchmarking studies (Blais and Akhloufi, 2025). It consists of a transformer-based encoder and a lightweight MLP decoder, which enables a more efficient architecture compared to CNN-based algorithms. The segmentation model was trained for 20 epochs using the F1-Score, edge F1-Score (eF1) and Edge Pixel Proximity Accuracy (ePPA) as evaluation metrics. The eF1 metric applies the F1-Score to extracted edges obtained using the Canny operator on both the prediction and GT. Similarly, ePPA measures the proportion of predicted edge pixels that lie within N pixels of the GT boundary. The epoch with the highest eF1 was selected for evaluation to ensure strong boundary delineation from the model. Once both phases were trained, the impact of SR on segmentation was evaluated by comparing the results of SRCNN and ESRGAN against bicubic interpolation using 100 unseen test samples. These results were compared against the original resolution using both quantitative metrics and visual assessments. To further assess the generalization of this approach, temporal scenes from two new data sources were also analyzed.

4. Results

4.1 Super-Resolution

Using 100 test samples from the Blais dataset, the PSNR, SSIM, Natural Image Quality Evaluator (NIQE) and LPIPS were evaluated for each scaling factor. PSNR measures distortion differences between samples, while SSIM primarily focuses on luminance, contrast and structure. LPIPS computes the difference between the features of the HR and super-resolved LR samples using AlexNet. NIQE uses statistical information to analyze super-resolved LR samples without their HR counterparts. Table 1 presents the various results for bicubic interpolation, SRCNN and ESRGAN, showing ESRGAN outperforming both methods at all scales. Figure 1 presents an example at each scale using no upscaling, bicubic interpolation and ESRGAN.

Table 1. PSNR, SSIM, NIQE and LPIPS on Test Set.

Scale	Method	PSNR ↑	SSIM ↑	NIQE ↓	LPIPS ↓
×2	Bicubic	30.56	0.7803	**5.311**	0.3633
	SRCNN	31.07	0.8088	5.767	0.2296
	ESRGAN	**31.59**	**0.8237**	5.633	**0.1846**
×3	Bicubic	28.91	0.6763	7.364	0.5562
	SRCNN	29.24	0.7057	6.733	0.4507
	ESRGAN	**29.66**	**0.7247**	**6.066**	**0.3955**
×4	Bicubic	28.08	0.6149	7.681	0.6456
	SRCNN	28.33	0.6442	7.472	0.5545
	ESRGAN	**28.66**	**0.6611**	**6.851**	**0.4799**
×5	Bicubic	27.54	0.5765	9.461	0.6986
	SRCNN	27.79	0.6004	7.911	0.5994
	ESRGAN	**28.03**	**0.6172**	**7.562**	**0.5345**
×6	Bicubic	27.16	0.5514	10.501	0.7333
	SRCNN	27.36	0.5718	8.248	0.6466
	ESRGAN	**27.57**	**0.5870**	**7.775**	**0.5874**
×8	Bicubic	26.61	0.5225	11.673	0.7745
	SRCNN	26.78	0.5368	**8.753**	0.6946
	ESRGAN	**26.96**	**0.5482**	8.893	**0.6522**
×10	Bicubic	26.23	0.5076	12.451	0.8255
	SRCNN	26.38	0.5176	**9.363**	0.7494
	ESRGAN	**26.57**	**0.5283**	9.582	**0.6996**
×12	Bicubic	25.93	0.4988	13.494	0.8556
	SRCNN	26.07	0.5057	**9.993**	0.7724
	ESRGAN	**26.27**	**0.5139**	10.224	**0.7342**

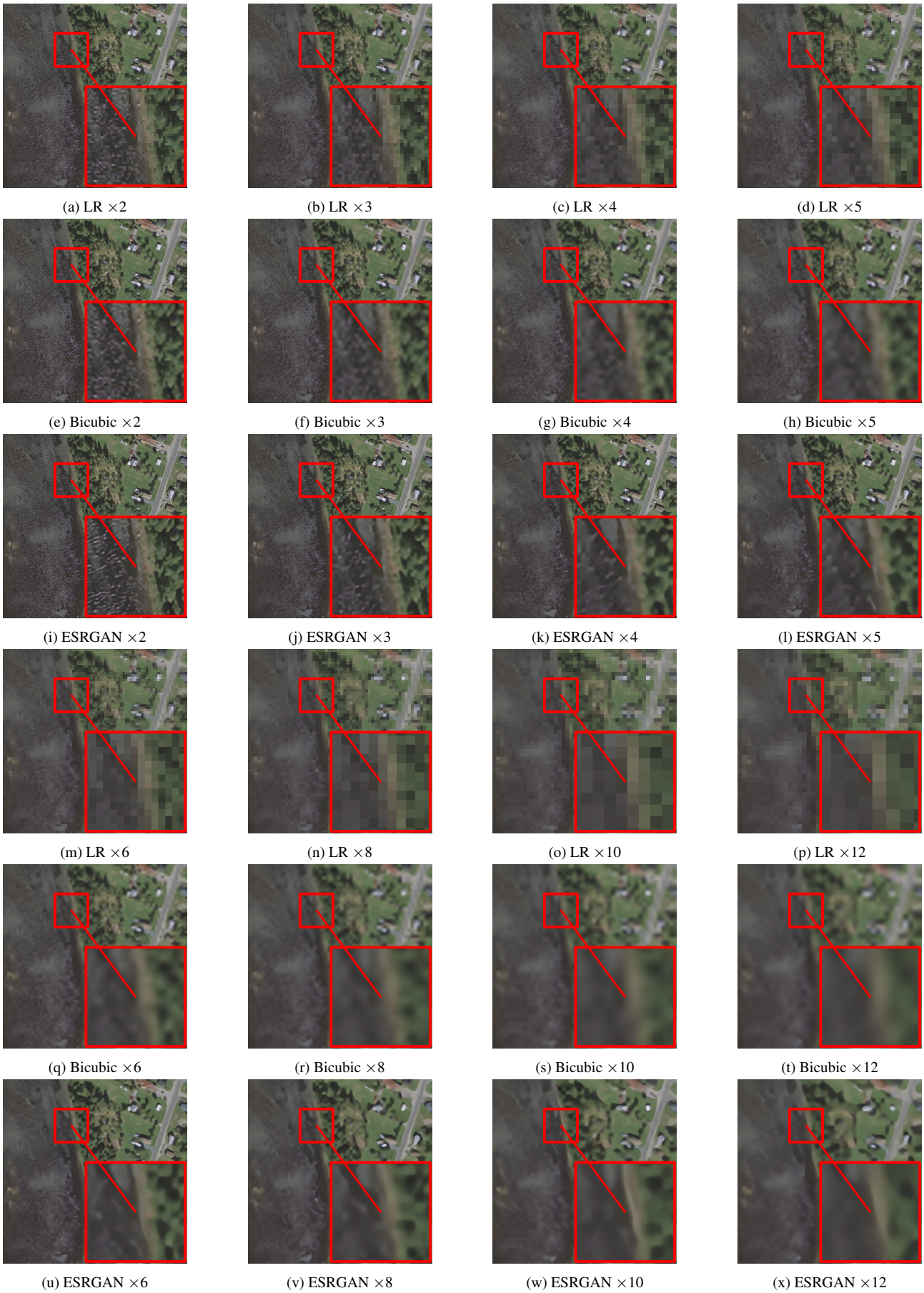
4.2 SR+Segmentation

To analyze the impact of SR on segmentation, the first step involved increasing the resolution of the downsampled test samples. A total of 100 test samples from the dataset were downsampled, upsampled and subsequently segmented. Table 2 presents the Intersection over Union (IoU), eF1, ePPA1 and F1-Score (F1) for the various segmentation results on the test set. This table shows that SRCNN and ESRGAN outperformed bicubic interpolation across all scales. Figure 2 presents three scenes of extracted edges at scales ×2, ×6 and ×12 using the GT, bicubic interpolation, SRCNN and ESRGAN.

Table 2. Segmentation Metrics on Test Set (n=100).

Scale	Method	IoU (%)	F1 (%)	eF1 (%)	ePPA@1 (%)
HR	HR	95.37	97.54	6.38	19.05
×2	Bicubic	95.03	97.34	5.66	16.16
	SRCNN	95.69	97.71	7.53	21.59
	ESRGAN	**95.78**	**97.75**	**8.75**	**23.92**
×3	Bicubic	94.63	97.11	4.43	13.22
	SRCNN	**94.92**	**97.28**	5.26	15.06
	ESRGAN	94.27	96.90	**7.14**	**19.77**
×4	Bicubic	94.48	97.02	4.23	12.96
	SRCNN	**94.75**	**97.18**	3.70	11.39
	ESRGAN	93.10	96.26	**5.79**	**16.10**
×5	Bicubic	94.23	96.88	4.69	13.77
	SRCNN	**95.12**	**97.39**	4.19	13.59
	ESRGAN	94.32	96.92	**6.25**	**17.76**
×6	Bicubic	92.46	95.77	5.27	14.63
	SRCNN	94.73	97.17	4.58	13.85
	ESRGAN	**95.29**	**97.49**	**6.19**	**18.52**
×8	Bicubic	84.31	88.63	5.01	13.45
	SRCNN	93.42	**96.31**	6.13	16.12
	ESRGAN	**93.53**	96.26	**6.27**	**18.51**
×10	Bicubic	75.47	80.57	4.88	13.69
	SRCNN	90.68	94.03	**7.85**	**20.43**
	ESRGAN	**91.46**	**94.51**	7.20	19.45
×12	Bicubic	59.11	65.27	2.91	7.94
	SRCNN	84.95	89.26	**6.48**	**16.81**
	ESRGAN	**85.55**	**89.92**	5.96	16.22

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-XI-3-2026-35-2026



(a) LR ×2 (b) LR ×3 (c) LR ×4 (d) LR ×5

(e) Bicubic ×2 (f) Bicubic ×3 (g) Bicubic ×4 (h) Bicubic ×5

(i) ESRGAN ×2 (j) ESRGAN ×3 (k) ESRGAN ×4 (l) ESRGAN ×5

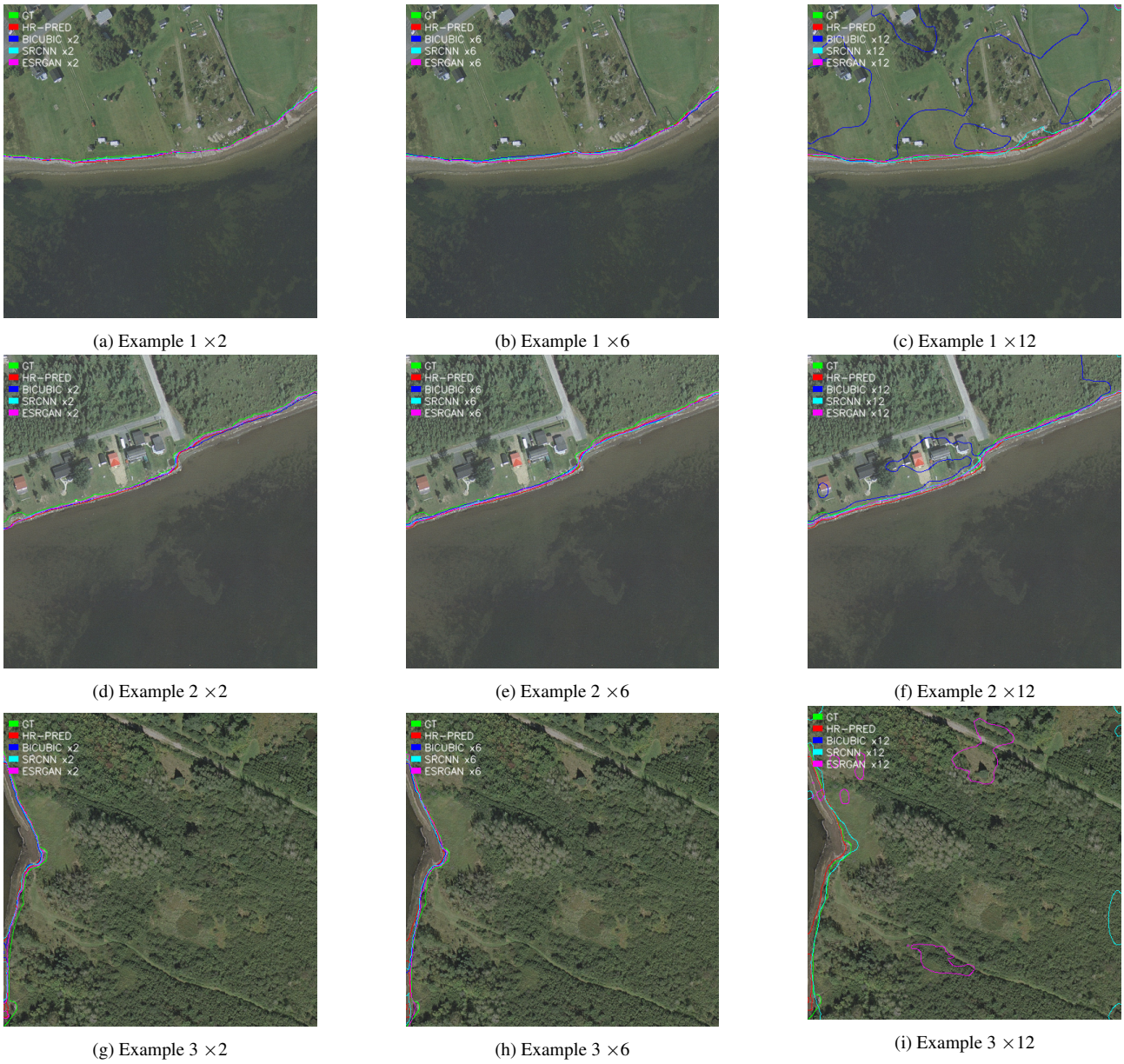
(m) LR ×6 (n) LR ×8 (o) LR ×10 (p) LR ×12

(q) Bicubic ×6 (r) Bicubic ×8 (s) Bicubic ×10 (t) Bicubic ×12

(u) ESRGAN ×6 (v) ESRGAN ×8 (w) ESRGAN ×10 (x) ESRGAN ×12

Figure 1. Qualitative comparison for one scene. Rows 1 & 4: LR inputs; Rows 2 & 5: bicubic outputs; Rows 3 & 6: ESRGAN outputs.

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-XI-3-2026-35-2026



(a) Example 1 ×2 (b) Example 1 ×6 (c) Example 1 ×12

(d) Example 2 ×2 (e) Example 2 ×6 (f) Example 2 ×12

(g) Example 3 ×2 (h) Example 3 ×6 (i) Example 3 ×12

Figure 2. Extracted boundary overlays at multiple SR scales (×2, ×6, ×12) for three examples.

4.3 Cross-dataset validation

To further analyze the impact of SR on segmentation, a test case was conducted using two new data sources covering an unseen 240 m × 240 m region in southeastern New Brunswick, Canada. The first sample was captured by Pictometry in 2016, while the second was acquired by Maxar WorldView 2 (WV-02) in 2023 (Esri et al., 2023). The Pictometry image has a native resolution of 0.075 m/px, whereas WV02 has a native resolution of 0.6 m/px. When exported in ArcGIS, they were respectively resampled using cell sizes of 0.5 m and 2.0 m. This resulted in image sizes of 480 × 480 and 120 × 120, both covering the same ground footprint. To compare the impact of super-resolution on these two examples, their GT were manually delineated. As the HR sample already had sufficient resolution, it was segmented without upscaling. The LR sample was super-resolved using three scales: the ×2 scale was applied twice, the ×4 scale once and the ×6 scale, which was subsequently downsampled to 480 × 480. Table 3 presents the SR performance for both samples, with high-resolution 2016 (HR16) data and the 2023 low-resolution (LR23) data. The F1-Score (F1), eF1 and ePPA1 are presented alongside the Average Symmetric Surface Distance (ASSD) and Hausdorff Distance (HD). ASSD represents the average distance between predicted and GT pixels, while HD represents the maximum distance.

Figure 3 presents visual results comparing the 2016 sample, the original 2023 sample and the 2023 sample upsampled using SR-CNN ×4. Each image includes the GT and segmentation predictions for both years, as well as the super-resolved 2023 samples generated using the three methods. Table 4 presents the change in coastal boundary between 2016 and 2023 using the change in Area, ASSD and HD. It presents the GT of both samples (GT16, GT23), predictions using no upsampling (Pred16, Pred23) and results with upsampling.

5. Discussion

This study aimed to analyze the impact of SR on coastal boundary segmentation and extraction for high-resolution imagery.

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-XI-3-2026-35-2026 | © Author(s) 2026. CC BY 4.0 License.

(a) 2016. (b) 2023 (No Interpolation). (c) 2023 SRCNN ×4.

Figure 3. Comparison between 2016, 2023 (resized) and super-resolved 2023 (SRCNN ×4). © Esri, Maxar, Earthstar Geographics and the GIS User Community. Source: *ArcGIS World Imagery (WV-02)*. Map created using ArcGIS® software by Esri.

Table 3. Segmentation Performance of HR16 and LR23 Using Bicubic (B), SRCNN (S) and ESRGAN (E).

Method	**F1** (%)	**eF1** (%)	**ePPA@1** (%)	**ASSD** (px)	**HD** (px)
HR16	98.96	12.26	32.81	3.56	21.91
LR23	99.42	55.61	91.45	1.64	57.51
LR23-B (×2)	99.33	53.70	93.98	0.47	1.91
LR23-S (×2)	99.33	54.64	92.35	0.48	1.91
LR23-E (×2)	99.32	56.06	90.57	0.48	2.32
LR23-B (×4)	99.64	71.93	99.45	0.27	**1.37**
LR23-S (×4)	99.65	**74.19**	98.92	**0.25**	1.91
LR23-E (×4)	99.64	71.74	98.37	0.28	1.91
LR23-B (×6)	99.63	71.39	99.45	0.28	**1.37**
LR23-S (×6)	99.61	71.35	**99.46**	0.28	**1.37**
LR23-E (×6)	**99.65**	73.22	98.64	0.26	1.91

Table 4. Coastal Change (2016→2023) Using Bicubic (B), SRCNN (S) and ESRGAN (E) at ×4.

Pair	**ΔArea**	**ASSD**	**HD**
GT16 vs GT23	-1856	3.18	19.51
GT16 vs Pred23 (×1)	-2223	8.01	231.12
GT16 vs LR23-B (×4)	-2350	3.44	**20.06**
GT16 vs LR23-S (×4)	**-2142**	**3.27**	20.47
GT16 vs LR23-E (×4)	-2590	3.67	23.33
Pred16 vs GT23	-2699	4.86	21.91
Pred16 vs Pred23 (×1)	-3066	10.39	231.41
Pred16 vs LR23-B (×4)	-3193	5.55	25.60
Pred16 vs LR23-S (×4)	**-2985**	**5.29**	**24.65**
Pred16 vs LR23-E (×4)	-3433	5.68	26.02

Seven scaling rates were analyzed using two DL-based methods and compared against bicubic interpolation. Table 1 presents the SR performance using PSNR, SSIM, NIQE and LPIPS values for scaling factors of ×2, ×3, ×4, ×5, ×6, ×8, ×10 and ×12. Overall, both SR techniques outperformed bicubic interpolation, achieving higher metrics across most scales. ESRGAN achieved the largest improvement over bicubic interpolation at ×2, with a PSNR gain of 1.03. As expected, the difference in performance compared to bicubic interpolation declined in proportion to the scaling factor. This is caused by the inherent limitation of very low-resolution images, which lack information. Other metrics showed similar behaviour, with lower scales exhibiting a larger difference between SR and bicubic interpolation. All metrics also show that the overall performance of all methods decreases proportionally as the scaling factor increases. A visual comparison was also conducted, with Figure 1 presenting LR, bicubic and ESRGAN samples. This example illustrates that bicubic interpolation provides slightly smoother results than no interpolation. ESRGAN further improves the quality of the samples, providing much clearer pixel delineation. This effect is most evident at scales between ×4 and ×12, where ESRGAN produces sharper and more distinct boundaries. Although the differences in quantitative metrics decrease at higher scaling factors, the visual differences become more pronounced. This occurs because ESRGAN can reconstruct high-frequency details using learned priors from the training data, while bicubic interpolation is restricted to the information available in the sample. However, SR can introduce artifacts or hallucinations in samples, which can negatively impact model performance. This effect, although not visible in samples from this study, is noteworthy and could potentially hinder the implementation and applicability of SR in this domain. SRCNN and ESRGAN were chosen over other SR techniques, such as Real-ESRGAN and diffusion methods, for various reasons. The Real-ESRGAN implementation often limits the scaling factors to 2×, 4× and 8× due to the use of PixelShuffle. Diffusion models, on the other hand, vary in size but are often much larger than SRCNN and ESRGAN, making them less suitable for consumer-grade integration. A lightweight diffusion model was trained using similar parameters to SRCNN and ESRGAN. It required significantly more computational power and achieved PSNR values ranging from 20.12 to 17.43 for scales ×2 to ×12, with substantially more distorted visual performance. An in-depth benchmarking study of other DL-based methods, such as Real-ESRGAN and large diffusion models, would be valuable in future work.

SegFormer with MiT-B5 was trained on full-resolution samples for segmentation, enabling the analysis of the impact of SR on segmentation. Table 2 presents the segmentation performance for the HR test set and various upscaling methods. Overall, both DL-based SR methods outperformed bicubic interpolation, with a larger difference at higher scales. As expected, segmentation performance decreased as the scaling factor increased. This occurs because larger scaling factors correspond to more severe information loss in the input, which even DL models cannot fully recover. However, higher scales demonstrated the advantage of SR over bicubic interpolation. For the ×12 scale, bicubic interpolation achieved an F1-score of 65.27%, while SRCNN and ESRGAN achieved 89.26% and 89.92%, respectively. This demonstrates that at smaller scales, bicubic interpolation can be sufficient, but at larger scales, SR-based techniques are necessary. This is because DL-based SR reconstructs missing spatial details using learned features from data, while classical interpolation relies solely on the information present in the input sample. Interestingly, ESRGAN with

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-XI-3-2026-35-2026 | © Author(s) 2026. CC BY 4.0 License.

a scaling factor of ×2 achieved better results than the HR test set. This suggests that upsampling can provide sharper details, reduce noise, correct defects and normalize images affected by variations. SRCNN and ESRGAN performed similarly, with ESRGAN achieving higher IoU and F1-scores at higher scales, while SRCNN performed better at lower scales. Figure 2 presents three HR scenes with SR-based segmentation at scales ×2, ×6 and ×12. Overall, both SR-based solutions outperformed bicubic interpolation in terms of boundary delineation and consistency. The examples demonstrate that at higher scales, bicubic interpolation struggles, while SR-based techniques achieve more consistent results. This confirms the metric-based findings, reinforcing the benefits of SR techniques. However, in some samples, namely in Figure 2i, both SR methods struggled due to the sample consisting of a dense forest region. This indicates that segmentation models still encounter challenges in specific scenarios, particularly at very low scales. Although additional training could improve performance, a ×12 scaling factor remains a substantial increase. Nonetheless, both SR methods were still able to extract some coastal boundaries, while bicubic interpolation failed to do so.

To further validate our findings, we tested our approach on an unseen region using two new data sources. Table 3, which shows segmentation performance on the new region samples, indicates that ESRGAN and SRCNN outperformed bicubic interpolation. Specifically, SRCNN and ESRGAN achieved maximum eF1-scores of 74.19% and 73.22%, while bicubic achieved a maximum of 71.93%. Furthermore, to avoid favouring the SR-based models, the GT of LR23 was manually drawn at a very coarse resolution, which can favour bicubic interpolation. Nonetheless, SR demonstrated its ability to reduce noise in coarse and diverse samples, a common challenge in older satellite imagery. Table 4 and Figure 3 present the boundary changes over time using the two samples and ESRGAN. In Table 4, similar improvements can be observed in the area change, ASSD and HD when compared to the non-SR sample. When comparing both GT annotations, with GT 2023 resampled to the GT 2016 size using nearest-neighbor interpolation, an area loss of 1856 pixels is observed, with an ASSD of 3.18 pixels. Using a spatial resolution of 0.5 m per pixel, this corresponds to an average displacement of 1.59 m over seven years. This equates to a boundary change rate of 0.227 m per year, consistent with values reported for the region. When analyzing the 2016 prediction, it often caused issues due to a notable misprediction in the bottom right corner, as shown in Figure 3. When comparing GT 2016 against 2023 predicted samples, a clear benefit of SR can be observed, with the ASSD decreasing from 8.01 without SR to 3.27 using SRCNN ×4. These results, combined with the visual analysis, confirm that SRCNN ×4 achieved similar boundaries to the 2023 GT. Using both segmentation predictions, SRCNN ×4 also achieved the best results, with an ASSD of 5.29. The difference between using GT 2016 and predicted 2016, which increased the ASSD from 3.27 to 5.29, is consistent with the visual results and is attributed to a misprediction in the 2016 sample. This difference represents a change in erosion rate from 0.233 m per year using the GT to 0.377 m per year with the 2016 prediction. It shows that an incorrect prediction from the segmentation model on HR data can negatively impact the overall results. This limitation can be mitigated with larger datasets, greater data diversity and improved segmentation algorithms. Nonetheless, this experiment demonstrates that segmentation algorithms can be used to track erosion over time, with SR enabling a better representation than non-SR samples. This comparison also highlights a common issue in remote sensing data, as the two images differ significantly in quality and normalization, which can negatively impact segmentation results, even at similar resolutions. DL-based segmentation models depend heavily on the training data and if inference data differs substantially, performance can be negatively affected. Although SR addresses resolution discrepancies, techniques such as domain adaptation and style transfer could further improve cross-dataset generalization. Future work integrating these approaches into comparative and benchmarking studies would be valuable. Additionally, analyzing the impact of resolution, normalization, sensor modality and image quality through ablation studies would be beneficial for this field.

6. Conclusion

In conclusion, coastal areas are crucial for both human populations and the environment. They provide economic stimuli while serving as habitats, mating grounds and feeding areas for many species. However, these regions are at risk of erosion, which is accelerated by human activities. Conservation efforts can help mitigate and reduce coastal erosion but require extensive tracking and monitoring. Due to the vast extent of coastlines, automated solutions based on remote sensing currently offer the most viable approach. Many studies have already demonstrated the ability of DL-based segmentation to extract coastal boundaries. However, as remote sensing data span multiple decades, their resolution and quality are often suboptimal. Therefore, we proposed analyzing the impact of SR on coastal boundary segmentation using high-resolution imagery. Our study also examined the performance of our approach on different data sources with varying resolutions and quality. Overall, we demonstrated that SR outperforms classical bicubic interpolation for image upscaling. We also showed that this effect becomes more significant when using scaling factors between ×4 and ×12. In fact, we achieved a substantial increase in segmentation performance at a scale of ×12. Visually, our approach enabled clearer boundary delineation, resulting in improved segmentation accuracy. When applied to coarse-resolution imagery from an unseen region, SR-based techniques also outperformed bicubic interpolation. Furthermore, we demonstrated that the proposed approach was able to track a visible coastal boundary recession between 2016 and 2023. The integration of high-resolution imagery provides additional insight into the effect of SR on segmentation performance. Finally, the inclusion of new data sources illustrates the ability of SR and DL-based segmentation to generalize to real-world applications. However, limitations remain in this field, particularly regarding datasets, boundary definitions and enhancement methods. Increasing dataset size, diversity and temporal aspects would be beneficial in future work. Integrating large diffusion-based SR models and newer SR models would also strengthen this field.

7. Funding

This research was made possible in part by the support provided by the Natural Sciences and Engineering Research Council of Canada (NSERC), funding reference number RGPIN-2024-05287.

References

Blais, M.-A., Akhloufi, M. A., 2021. Deep learning for low altitude coastline segmentation. *Ocean Sensing and Monitoring XIII*, 11752, SPIE, 103–111.

Blais, M.-A., Akhloufi, M. A., 2025. Advances in Remote Sensing and Deep Learning in Coastal Boundary Extraction for Erosion Monitoring. *Geomatics (2673-7418)*, 5(1).

Chaurasia, A., Culurciello, E., 2017. Linknet: Exploiting encoder representations for efficient semantic segmentation. *2017 IEEE visual communications and image processing (VCIP)*, IEEE, 1–4.

Chen, L.-C., Papandreou, G., Schroff, F., Adam, H., 2017. Rethinking atrous convolution for semantic image segmentation. *arXiv preprint arXiv:1706.05587*.

Cosby, A., Lebakula, V., Smith, C., Wanik, D., Bergene, K., Rose, A., Swanson, D., Bloom, D., 2024. Accelerating growth of human coastal populations at the global and continent levels: 2000–2018, sci. rep., 14, 22489.

Dong, C., Loy, C. C., He, K., Tang, X., 2015. Image super-resolution using deep convolutional networks. *IEEE transactions on pattern analysis and machine intelligence*, 38(2), 295–307.

Esri, Maxar, Geographics, E., Community, G. U., 2023. Arcgis world imagery (wv-02). `https://www.arcgis.com/home/item.html?id=10df2279f9684e4a9f6a7f08febac2a9`. Esri, Redlands, CA. Accessed: 2025-10-15.

Fernandez-Beltran, R., Latorre-Carmona, P., Pla, F., 2017. Single-frame super-resolution in remote sensing: A practical overview. *International journal of remote sensing*, 38(1), 314–354.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770–778.

Khurram, S., Pour, A. B., Bagheri, M., Ariffin, E. H., Akhir, M. F., Hamzah, S. B., 2025. Developments in deep learning algorithms for coastline extraction from remote sensing imagery: a systematic review. *Earth Science Informatics*, 18(3), 292.

Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z. et al., 2017. Photo-realistic single image super-resolution using a generative adversarial network. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 4681–4690.

Lei, D., Luo, X., Zhang, Z., Qin, X., Cui, J., 2025. Research on Super-Resolution Reconstruction Algorithms for Remote Sensing Images of Coastal Zone Based on Deep Learning. *Land*, 14(4), 733.

Li, J., Bhatti, U. A., Nawaz, S. A., Huang, M., Ahmad, R. M., Ghadi, Y. Y., 2024. Remote-sensing image classification: A comprehensive review and applications. *Deep learning for multimedia processing applications*, 18–47.

Li, X., Cooper, J. R., Plater, A. J., 2021. Quantifying erosion hazards and economic damage to critical infrastructure in river catchments: Impact of a warming climate. *Climate Risk Management*, 32, 100287.

Lin, T.-Y., Dollár, P., Girshick, R., He, K., Hariharan, B., Belongie, S., 2017. Feature pyramid networks for object detection. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2117–2125.

McLachlan, A., Wooldridge, T., Schramm, M., Kühn, M., 1980. Seasonal abundance, biomass and feeding of shore birds on sandy beaches in the Eastern Cape, South Africa. *Ostrich*, 51(1), 44–52.

Pala, İ., Algancı, U., 2025. Investigating the performance of super-resolved remote sensing images on coastline segmentation with deep learning based methods. *International Journal of Engineering and Geosciences*, 10(1), 93–106.

Pang, T., Wang, X., Nawaz, R. A., Keefe, G., Adekanmbi, T., 2023. Coastal erosion and climate change: A review on coastal-change process and modeling. *Ambio*, 52(12), 2034–2052.

Pasquali, D., Marucci, A., 2021. The effects of urban and economic development on coastal zone management. *Sustainability*, 13(11), 6071.

Phillips, M. R., Jones, A. L., 2006. Erosion and tourism infrastructure in the coastal zone: Problems, consequences and management. *Tourism Management*, 27(3), 517–524.

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*, Springer, 234–241.

Seitz, R. D., Wennhage, H., Bergström, U., Lipcius, R. N., Ysebaert, T., 2014. Ecological value of coastal habitats for commercially and ecologically important species. *ICES Journal of Marine Science*, 71(3), 648–665.

Silvano, R. A., MacCord, P. F., Lima, R. V., Begossi, A., 2006. When does this fish spawn? Fishermen’s local knowledge of migration and reproduction of Brazilian coastal fishes. *Environmental Biology of fishes*, 76(2), 371–386.

Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

Smedes, H. W., Pierce, K. L., Tanguay, M. G., Hoffer, R. M., 1970. Digital computer terrain mapping from multispectral data. *Journal of Spacecraft and Rockets*, 7(9), 1025–1031.

United Nations Conference on Trade and Development, 2025. Shipping data: Unctad releases new seaborne trade statistics. `https://unctad.org/news/shipping-data-unctad-releases-new-seaborne-trade-statistics`. Accessed: 22 September 2025.

Vikas, M., Dwarakish, G., 2015. Coastal pollution: a review. *Aquatic procedia*, 4, 381–388.

Wang, X., Yu, K., Wu, S., Gu, J., Liu, Y., Dong, C., Qiao, Y., Change Loy, C., 2018. Esrgan: Enhanced super-resolution generative adversarial networks. *Proceedings of the European conference on computer vision (ECCV) workshops*, 0–0.

Xie, E., Wang, W., Yu, Z., Anandkumar, A., Alvarez, J. M., Luo, P., 2021. SegFormer: Simple and efficient design for semantic segmentation with transformers. *Advances in neural information processing systems*, 34, 12077–12090.

Zhou, X., Wang, J., Zheng, F., Wang, H., Yang, H., 2023. An overview of coastline extraction from remote sensing data. *Remote Sensing*, 15(19), 4865.