Improving Segmentation Pipelines for Surfzone Rock Mapping

Alex Washburn, Ashton Pihl

Abstract—Rocky shores make up the majority of coastline worldwide and support diverse ecosystems, including invertebrates, intertidal plants, and algae, whose distribution depends on wave breaking and circulation. However, our understanding of how surface-piercing rocks—common features of rocky shores which are irregular in shape—impact hydrodynamics in the surfzone remains poor. To address this, we developed and evaluated different methods to achieve the most accurate rock segmentation possible from high resolution aerial imagery. Three Object-Based Image Analysis (OBIA) approaches were evaluated that progressively incorporated connected component grouping, morphological fragment healing, and spectral smoothing to improve rock segmentation. These results were compared to the Segment Anything Model (SAM), a state-of-the-art pretrained model. We found that SAM performed best on area-based metrics, whereas OBIA methods preserved rock edges best and achieved perfect cardinality. These findings provide a transferable method for segmenting rocks in the surfzone from aerial imagery, supporting further research into wave breaking and circulation patterns along rocky coasts.

Index Terms—Remote Sensing, Image Segmentation, Rock Morphology, Nearshore Hydrodynamics

1 Introduction

The the use of remote sensing as a tool to observe the ocean began in the 1970s with orbiting satellites, enabling oceanographers to make advances in our understanding of basin-scale and mesoscale patterns in the ocean such as the circulation of heat, salinity, and nutrients [1]. However, the limited spatial resolution of these data implied that small-scale coastal processes such as beach evolution required expensive in-situ sensor deployments. Remote sensing techniques with greater spatial resolution—including aircraft-based observations—began being used near the turn of the century and have substantially expanded the applications of remote sensing to smaller-scale processes in the coastal environment [2].

Currently, remote-sensed products exist with pixel resolutions of O(1m), which have several applications for small-scale oceanographic research, such as quantifying rates of coastal erosion [3] and estimating sea level rise by quantifying the position of the shoreline and how it changes over time [4]. Much of existing coastal remote-sensing research has focused on sandy beach processes [5], while comparatively little effort has gone into understanding coastal processes along rocky coastlines. Given that the rocky shoreline environment is highly dynamic and difficult to navigate, in-situ sensor deployments are dangerous and observational data are limited. Remote sensing and image processing, therefore, serves as a particularly useful tool for studying rocky shores.

Comprising 75% of coastline worldwide, rocky shores support diverse ecosystems of invertebrates, intertidal plants, and algae. The stability and connectivity of these ecosystems depends on wave breaking and the subsequent water circulation it generates. Rocks located in the surfzone (the region of wave breaking) have been shown to control wave breaking and circulation patterns [6]. The size, surface roughness, and shape of rocks can modify flow structure in the surfzone by causing flow deceleration, wake formation, turbulence, and more [7]. Therefore, to understand the impact of rocks on wave breaking and circulation patterns in the rocky shoreline environment, an accurate identification of characteristic morphological features on rocky shores near the region of breaking waves is necessary. Precise rock segmentation in the surfzone, allowing us to compute important statistics such as the density and size of surface-piercing rocks, is an important component of this.

- *A. Pihl and A. Washburn are with the Department of Civil and Environmental Engineering, Stanford University, Stanford, CA, 94305.*

2 Background and Related Work

Publicly available segmentation pipelines exist to segment various coastal features such as sand, whitewater, and deeper offshore water [5]. However, these approaches are limited to sandy beaches, and lack the functionality to segment rocky features, such as the position of surface-piercing rocks. Yet, there are several segmentation techniques that are applied broadly to satellite imagery. Two of these—Object-Based Image Analysis (OBIA) and The Segment Anything Model (SAM)—are discussed below.

2.1 Object-Based Image Analysis (OBIA)

OBIA identifies objects of interest by taking in a pixel-level classification and grouping pixels into meaningful shapes which can be analyzed based on geometry, color, relative location, and more [8]. This extends beyond a pixel-based approach—which is also commonly used for remote sensed image segmentation by doing thresholding and segmentation purely in feature space [9]—by taking into account spatial context to better determine whether a candidate group of pixels represents the object of interest.

2.1.1 Past Application of OBIA for General Aerial Image Segmentation

OBIA has been used extensively to segment objects in both urban and non-urban aerial imagery. For example, past work using both satellite and aerial imagery has shown that it is effective in segmenting water bodies, low vegetation/grass, trees, buildings, and roads [10,11].

2.1.2 Applicability of OBIA for Rock Segmentation Along Rocky Shores

While OBIA has not been used to date to segment rocks in the coastal environment, it is a good candidate method for rock segmentation due to its unique ability to integrate fragments of rocks into a complete rock object using morphological operations. Segmenting rocks in the surfzone is challenging because breaking waves rush over rocks, making a singular rock appear as a set of several small, fragmented rocks, where fragments are separated by whitewater. To address this, OBIA supports morphological operations such as erosion and dilation, which have been used for applications such as expanding fragmented segments of macrophages to reconstruct a complete macrophage [12]. In the context of rock segmentation, applying morphological erosion and dilation can aggregate fragmented rocks into a singular desired rock by filling small internal gaps while preserving the edge structure of the rocks using erosion.

2.2 Segment Anything Model (SAM)

The Segment Anything Model (SAM) is a pre-trained segmentation model trained on a dataset with over 1 billion masks, allowing it to perform zero-shot segmentation and automatic mask generation for objects in images without providing it with ground truth masks of the objects of interest [13]. SAM has been shown to perform segmentation at a level at or above that of supervised models [14]. Additionally, it is particularly advantageous when ground truth data is lacking and laborious to produce.

2.2.1 Past Applications of SAM for General Aerial Image Segmentation

Given its generalizability, SAM has been used in a wide range of aerial imagery segmentation applications spanning the geosciences. Applications of particular interest include segmenting images of rocks drilled underwater [15], classification of coastal wetland regions [16], and tracking of floating marine debris in satellite images via segmentation [17]. To our knowledge, the application of SAM to segmentation tasks involving coastal imagery is limited to date.

2.2.2 Applicability of SAM for Rock Segmentation Along Rocky Shores

Beyond being considered the state-of-the-art approach to segmenting imagery in general, SAM is especially useful in coastal rock segmentation due to the difficult nature of creating data for supervised training. Rocks are high density, ubiquitous across the coast of California, and fractal in shape, implying that manually labeling ground truth rock data for training is extremely time consuming. Given its zero-shot nature, SAM does not require training data, and thus can segment large stretches of rocky shorelines easily.

3 Theory and Method

3.1 Data and Models

For rock segmentation, 0.35-meter publicly available aerial imagery of the coastline at Garrapata State Park, CA was captured by aircraft and downloaded from Google Earth. The image used for segmentation in this paper was captured during February 2021. The SAM1 model is publicly available as well [13].

3.2 Surfzone Segmentation using Otsu's Thresholding

The first step in segmenting rocks in the region of breaking waves is to isolate the surfzone. To do so, a spectral score is computed per-pixel as shown below.

$$score = (B + 0.7G) - 1.1R \tag{1}$$

Given that whitewater is high in B and G relative to R, this score creates a bimodal intensity distribution. Otsu's thresholding method is then applied to this score: by minimizing the intra-class variance between the whitewater and non-whitewater distributions, it optimizes a separation threshold to distinguish between whitewater and non-whitewater pixels. From here, morphological operations such as hole-filling of small gaps in whitewater are done, and the largest connected component of whitewater is returned as the surfzone.

3.3 Rock Pixel Detection

Rock pixel detection is performed within the surfzone only. To obtain a rock pixel classification in the surfzone, the same spectral score is used as in equation (1) along with Otsu's thresholding method, as this score is high for whitewater and low for rocks.

3.4 Rock Segmentation using OBIA

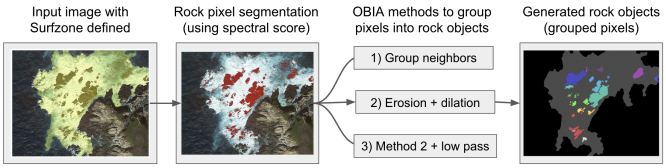

Fig. 1. Rock segmentation pipeline for three different OBIA approaches.

Given an input of a binary pixel mask within the surfzone, we performed OBIA, grouping pixels together to form rock objects following three separate methods (Figure 1). In OBIA 1, rocks are formed by grouping adjacent pixels into rock objects using 8-pixel connectivity, with contiguous regions representing rocks. OBIA 2 builds upon OBIA 1 by using morphological erosion and dilation to heal small rock fragments from OBIA 1 that are assumed to be part of a broader fragmented rock into one region. Furthermore, detected rocks were filtered: all classified rocks were required to surpass minimum area and extent and maximum eccentricity thresholds. In OBIA 3, we extend the approach from OBIA 2 (erosion + dilation + geometric filtering) but Gaussian low-pass filter the spectral image to remove noise and false positives. This smoothing blurs rock edges; to

restore edges we use the original unblurred mask in the broad, low-pass filtered mask to reduce blur. In OBIA 3, the standard deviation of the Gaussian blur kernel was optimized (§5) using segmentation metrics discussed in §4.

3.5 Rock Segmentation using SAM

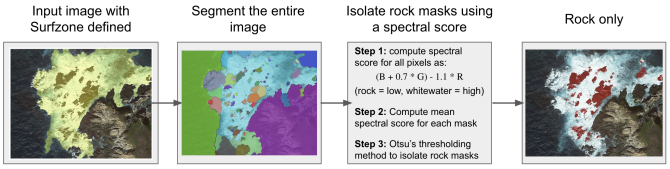

Fig. 2. Rock segmentation pipeline for Segment Anything Model (SAM).

SAM takes in an RGB image—in our case the surfzone—as input and produces candidate masks of detected objects in the surfzone as the output. When running SAM, we chose a dense sampling grid (64x64 points) to detect objects spanning spatial scales. Candidate masks were further filtered using minimum thresholds for predicted IoU, stability score, and minimum area.

Although SAM supports promptable segmentation for objects of interest, it failed to successfully do so for "surface rocks." Instead, to isolate the rock masks, the same spectral score as in equation (1) was used which produces high values for whitewater and low values for rocks. Here, each mask was given a spectral score computed as the mean score of all of its pixels; Otsu's thresholding method was used to classify segments as rock or other.

4 Analysis and Evaluation

To evaluate and compare pipeline performances, ground truth data of rock segments were hand-crafted, consisting of 143 ground-truth labeled rocks. Our analysis and evaluation focused on a subset of the entire image as it was easier to visualize errors, yet we evaluated segmentation on the entire labeled region to assess whether our methods are generalizeable.

Standard semantic segmentation error metrics were selected to evaluate errors stemming from pixel misidentification, boundary misrepresentation, and fictitious and missed groupings [18]. Standard set notation is assumed when defining these metrics with P denoting the predicted rock mask pixels and G denoting the set of ground truth rock mask pixels. Qualitative illustrations of errors are also explained here.

4.1 Qualitative Evaluation Metrics

In order to visually identify flow conditions and rock morphological features that may lead to errors in the pipeline outputs, masks of false positives and negatives are plotted alongside the rock masks for a region with the most variation in rock sizes.

4.2 Quantitative Evaluation Metrics

The first two evaluation metrics belong to a class of metrics called single-threshold counting (overlap-based) metrics which operate directly on the confusion matrix and express the metric value as a function of the pixel cardinalities. They generally measure total area error.

4.2.1 Dice Similarity Coefficient (DSC)

$$\mathrm{DSC}(P,G) = \frac{2|P \cap G|}{|P| + |G|} \tag{2}$$

DSC measures the ratio of the intersection of prediction and ground truth to the total area of the prediction and ground truth masks and is bounded from zero to one [19]. It gives greater weight to true positive pixels than false positives and negatives. Here, a DSC close to zero suggests that there is zero overlap between the predicted and ground truth pixel sets while one suggests a perfect overlap.

4.2.2 Intersection over Union (IoU)

$$\mathrm{IoU}(P,G) = \frac{|P \cap G|}{|P \cup G|} \tag{3}$$

IoU measures the ratio of the intersection of prediction and ground truth pixels to the union of the prediction and ground truth pixel sets and is bounded from zero to one [20]. In contrast to DSC, it gives equal weight to true positives and false positives and negatives. Similarly to DSC, an IoU close to zero suggests that there is zero overlap between the predicted and ground truth pixel sets while one suggests a perfect overlap.

4.2.3 Hausdorff Distance (HD)

$$d_H(P,G) = \max\left\{ \sup_{p\in P} \inf_{g\in G} \|p-g\|, \sup_{g\in G} \inf_{p\in P} \|g-p\| \right\} \tag{4}$$

The HD belongs to the distance-based metrics class which operate exclusively on the true positives and rely on the explicit definition of object boundaries [18]. It measures the greatest of all of the distances from a point on the boundary of one set to the closest point in the other set for each rock (in this case) and is the L_∞ norm. A small HD means the edge features of the predicted set is well reproduced from the ground truth, while a large HD means the edges are not an accurate representation of the ground truth edge. The median and distribution of HD values are reported and obvious individual rock grouping mismatches were thrown out.

4.2.4 Average Surface Distance (ASD)

$$\mathrm{ASD}(P,G) = \frac{1}{|P| + |G|} \left(\sum_{p\in P} \min_{g\in G} \|p-g\| + \sum_{g\in G} \min_{p\in P} \|g-p\| \right) \tag{5}$$

The ASD is an L_1 norm representation of the HD [21]. It is an average value of the minimum distance between a point on the boundary of one set to the closest point in the other set for each rock. This is a more robust estimate of overall rock dilation or shrinkage than the HD. The median and distribution of HD values are reported and obvious individual rock grouping mismatches were thrown out.

4.2.5 Fragmentation Ratio (FR)

$$\mathrm{FR} = \frac{N(P)}{N(G)} \tag{6}$$

The FR was designed for this purpose to evaluate the ratio between the number of predicted rocks to ground truth rocks. An ideal FR is one where the number of predicted and ground truth rocks are equal, while an FR greater than one and less than one indicate over and under prediction respectively.

5 Results

5.1 Rock Masks and Qualitative Error

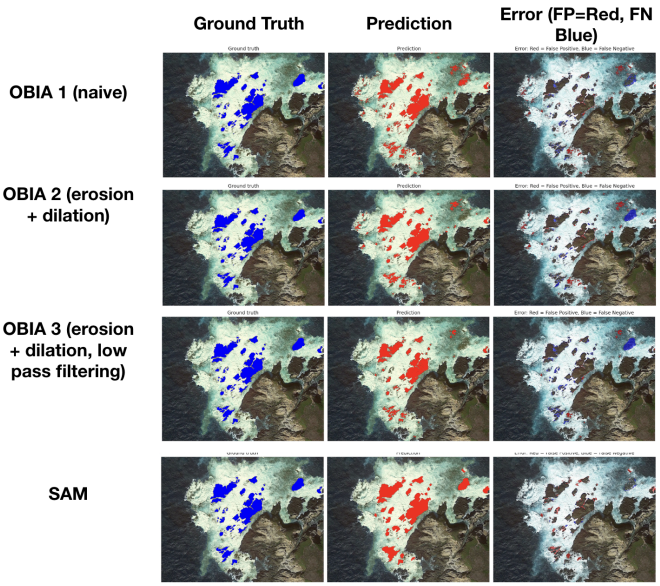

Fig. 3. Ground truth and rock pixel prediction shown for each pipeline rock mask output, along with false positives (red) and false negatives (blue) highlighted in the third column.

The rock mask output of each pipeline along with the false positive and negative pixels are shown in Figure 3 for the region of interest. Notably, SAM avoids the large false positive and negative patches in the upper right quadrant of the image.

5.2 Quantitative Error

TABLE 1
Quantitative metrics described in §4 for each pipeline on the region of interest, where HD and ASD are median values.

Pipeline	DSC	IoU	HD (pixels)	ASD (pixels)	FR
OBIA 1	0.822	0.698	7.62	1.59	26.0
OBIA 2	0.812	0.683	5.36	0.61	1.37
OBIA 3	0.762	0.615	12.1	2.56	1.00
SAM	0.889	0.800	12.8	2.46	0.76

The quantitative metrics for the region of interest are reported in Table 1. An inverse relationship exists between the number of steps in the OBIA pipeline and both area-based metrics (DSC, IoU). SAM significantly performs the best on the area-based metrics. OBIA 2 has the best performance given the distance-based metrics (HD, ASD). The FR decreases with increasing number of steps in the OBIA pipeline to a perfect score of one due to low-pass filter optimization. The standard deviation of the Gaussian blur kernel that yielded the optimal FR was found to be four pixels (Figure 4). Unlike the OBIA cases, SAM underestimates the number of ground truth rocks.

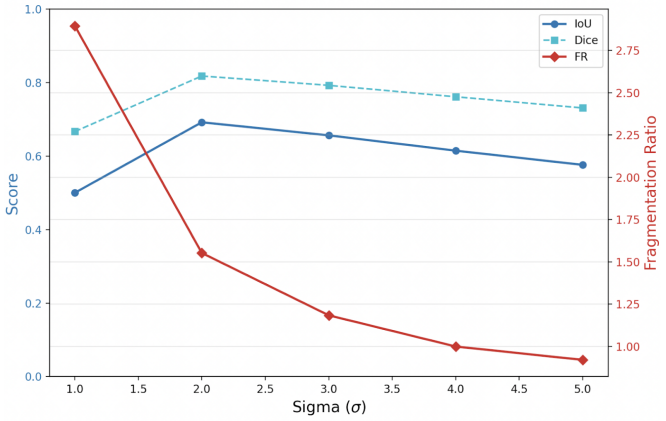

Fig. 4. Optimization for the standard deviation for the Gaussian kernel used in low pass filtering for OBIA 3, optimized on data for the small region.

5.2.1 Hausdorff Distance and Average Surface Distance Distributions

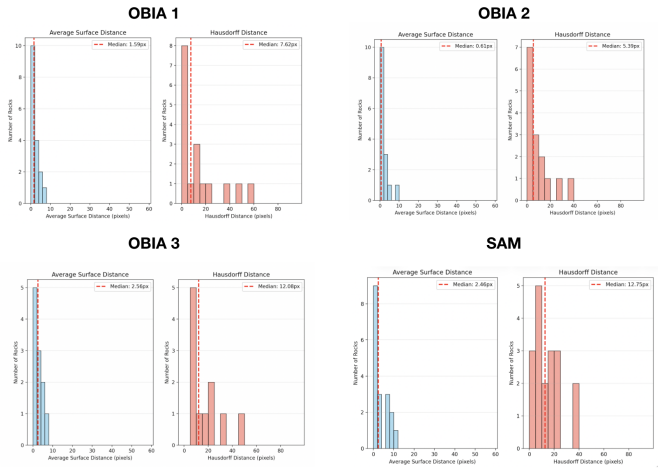

Fig. 5. Distribution of Hausdorff distances and average surface areas for each rock in the region of interest for all pipelines.

The distributions of HDs and ASDs for all rocks in the region of interest are displayed in Figure 5.

Given this bin size, the distributions appear to be exponentially or lognormally distributed. The SAM distributions have greater weight in the tail values than the OBIA distributions.

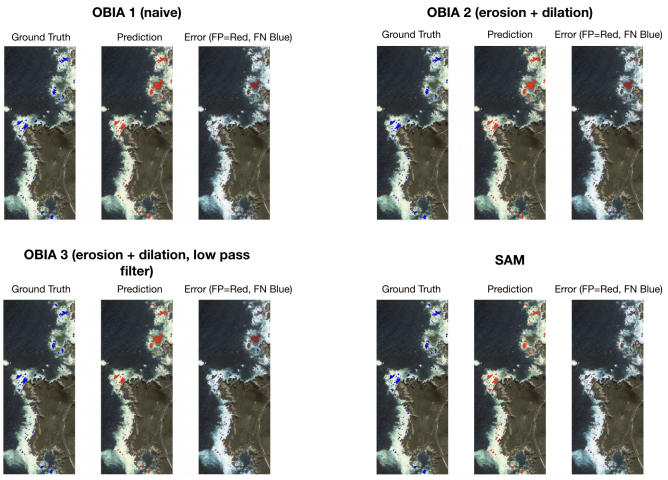

Fig. 6. Ground truth and rock pixel prediction for the entire coastline for each pipeline rock mask output, along with false positives (red) and false negatives (blue) highlighted in the rightmost subplot for each method.

5.3 Assessing Generalizability to the Entire Coastline

Although our analysis focused on evaluating segmentation pipelines on a small region, we segmented surfzone rocks along the entire coastline to assess how our segmentation methods generalize. Qualitative results are shown in Figure 6. Similarly to the small region we evaluated, SAM is particularly effective in avoiding false positives.

TABLE 2
Quantitative metrics described in §4 for each pipeline on the entire coastline, where HD and ASD are median values.

Pipeline	DSC	IoU	HD (pixels)	ASD (pixels)	FR
OBIA 1	0.622	0.4508	6.32	1.55	49.68
OBIA 2	0.603	0.431	7.45	1.66	2.17
OBIA 3	0.578	0.407	9.03	2.36	1.03
SAM	0.788	0.651	8.06	1.99	0.34

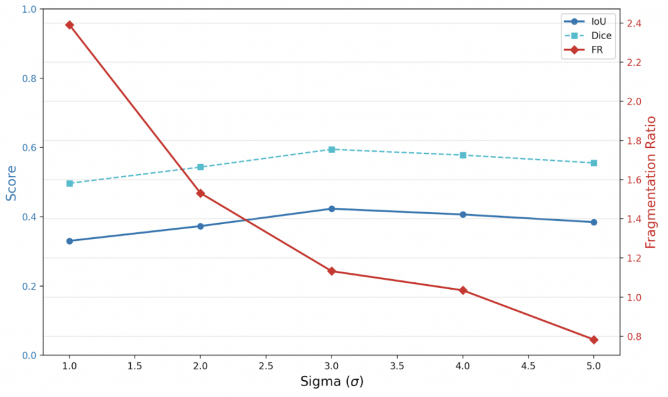

Fig. 7. Optimization for the standard deviation for the Gaussian kernel used in low pass filtering for OBIA 3, optimized on data for the entire coastline.

Quantitatively, evaluation metric trends for the entire coastline (Table 2) are largely consistent with what we observed for the smaller region (Table 1). Specifically, with increased steps in the OBIA pipeline, DSC and IoU decrease, whereas SAM performs better on these area-based metrics than any of the OBIA methods. The standard deviation of the Gaussian blur kernel corresponding to the optimal FR was also found to be four pixels (Figure 7), just as for the small region. OBIA 3 achieves near-perfect FR, whereas OBIA 1 and 2 overpredict the number of ground truth rocks and SAM underpredicts them.

6 Discussion and Conclusion

For the task of segmenting rocks in the surfzone, both OBIA-based and SAM approaches are successful, yet their performance varies based on the segmentation metric being analyzed. The metrics analyzed were chosen such that they would comment on various aspects of segmentation performance, and Table 1 illustrates that each pipeline tends to excel in some aspects given a tradeoff with others. These tradeoffs inform when each segmentation method may be appropriate to use, and how each method can be mechanistically improved.

6.1 Comparison of Pipelines

The thresholding only case (OBIA 1) acts as a naive reference state for this task. Given the inherent difficulty of matching the ground truth rock mask due to high-frequency morphological features and noise from whitewater, OBIA 1 performs well on the area-based metrics. It is able to resolve high-frequency rock edges relatively well as no smoothing steps were taken, however this results in several fictitious low area rocks which inflate the fragmentation ratio.

To address this issue, erosion and dilation was used in OBIA 2. This resulted in the best distance-based metric performance across all pipelines suggesting it has the best edge matching capabilities. The area-based metrics are slightly degraded but not significantly. The fragmentation ratio is greatly improved, but still overestimates the number of rocks present. This suggests that erosion and dilation is an effective remedy for reducing high-frequency small area rock artifacts that result from thresholding only.

In order to address the overestimation and false positives issue, an optimized low-pass filter was introduced in OBIA 3 which resulted in a perfect rock cardinality prediction. The area-based metrics were slightly degraded but continue to be relatively high for this use case. The distance-based metrics were the major tradeoff resulting in significantly worse performance than the naive case. This suggests that OBIA 3 is an effective tool for assessing rock cardinality but not necessarily rock morphological statistics.

SAM was much more effective than the OBIA pipelines for the area-based metrics. However, SAM had worse distance-based metrics than OBIA 3 and underpredicted the rock cardinality. This is likely a result of SAM closing high-frequency (narrow and long) gaps in rocks which would not significantly alter area-based metrics but would significantly worsen the distance-based metrics. This suggests that SAM is an effective tool for identifying rocks and preserving their area, but lacks effective edge-detection and ability to identify real small area rocks.

6.2 Extrapolating to Future Use Cases

Publicly available high resolution imagery is only available in select areas of the globe. Lower resolution imagery is effectively low-pass filtered such that high-frequency rock features like cracks and edges are not preserved. The results suggest that SAM would have the best performance as it preserves and identifies rock area accurately and its limitations with edge and high-resolution detection wouldn't degrade distance-based metrics for low-resolution imagery.

In use cases where accurate cardinality is necessary, OBIA 3 is the obvious choice for its optimization on the fragmentation ratio. In use cases where getting morphological statistics is most important, OBIA 2 would likely be the best choice for its ability to preserve edges and high-frequency features while retaining high area identification accuracy.

6.3 Limitations and Future Work

The rock segmentation methods described in this paper are optimized for high-resolution aerial imagery, which, as described above, is largely concentrated around the Western United States and Europe. Coastal rock segmentation in regions with only lower-resolution imagery available opens up future research in image resolution enhancement using methods such as Gaussian deconvolution [22]. Additionally, rock segmentation depends on the presence of whitewater; if wave breaking is minimal and whitewater is not present, this method may not work. Future work involves using time-averaged imagery rather than an instantaneous images for segmentation, which increases the likelihood of abundant whitewater in the surfzone.

6.4 Conclusion

In this paper, OBIA methods and SAM were modified for the purpose of segmenting rocks in the surfzone. The thresholding only OBIA method was found to perform relatively well on area and distance-based metrics, but poorly on cardinality. The cardinality problem was solved by increasing the number of steps in the pipeline with erosion-dilation and low-pass filtering which in turn slightly degraded the area metrics. The erosion-dilation step increased distance-based performance, but the low-pass filter degraded it. SAM performed better than OBIA on the area-based metrics, but suffered on the distance-based metrics and underestimated rock cardinality. SAM was found to be a promising candidate for future use on lower resolution imagery.

Acknowledgments

The authors would like to thank our project mentor and course instructor Dr. Gordon Wetzstein for help in many office hour chats.

References

[1] J. J. Lee et al., "Advances in Coastal Monitoring and Prediction," *Meteorol. Monogr.*, vol. 59, no. 1, pp. 1–25, 2019. [Online]. Available: https://journals.ametsoc.org/view/journals/amsm/59/1/amsmonographs-d-18-0010.1.xml

[2] A. H. Sallenger et al., "Evaluation of Airborne Topographic Lidar for Measuring Beach Changes," *J. Coastal Res.*, vol. 19, no. 1, pp. 125–133, 2003.

[3] C. Holzner et al., "Emerging Trends in Biomedical Image Analysis," *PMC*, 2025. [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC11873248/

[4] S. Vitousek et al., "The future of coastal monitoring through satellite remote sensing," *Cambridge Prisms: Coastal Futures*, vol. 1, 2022.

[5] K. Vos et al., "CoastSat: A Google Earth Engine-enabled Python toolkit to extract shorelines from publicly available satellite imagery," *Environ. Model. Softw.*, vol. 122, p. 104528, 2019.

[6] J. MacMahan et al., "Breaking Wave Observations in the Surfzone," *J. Geophys. Res. Oceans*, vol. 128, no. 5, 2023.

[7] J. Fang et al., "A deep learning approach for saltmarsh vegetation mapping," *J. Geophys. Res. Earth Surf.*, vol. 122, no. 6, 2017.

[8] T. Blaschke et al., *Object-Based Image Analysis: Spatial Concepts for Knowledge-Driven Remote Sensing Applications*. Springer, 2008.

[9] J. Schiewe, "Segmentation of high-resolution radar imagery," *ISPRS Proc.*, vol. 34, 2002.

[10] M. Li and Z. Shao, "Object-based bidirectional method for multiscale image segmentation," *Remote Sens.*, vol. 6, no. 11, pp. 11372–11391, 2014.

[11] X. Zou et al., "A novel object-oriented method for rock detection," *ISPRS Arch.*, vol. 41, 2016.

[12] S. Park and K. Mikula, "Automated completion of segmented fragments of macrophages using weighted dilation and erosion in 2D+time microscopy videos," *ResearchGate*, 2024.

[13] A. Kirillov et al., "Segment Anything," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2023, pp. 4015–4026.

[14] J. Pei et al., "Recent Advances in AI-driven Image Analysis," *AI Res.*, 2025.

[15] L. Shan et al., "Advanced Geological Segmentation Techniques," *AGER*, 2024.

[16] Y. Zou et al., "Deep learning for coastal landform classification," *J. Appl. Geophys.*, 2025.

[17] R. Van Dalen et al., "Robotic Vision in Coastal Environments," *IEEE Trans. Rob.*, 2025.

[18] A. Reinke et al., "Common Limitations of Image Processing Metrics: A Picture Story," *arXiv preprint*, 2022. [Online]. Available: https://arxiv.org/abs/2104.05642

[19] A. A. Taha and A. Hanbury, "Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool," *BMC Med. Imaging*, vol. 15, no. 29, 2015.

[20] V. Yeghiazaryan and I. Voiculescu, "Family of Hausdorff distances for segmentation dynamics," *Pattern Recognit.*, vol. 75, pp. 1–15, 2018.

[21] J. Ma et al., "Loss odyssey in medical image segmentation," in *Proc. Med. Imaging Deep Learn. (MIDL)*, 2020.

[22] A. Jalobeanu et al., "Satellite image deconvolution using complex wavelet packets," *IEEE Trans. Image Process.*, vol. 13, no. 6, pp. 804–817, 2004.