Edge-Attention U-Net for Shoreline Detection

Vihaan Akshaay
University of California, Santa Barbara
Santa Barbara, CA 93117
vihaanakshaay@ucsb.edu

Aman Sariya
University of California, Santa Barbara
Santa Barbara, CA 93117
amansariya@ucsb.edu

Abstract

Coastline mapping and change detection are crucial for safe navigation, resource management, environmental protection and sustainable coastal development and planning. This paper proposes a new methodology for extracting coastlines from images. This is based on semantic segmentation and it includes several steps and methods to improve the accuracy of the extracted coastlines. We propose a new Edge-based Attention module called Edge Attention. We integrate this with the traditional Attention U-Net to observe improvement in the aforementioned method. All the models are tested on the Sentinel-2 Water Edges Dataset to output segmentation masks and classify pixels as land or water pixels. We then perform a qualitative and quantitative analysis of these models and present the segmentation evaluation metrics to find that our model performs better, if not, at par with the existing models. Link to the GitHub repository: https://github.com/VihaanAkshaay/EdgeAttn-U-Net

1. Introduction

With global warming having an increasingly devastating impact on the environment and weather conditions, it is important to focus on safeguarding communities near the coastline. Computer vision based shoreline detection can be used for:

- Environmental monitoring: Identifying erosion, sedimentation, and sea-level rise can help in early warning systems for potential disasters and also in identifying areas that are at high risk.
- Coastal management: Identifying areas that need protection from coastal erosion or where restoration projects are needed.
- Navigation and safety: Identifying navigational hazards, such as submerged rocks or sandbars, that can be a danger to ships and boats.
- Urban planning: Assess the impact of coastal development on the coastline and to plan for sustainable development.

There is now a chance to build deeper learning-based shoreline detection systems that are more precise and reliable thanks to recent improvements in the quality and accessibility of remotely sensed information, including satellite and aerial imagery.

Despite being an important problem, it does not seem like it gets enough attention as all the state-of-the-art results for a standardized dataset still use traditional segmentation models. While they work well on generic segmentation tasks, the do not generalize well for specific tasks. We wish to target this problem statement by building and testing custom models. We try to bring the intuition that humans use while having to segment regions as land and sea by considering boundaries as an important deciding criteria rather than evaluating pixel-by-pixel. To use this idea we leverage edge information at various feature abstraction levels to perform semantic segmentation.

Semantic segmentation associates a label or category with every pixel in the image. For our purposes, we have two categories: 0 for land pixels and 1 for water pixels.

2. Related Work

2.1. Shoreline Detection

Previous work [5] involved using image processing techniques along with active contours for coastline modeling. Active contours are deformable curves that move to the edges of objects in the image. The method was based on a deformable curve or contour that moves iteratively in the image domain to locate the boundaries of the object of interest and involves minimizing an energy function. While it was effective, the work only used qualitative metrics and lacked an understanding of whether the proposed methodology would generalize well beyond the data it was trained on (coastline of Greece).

There has also been work [9] in using traditional machine learning algorithms Random Forest, XGBoost and

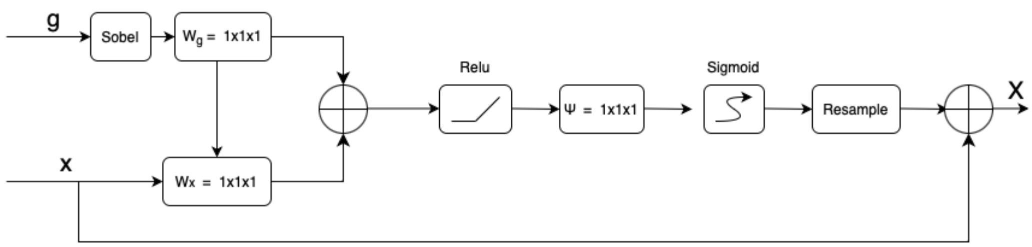

Figure 1. Proposed Edge Attention Module

LGBM for shoreline recognition. However, these studies were not very intensive and did not generalize very well.

The idea of our work was inspired by Seale et al. [7]. The authors explored novel convolutional neural networks based on the U-Net architecture on the Sentinel-2 Water Edges Dataset (SWED).

2.2. Variants of U-Net

When we performed a quick literature review of Image segmentation models, either as variation of the U-Net for binary segmentation or for land/sea segmentation specifically, we came across the following paper related to our project. The DeepUNet [2] was one of the first semantic segmentation models that was used to classify land/sea from images. HED-UNet: [1] tried predicting edges and segmentation results in parallel. Xiao et al [8] provided an model that uses data injection into layers while using attention that acts as the main inspiration to develop methods to inject in edge information to improve shoreline detection. We began by reading [3] for new methods for combining edge information but later realised the need to build a more deep learning friendly architecture that is easy to scale and train on large datasets.

With the rise of Attention as a powerful feature correlation method and multiple papers using the Attention U-Net [4], we wished to build a new U-Net with an attention module that could be more task specific (For binary land/sea segmentation) than using a generic segmentation model that performs better in traditional multi-class classification, to our use case.

3. Methods

Since we know that the specific problem statement we are trying to tackle is binary segmentation (land/sea), we wish to leverage this to build a model that is tailored for this best. Analysing this data from a human perspective, it is a little obvious that we humans don't always annotate/classify each pixel or patches of pixels directly as land or sea, and rather try to draw the boundaries to segregate the observable space and then decide if the clusters formed are land or sea.

We propose a conjecture based on this behaviour that for this specific problem, using the edge information somehow, along side the original image data to be able to segment the image into land and sea, might be fruitful.

The Edge Attention Gate Module above As the feature-map grid in a U-Net is gradually downsampled, features on the coarse spatial grid level model location and relationship between tissues at global scale. These maps only get more broader and keeps missing on finite details. The traditional attention module proposed, helps in bringing in more local contexts from previous layers. We propose a sobel edge detection module addition to the attention gate. We believe that the edge features are good enough (probably more useful than other information propagation) and might help the model converge better and produce more accurate predictions.

3.1. Edge-Attention Module

We propose a new edge-based attention module (Figure 1) that produces maps by taking into consideration the edges in the feature map rather than just the feature itself. We take the attention approach proposed in [4] a step further by adding a Sobel Filter on the first feature set to gate that, allowing attention coefficients to take advantage of regions formed by edges. We believe this enables the model to use regions in the image more efficiently.

3.2. Edge-Attention U-Net Model specification

Now that we have an edge-attention module that uses the edge data for finding correlations between segments of the image for segmentation, we go ahead adopting a U-Net architecture similar to the traditional Attention U-Net [4]. Our proposed network (Edge Attention U-Net 2) uses edges of features at different levels.

While using edge features as gates for the attention module in the second half of the U-net, we use Sobel module on to obtain edges from the features in the first half as our main data gets compressed. Borrowing edges from here provides us with the advantage of having different scales and depths of the edges provided. Since we know that the main features of the image keep getting compressed and the scope of features keep going wide with layers, applying sobel on these give us edges on different scales as well. Earlier edges give

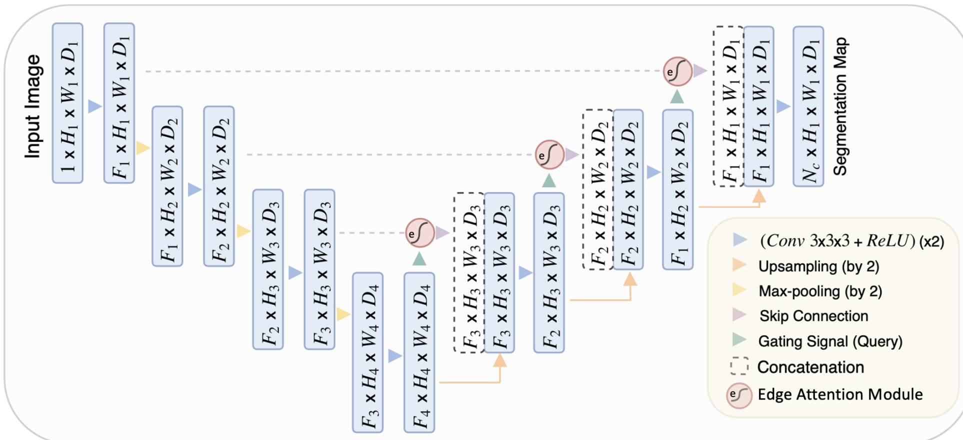

Figure 2. Edge-Attention U-Net model architecture (inspired from Attention U-Net [4])

us minute and finer edge details that will help the model handle small patches, while deeper in the architecture, the edges become more broad and prominent.

4. Experiments

To evaluate our model, we chose the Traditional U-Net [6] and the Attention U-Net [4]. The U-Net has a total of 50 convolutional layers, 4 up-convolutional layers, and 1 1x1 convolutional layer. Our model has the same configuration but has extra additional attentions from first half to second. The attention U-net also has the same configuration but has attention model from features in the first half. We conducted our experiments on an Alienware m17 laptop with the following specifications: an AMD Ryzen 9 - processor, 16GB of DDR4 RAM, a NVIDIA GeForce RTX 3070-Ti graphics card, and a 512GB PCIe M.2 SSD. We used Python 3.8.5 as our programming language and PyTorch 1.7.1 as our deep learning framework. We also used CUDA 11.0 and cuDNN 8.0 for GPU acceleration. The laptop was running on Windows 11 operating system.

We used the SWED_sample dataset that has 1764 images and labels. We split it into Train, Validation and Test datasets with 1411, 177 and 176 data points. We trained each of the models for 30 epochs with batch size of 8. Post training, we evaluated these models on the test set to compare the models based on the standard pre-defined metrics.

5. Results

5.1. Qualitative Analysis

The output of all 3 models was plotted against the ground truth and the image [Figure 3]. Qualitative analysis revealed that our model was performing quite well when compared with the true labels of unseen data. It was able to capture minute details in most predictions and did not have major detail issues.

We also notice similarities between the capabilities of the Attention U-Net and Edge Attention U-Net in detecting water pixels [Figure 3 (b)]. They both, however, were clearly superior to U-Net. U-Net fails in areas where there are very small patches of desired regions surrounded by the other. This is especially visible in Figure 3(a).

Additionally, in specific scenarios [Figure 3 (c) and (d)] where either land or water pixels are larger in number, Attention U-Net and Edge Attention U-Net adequately identify the difference and perform well.

[Figure 3 (e)] displays an example of a situation where Attention U-Net hallucinates and generates water pixels while the other models don't. It is likely that using edge features helped the model realise that these small patches aren't truly different classes.

5.2. Quantitative Analysis

Quantitative evaluation metrics used are as follows:

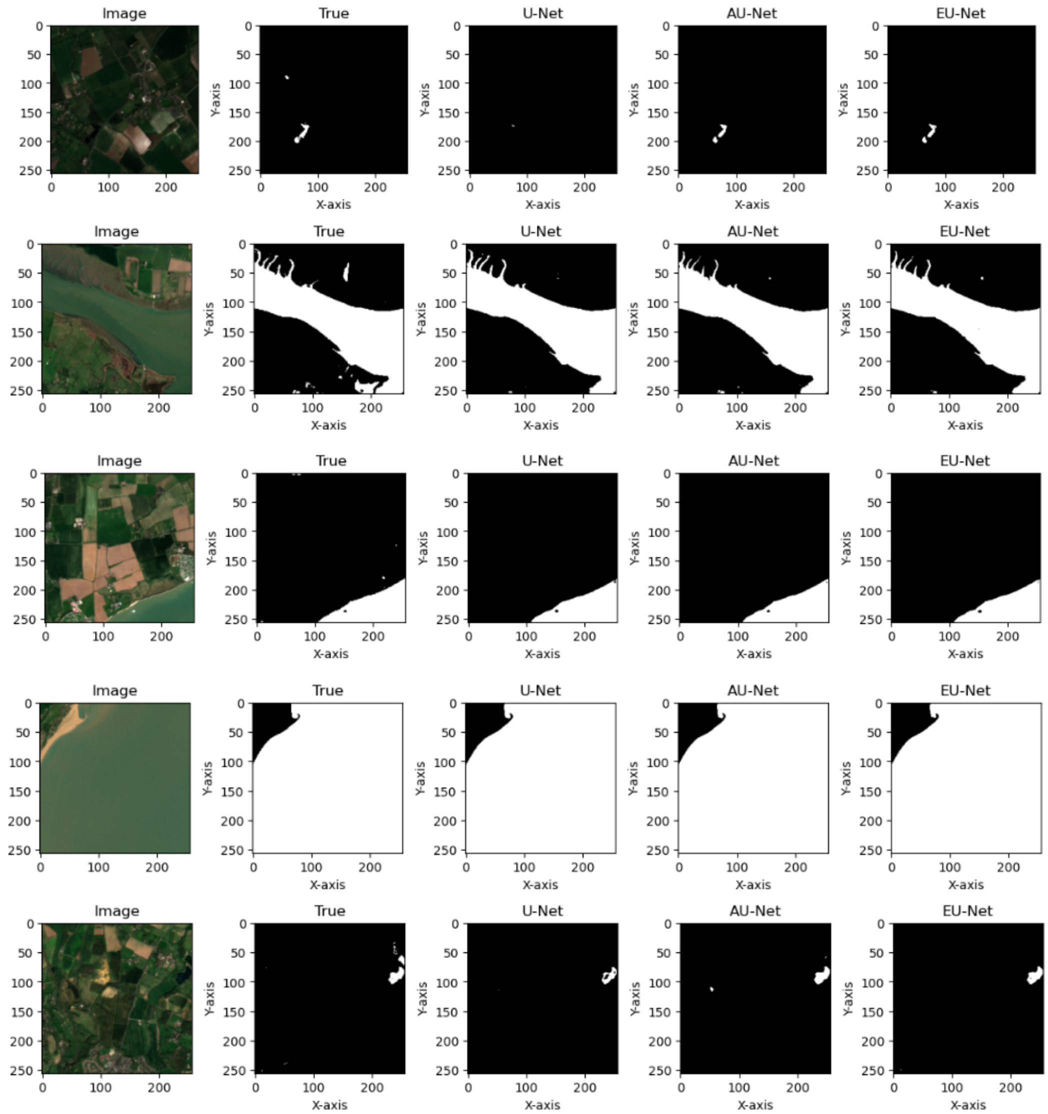

Figure 3. (a),(b),(c),(d) & (e) are 5 SWED-2 Images from the test set compared with the true labels and predictions of all 3 models.

$$\text{Sensitivity or Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$
$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}$$
$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$
$$\text{F1-Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$
$$\text{Jaccard Similarity} = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}}$$
$$\text{Dice Coefficient} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}}$$

where:

- TP - True Positive or correctly predicted water pixels;
- TN - True Negative or correctly predicted land pixels;
- FP - False Positive or land pixels predicted to be water;
- FN - False Negative or water pixels predicted to be land.

From 1, we observe that our model performs better, if not, as good as, other models in all performance metrics. Despite Attention U-Net performing as well as our model in the Qualitative study, we can see here that when these models are deployed over all images in the test set, our model outperforms Attention U-Net comfortably.

Jaccard Similarity and Dice coefficient, the most widely used metrics for semantic segmentation [6], are highest for our model. This signals a slightly better pixel classification over Attention U-Net.

As it stands, the model performs well on unseen images as evident from the evaluation metrics. From Table 1 we can see the increase in the metric values moving from U-Net to Attention U-Net to Edge Attention U-Net. With some hyperparamter tuning, Edge Attention U-Net outperforms the baseline U-Net and the theorized upper limit of Attention U-Net.

	U-Net	Atten U-Net	Our Model
Sensitivity	0.8886	0.9160	0.9123
Specificity	0.9993	0.9977	0.9987
Precision	0.9888	0.9537	0.9788
F1 Score	0.9186	0.9293	0.9358
Jaccard Similarity	0.8822	0.8957	0.9004
Dice Coefficient	0.9186	0.9293	0.9358

Table 1. Comparison of model performances

6. Conclusion & Future Work

Our proposed model, that uses edges detected from various levels of feature abstraction in a typical U-Net as part of the segmentation task, is able to perform better than traditional U-Net. We plan on improving our results for land-sea segmentation by building methods that could be more robust.

Since the model has currently been trained an a subset of the SWED dataset due to memory restrictions, the first step would be expanding to using the entire SWED dataset.

Most satellite imagery consist of 12 bands or channels. While our project handles this in the preprocessing stage to work with the RGB bands, our model currently does not generalize to consider older data that consists of just 1 band. The next step would be to make the model applicable to a wider range of images.

We also want to extend to multi-class segmentation and panoptic segmentation to be able to identify differing land and water bodies. This can especially be useful in images where water (or a river) divides two land masses or vice versa.

7. Contributions

The project team consists of Vihaan Akshaay and Aman Sariya. Vihaan's focus through the project was mainly on studying and understanding the dataset along with conceptualizing the flow of the project with respect to the neural network architectures that need to be identified and implemented for the purpose of what we want to achieve.

Aman's focus was towards finding the right segmentation model (i.e. semantic segmentation) and studying it's implementation in the context of the dataset that we have. He also focused on preprocessing data along with some work on the evaluation metrics.

References

[1] Konrad Heidler, Lichao Mou, Celia Baumhoer, Andreas Dietz, and Xiao Xiang Zhu. HED-UNet: Combined segmentation and edge detection for monitoring the antarctic coastline. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–14, 2022. 2

[2] Ruirui Li, Wenjie Liu, Lei Yang, Shihao Sun, Wei Hu, Fan Zhang, and Wei Li. Deepunet: A deep fully convolutional network for pixel-level sea-land segmentation, 2017. 2

[3] X Muñoz, J Freixenet, X Cufí, and J Martí. Strategies for image segmentation combining region and boundary information. *Pattern Recognition Letters*, 24(1-3):375–392, 2003. 2

[4] Ozan Oktay, Jo Schlemper, Loic Le Folgoc, Matthew Lee, Mattias Heinrich, Kazunari Misawa, Kensaku Mori, Steven McDonagh, Nils Y Hammerla, Bernhard Kainz, Ben Glocker, and Daniel Rueckert. Attention u-net: Learning where to look for the pancreas, 2018. 2, 3

[5] Vasilis Paravolidakis, Lemonia Ragia, Konstantia Moirogiorgou, and Michalis E. Zervakis. Automatic coastline extraction using edge detection and optimization procedures. *Geosciences*, 8(11), 2018. 1

[6] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation, 2015. 3, 5

[7] Catherine Seale, Thomas Redfern, Paul Chatfield, Chunbo Luo, and Kari Dempsey. Coastline detection in satellite imagery: A deep learning approach on new benchmark data. *Remote Sensing of Environment*, 278:113044, 2022. 2

[8] Jiajun Xiao, Jie Li, Qiangqiang Yuan, and Liangpei Zhang. A dual-unet with multistage details injection for hyperspectral image fusion. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–13, 2022. 2

[9] Yunshun Zhong and Tamer El-Diraby. Shoreline recognition using machine learning techniques. *IOP Conference Series: Earth and Environmental Science*, 1101(2):022025, nov 2022. 1