

A comparison between image processing models featuring

Human-drawn, Segment Anything (SAM), and Mask R CNN

Hoan Lam and Dat Nguyen

Abstract— Instead of spending large amounts of time identifying corals manually, researchers can now use image processing tools to automatically detect corals in images. Tools such as Segment Anything (SAM) and Mask R-CNN are capable of quickly identifying objects in images. Comparing these tools is essential to help researchers decide which is more suitable for their needs. In this paper, users input an image along with a hand-drawn outline of an object they want the tool to detect. The image is processed using SAM and Mask R-CNN, and the detected objects are compared with the hand-drawn outline to find the most overlapping mask. An Intersection over Union (IoU) score is generated for each model based on how well it overlaps with the hand-drawn mask. Another IoU score is calculated to compare the best masks from each model for validation. Both scores are plotted to show side-by-side trends. Surprisingly, results showed that Mask R-CNN occasionally outperformed SAM in certain cases, particularly when outlines were loosely drawn or when the coral was large and visually distinct. SAM generally performed better with detailed outlines but was sometimes over-segmented. These findings suggest that each tool has strengths and limitations depending on the user's drawing style and the coral's visual complexity, making them better suited for teaming up rather than direct comparison.

1 INTRODUCTION

This project compares the performance of the Segment Anything Model (SAM) and Mask R-CNN in detecting coral outlines. A web page is set up where users can input a coral image and a hand-drawn outline. The program then runs both models and compares their detected masks with the hand-drawn outline using the Intersection over Union (IoU) equation. This study aims to provide helpful insight into which tool suits marine biology researchers. Tasks, Insights, and Research Questions:

- What is the IOU of SAM vs Human Drawn Outline?
- What is the IOU of Mask R-CNN vs Human Drawn Outline?
- What is the IOU of Mask R-CNN vs SAM Outline?
- What is better for finding outlines automatically?

Notes: IOU is intersection over union, a metric to measure how much two areas are aligned

2 BACKGROUND

Tracking coral populations on a reef requires researchers to identify corals in the area accurately. Traditionally, this is done manually—researchers outline each coral by hand, introducing cost, time, and human error. Given the ever-changing environment, comparing old and new coral layouts to identify new growth or coral loss can be tedious. While the manual method works, the advancement of technology now makes it possible to look for better alternatives. This project explores two such tools: SAM and Mask R-CNN.

2.1 SAM

Meta AI developed the Segment Anything Model, a general-purpose image segmentation tool suitable for detecting objects of any kind—including corals, in this case. Its capabilities come from the largest segmentation dataset ever created, SA-1B, which includes over 1.1 billion segmentation masks across more than 11 million images from various categories [1]. While SAM does not assign labels to objects, its ability to generate accurate masks makes it well-suited for coral detection.

2.2 Mask R-CNN

Mask R-CNN is a deep learning framework for instance segmentation, also developed by Facebook AI Research. Unlike SAM, Mask R-CNN must be trained on a labeled dataset with predefined object categories [2]. This project uses a pre-trained Mask

R-CNN model trained on the Microsoft Common Objects in Context (MS COCO) dataset, which contains approximately 330,000 images across 80 object categories. Since corals are not included in COCO, Mask R-CNN cannot label corals. However, it can still generate masks around visually distinct coral outlines, although the labels may be incorrect.

3 IMPLEMENTATION

3.1 Front-end

Users upload a coral image on the webpage, which is then displayed for manual outlining. Drawing instructions and Undo and Redo buttons are provided to support editing.

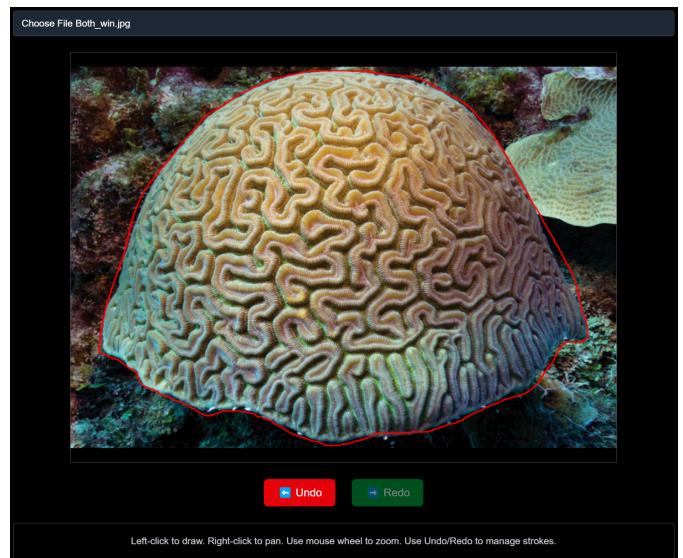


Fig. 1. User input area

After outlining the coral, the user clicks the “Send data to Flask” button to begin processing. A status message lets the user know when the program is finished. Once complete, the user can click the “Get data from Flask” button to retrieve the results.

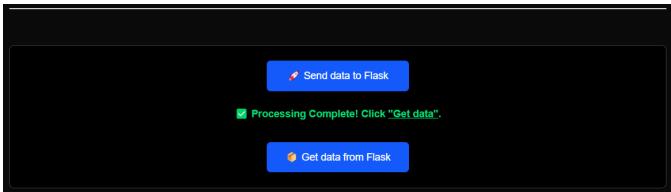


Fig. 2. Send and fetch

3.2 Back-end

Behind the scenes, Python analyzes the data received from the front end. The uploaded image and the corresponding hand-drawn outline are saved into a local folder. New segmented images are generated based on the original input using the SAM and Mask R-CNN models, called the model interpretation images (Figure 3). These images represent how each model interprets the coral outline.

The hand-drawn mask is compared with each segment produced by SAM and Mask R-CNN. The most overlapping segment (based on Intersection over Union, or IoU) is selected from each model. These selected masks are saved and sent back to the front end for visual comparison, showing how well each model matches the human-drawn outline (Fig. 3)

IoU scores are also saved to a .txt file for later use in generating a trend chart, which helps visualize each model's performance across multiple images.

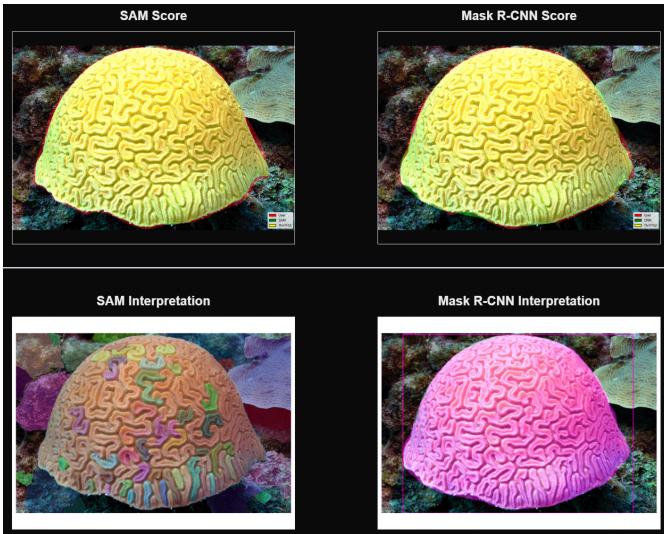


Fig. 3. Sam vs Mask R-CNN visualizations

4 RESULTS & EVALUATION

The results vary based on image size and how the user outlines the coral. If the user draws a detailed outline, SAM typically scores higher. SAM's parameters can also be adjusted to look at larger regions, filtering out the finer details. However, this depends on the individual user and the shape of the coral—some coral structures are irregular. In general, SAM is more capable and flexible than Mask R-CNN when dealing with various outlines and coral types.

On the other hand, Mask R-CNN performs better when the user is less focused on precision—for example, when they just want to circle a coral and move on. It also does relatively well when the coral is large and visually prominent, even though it was not trained on coral-specific datasets. This makes it a valuable tool for quick segmentation tasks with minimal user effort.

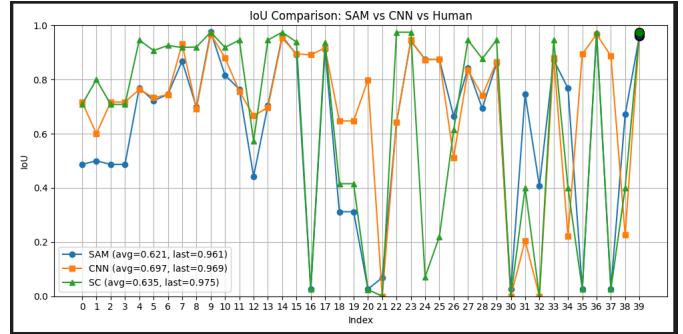


Fig. 4. Trend graph

Figure 4 shows the trend of model performance across different test cases. In the early tests, Mask R-CNN sometimes outperformed SAM because SAM focused too much on fine details, leading to lower overlap with loosely drawn outlines. However, in more challenging cases, such as tests with random outlines, SAM outperformed Mask R-CNN. In some of these tests, Mask R-CNN failed to detect any relevant segment, while SAM could still return a partial mask in the targeted region. In figure 4, SAM is indicated in blue, Mask R CNN is indicated in orange, and SAM vs Mask R CNN (SC) score is indicated in green. Using the image and human outline from Figure 1, the results for the task are as follows: SAM scored 0.961, while Mask R CNN scored 0.969, and SC scored 0.975. So, Mask R CNN outperformed SAM for this example, the SC score also indicates that the regions SAM and Mask R-CNN generated do have much in common, area-wise.

5 CONCLUSIONS

SAM is good at segmenting objects, and depending on the specific goal the user has in mind, it can do a great job identifying coral outlines. Configuring SAM properly is important to get the most accurate results. In general, users can expect SAM to perform well across a variety of coral types, but a quick check of the results is still recommended—SAM is powerful but not flawless.

Mask R-CNN, on the other hand, is also effective at detecting coral structures. It processes images much faster than SAM, mainly because it uses the model with less weight and generates fewer masks overall. With the right training dataset focused on coral detection, Mask R-CNN could outperform SAM in speed and accuracy. Its ability to be specialized for a chosen subject makes it a strong candidate for coral classification and labeling.

By comparing both models, we could highlight SAM and Mask R-CNN's strengths and limitations. Looking back, it's clear that SAM and Mask R-CNN are not competing tools for image segmentation but rather complementary ones. They can offset each other's weaknesses and enhance the overall coral detection workflow.

REFERENCES

- [1] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, W.-Y. Lo, P. Dollár, and R. Girshick, “Segment Anything,” *arXiv preprint arXiv:2304.02643*, Apr. 2023. [Online]. Available: <https://segment-anything.com>
- [2] K. He, G. Gkioxari, P. Dollár, and R. Girshick, “Mask R-CNN,” *arXiv preprint arXiv:1703.06870*, Mar. 2017. [Online]. Available: <https://arxiv.org/abs/1703.06870>