# Automatic Image-Based Coral Polyp Analysis through Multi-View Instance Segmentation

S. Dutta[1], G. Pavoni[1], M. Corsini[1], F.Ganovelli[1], P.Cignoni[1], P. Rossi[2], E. Cenni[2], R.Simonini[2], F.Grassi[2], D.Cassanelli[2], S. Cattini[2], L.Rovati[2], A.Capra[2], and C.Castagnetti[2]
[1]CNR-ISTI, Pisa, Italy, [2]University of Modena and Reggio Emilia, Modena, Italy

## ABSTRACT

We present an automated framework for counting and measuring the polyps of Cladocora caespitosa, a Mediterranean reef-building coral. To our knowledge, the most practical method for counting polyps currently involves ecologists' visual inspection of a 3D model. However, measuring polyps from the model can lead to inaccuracies due to distortions in the reconstruction. Our method integrates deep learning-based instance segmentation on 2D images with 3D models for unique polyp identification, ensuring precise biometric extraction. The proposed pipeline automates polyp detection, counting, and measurement while overcoming the limitations of manual in situ methods. Laboratory validation demonstrates its accuracy and efficiency, paving the way for scalable, high-resolution phenotyping, and field monitoring of Mediterranean coral populations.

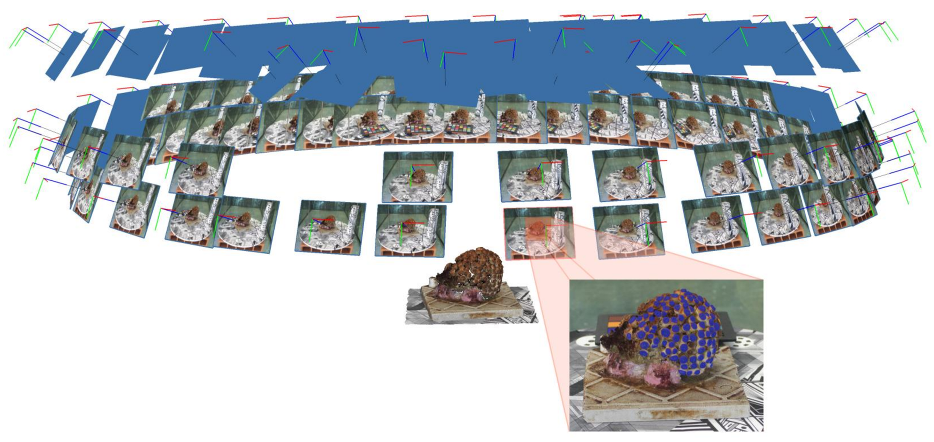

Figure 1: Pipeline of multi-view instance segmentation.

## OVERVIEW

Ecological monitoring is vital for evaluating underwater ecosystem health, but traditional methods for assessing fine-scale structures like coral colonies are slow and imprecise. While underwater photogrammetry is common for tropical corals, Mediterranean species like Cladocora caespitosa remain underexplored. Counting its densely packed, repetitive polyps on 3D models is time-consuming and inaccurate due to their transparent, dynamic nature, which hampers geometric reconstruction. To address this, we propose a hybrid method that uses 2D image-based phenotyping for polyp measurement, leveraging 3D models only for unique polyp identification via multi-view redundancy. Tests in controlled tank environments show that our approach automates fine-scale biometric extraction effectively, offering a foundation for scalable Cladocora monitoring in the field.

## METHODOLOGY

In Cladocora cespitosa, key biometric indicators for assessing the species' health status include the number of living polyps, the oral disc length (LP), the oral disc width (WP), and the oral disc surface area of each corallite. Determining these metrics would typically require segmenting each polyp from a 3D model.

The developed pipeline involves automatic instance segmentation of polyps from raw images, followed by Structure-from-Motion (SfM) reconstruction of a 3D model. The segmentation is then transferred from the images to the 3D model to ensure unique counting, and finally, biometric quantities are estimated based on a selected view of each polyp.

To our knowledge, no deep learning-based solution currently exists for automatic coral polyp instance segmentation. In [LBK∗21], an SVM classifier was developed for semantic segmentation of living corals in multispectral lab images. Due to the lack of annotated datasets, we created a dedicated dataset for training an instance segmentation network to detect corallite oral discs, using the interactive annotation tool TagLab [PCP∗22]. TagLab, enhanced with the agnostic segmentation model SAM [KMR∗23], streamlines semi-automatic labeling and dataset management. We extended it with an exporter compatible with YOLO11 [JQC23], our chosen segmentation model. YOLO11-nano, a lightweight architecture designed for real-time performance, enabled fast segmentation of polyp discs from 3D reconstruction image sets. Trained on ~8,000 oral disc instances, the network achieved 0.902 mAP@50 and 0.593 mAP@50-95 on a test set of 2,834 instances.

As discussed earlier, the most effective approach for segmenting polyps is to use the acquired images. However, since each polyp appears in multiple images, we employ image-based 3D reconstruction using Structure-from-Motion (SfM) as a common reference. SfM estimates camera parameters, enabling us to reproject each segmented pixel region onto the corresponding part of the 3D model. We associate each segmented polyp with its covered mesh triangles and pixel count per triangle to estimate overlap. When regions overlap in 3D, we infer they belong to the same polyp, preventing multiple counts (see Figure 2). A second advantage of 3D reconstruction is identifying the most frontal image of each polyp, ensuring accurate segmentation. This image and its label are then used to extract reliable biometric measurements, as shown in Figure 3 using TagLab.

## RESULTS

For our preliminary studies, we tested our measuring methodology on colonies of Cladocora caespitosa transplanted in water tanks. Colonies were photographed using a standard photogrammetric acquisition procedure, and the models were reconstructed and scaled with Metashape software. Polyps were manually counted by marine ecologists directly into Metashape to provide us with ground truth.

In the current version of the automatic counting tool, the number of polyps detected automatically aligns with the manually counted ones with an average accuracy of 96.4%. For instance, in the example shown in Figure 2, our system counts 196 polyps, compared to the ground truth of 191. This discrepancy is primarily due to the approximations in the 3D model, which lead to improper surface coverage, particularly in areas near the silhouette of the projection. Considering that a small margin of error is acceptable, as some dead polyps are challenging to distinguish even for experts. This greatly reduces the required working time—from 20–30 hours per model to just a few minutes with our automatic counting tool. In addition, it automatically provides the width and height of each polyp's oral section, further saving time. These measurements are taken from the most frontal view, avoiding the need for experts to manually select suitable images or spend time exploring the 3D model to extract measurements.

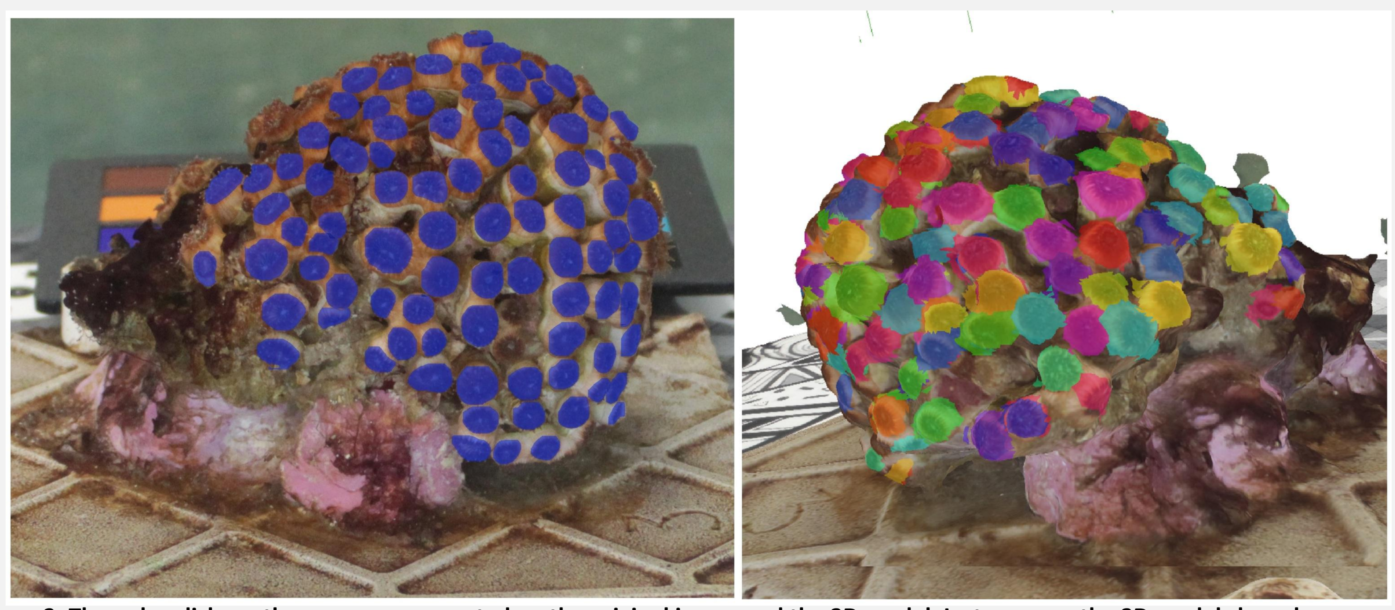

Figure 2: The polyp disks as they appear segmented on the original image and the 3D model. Instances on the 3D models have been segmented from multiple views.

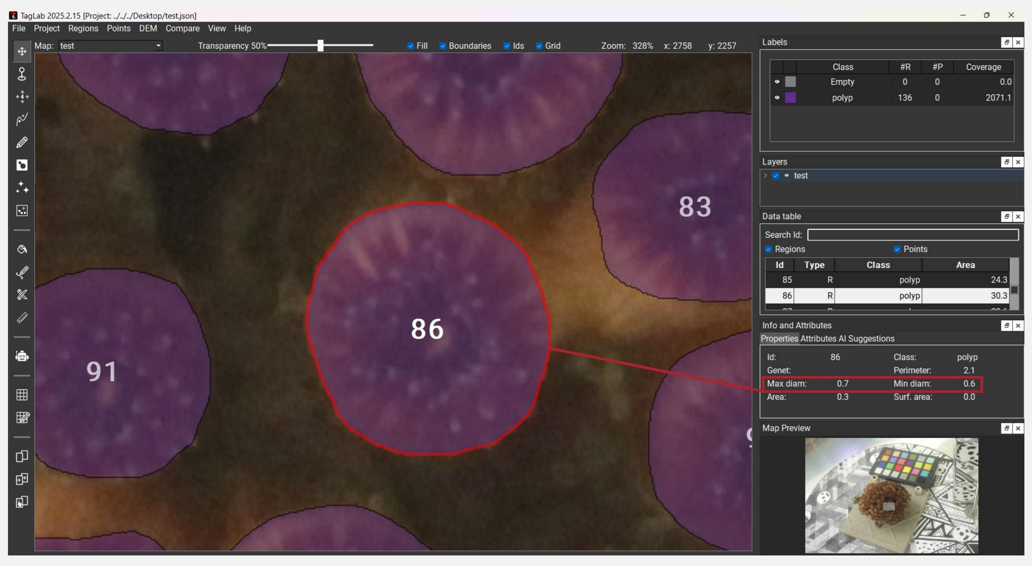


Figure 3: Metrics are extracted from the best view of a disc and dis-played within TagLab. The maximum diameter measurement rep-resents the oral disc length (LP), while the oral disc width (WP)corresponds to the minimum diameter measurement. The disk area is indicated on the line below. All measurements are given in centimeters.

## AFFILIATIONS

EUPHROSYNE HIGH-RESOLUTION CORAL MONITORING

UNIMORE UNIVERSITÀ DEGLI STUDI DI MODENA E REGGIO EMILIA

Consiglio Nazionale delle Ricerche

ISTI

## REFERENCES

[JQC23] JOCHER G., QIU J., CHAURASIA A.: Ultralytics YOLO, Jan. 2023. URL: https://github.com/ultralytics/ultralytics.

[KMR*23] KIRILLOV A., MINTUN E., RAVI N., MAO H., ROLLANDC., GUSTAFSON L., XIAO T., WHITEHEAD S., BERG A. C., LO W.-Y., DOLLÁR P., GIRSHICK R.: Segment anything. arXiv:2304.02643(2023). 1

[LBK*21] LIU H., BÜSCHER J. V., KÖSER K., GREINERT J., SONG H.,CHEN Y., SCHOENING T.: Automated activity estimation of the cold-water coral lophelia pertusa by multispectral imaging and computational pixel classification. Journal of Atmospheric and Oceanic Technology 38,2 (2021), 141–154. 1

[PCP*22] PAVONI G., CORSINI M., PONCHIO F., MUNTONI A., ED-WARDS C., PEDERSEN N., SANDIN S., CIGNONI P.: Taglab: Ai-assisted annotation for the fast and accurate semantic segmentation of coral reef ortho images. Journal of Field Robotics 39, 3 (2022), 246–262.doi:https://doi.org/10.1002/rob.22049. 1