



Archaeocyathids: Precursor to Corals

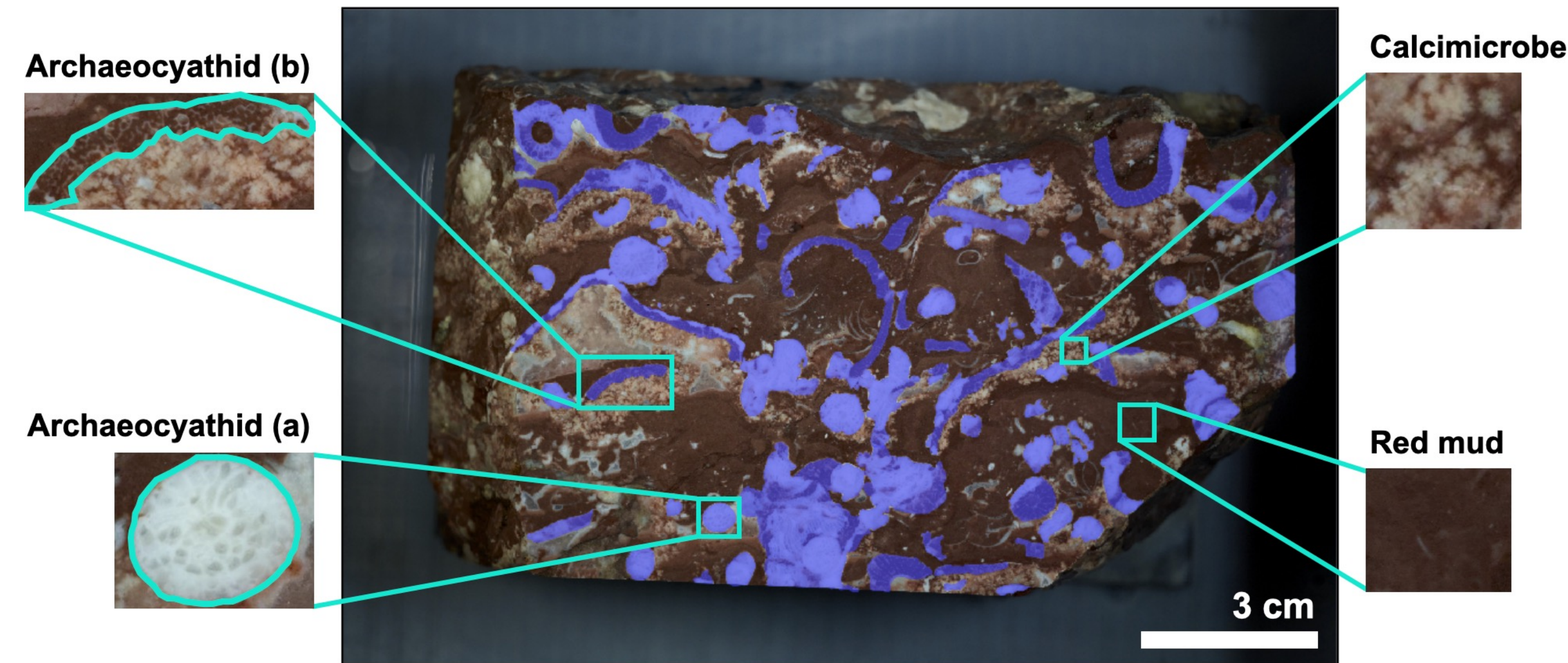


Fig. 1: Overview of rock sample¹.

Motivation & Goal

Why Archaeocyathids: Studying archaeocyathid reef structure will inform our understanding of coral reef impact on Earth's biosphere¹.

Goal: Reconstruct individual archaeocyathids from 2D images (Fig. 2).

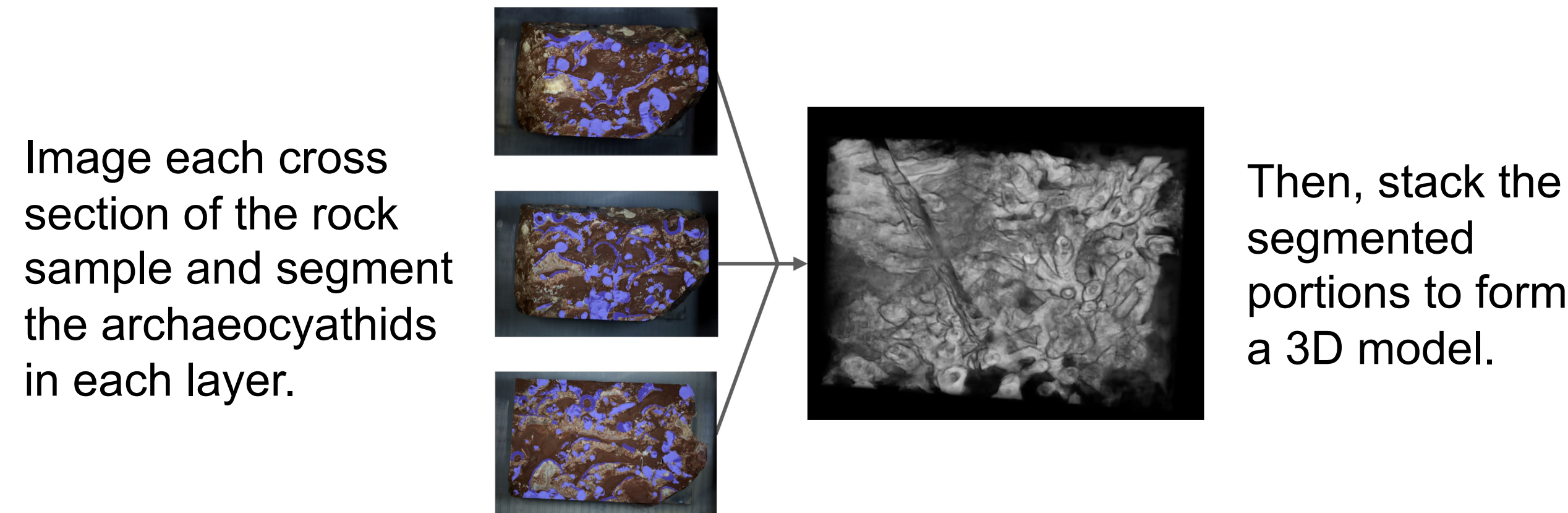


Fig. 2: 3D reconstruction of archaeocyathids from 2D segmentation of rock layers.

Problem: Need to segment over 3,000 images and have a **limited amount of labeled data** (10 annotated images) with which to train a segmentation model.

References & Acknowledgments

[1] R. Manzuk et al. "Branching archaeocyathids as ecosystem engineers during the Cambrian radiation". *Geobiology*, 2022.

[2] Zeiler and Fergus. "Visualizing and understanding convolutional networks". ECCV 2014.

[3] S. Santurkar, D. Tsipras, et al. "Editing a classifier by rewriting its prediction rules". NEURIPS 2021.

We acknowledge support from the Princeton SEAS IW funding (IP), the Princeton SEAS Project X Fund (RF), and Open Philanthropy (RF). We thank Devon Ulrich and Sunnie S. Y. Kim for helpful discussions.

Identifying Model weaknesses

By applying domain-informed image perturbations, we find that our model



Fig. 3: Domain-informed image perturbations include occlusion².

1. Experiences interclass confusion

- Inpaint archaeocyathids with red mud shade → Identifies calcimicrobes (Fig. 1) as archaeocyathids when nothing should be identified (Fig. 4).
- Inpaint non-archaeocyathid pixels with red mud shade → Mean IoU increases from 0.63 ± 0.29 to 0.78 ± 0.24 , model can distinguish between red mud and archaeocyathids (Fig. 5).

2. Is not robust to intraclass variation

- Substitute porous texture (archaeocyathid a in Fig. 1) for all archaeocyathids → Mean IoU increases from 0.63 ± 0.29 to 0.66 ± 0.31 (Fig. 6).



Fig. 4: Circled are two of the calcimicrobes misclassified as archaeocyathids.

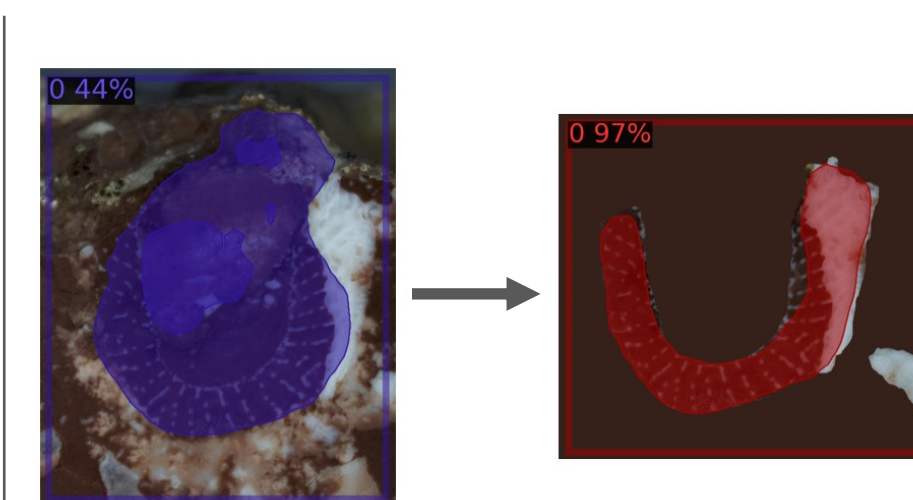


Fig. 5: Archaeocyathid segmentation improves when non-archaeocyathids are inpainted.

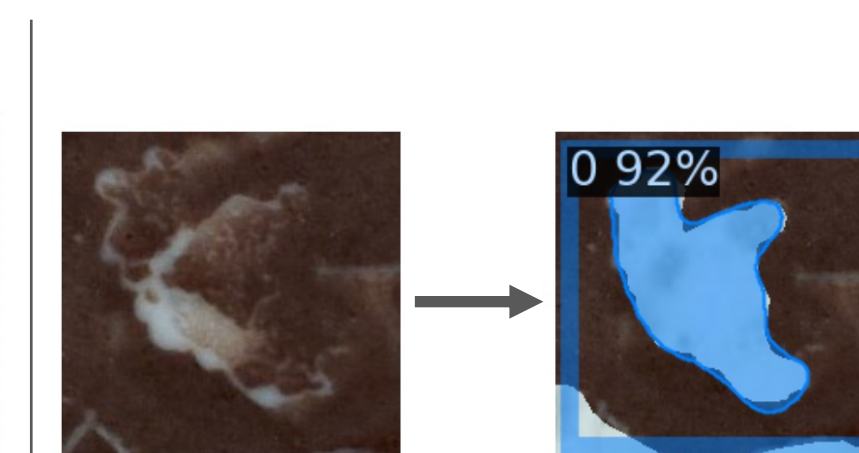


Fig. 6: Archaeocyathid segmentation improves when replaced with the porous texture.

Model Editing as a Data-Efficient Solution

We extend a model editing method for image classifiers³ to image segmentation via domain-informed image perturbations.

- Data-efficient method for debiasing models, updating model knowledge, and correcting systematic mistakes.
- Encourage feature maps for two paired images to be similar (Fig. 7).

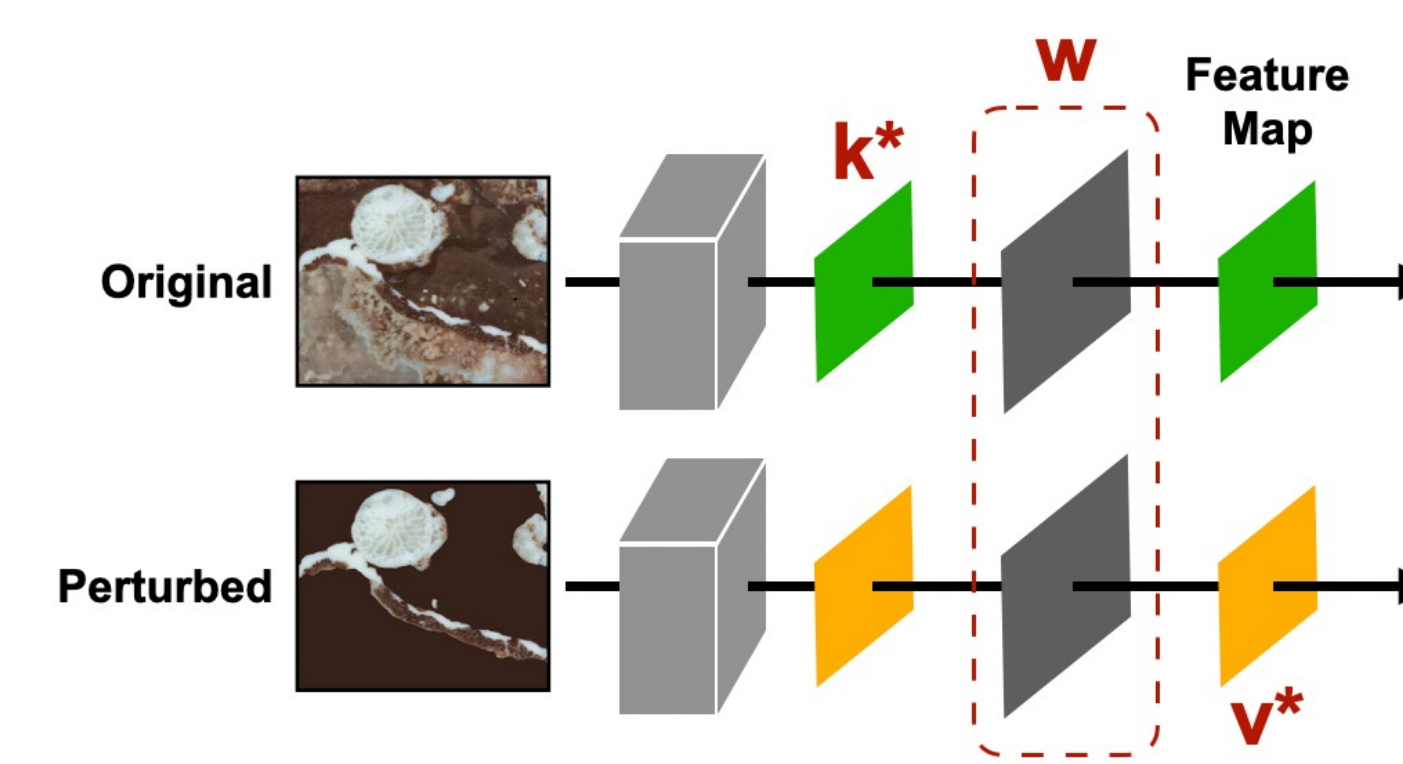


Fig. 7: Edit weights (w) such that inputs (k^*) map to specified outputs (v^*).

Results & Discussion

What is the best edit?

Image	Precision	Recall	IoU
None	0.86 ± 0.17	0.75 ± 0.26	0.63 ± 0.28
A	0.90 ± 0.13	0.59 ± 0.26	0.52 ± 0.27
B	0.88 ± 0.17	0.59 ± 0.25	0.52 ± 0.27
C	0.91 ± 0.12	0.63 ± 0.26	0.56 ± 0.27
D	0.89 ± 0.15	0.64 ± 0.25	0.56 ± 0.27
E	0.87 ± 0.17	0.64 ± 0.26	0.56 ± 0.28
F	0.90 ± 0.15	0.65 ± 0.26	0.57 ± 0.27
A	0.85 ± 0.16	0.70 ± 0.30	0.56 ± 0.32
B	0.84 ± 0.16	0.70 ± 0.29	0.55 ± 0.32
C	0.85 ± 0.18	0.73 ± 0.29	0.59 ± 0.31
D	0.84 ± 0.15	0.72 ± 0.28	0.55 ± 0.32
E	0.80 ± 0.20	0.69 ± 0.32	0.52 ± 0.34
F	0.85 ± 0.15	0.70 ± 0.30	0.57 ± 0.32

Mean across 215 instances.

A-F are the 6 training images.

Map non-archaeocyathids to red mud:

- Precision is the most important for our application and improves.
- Recall and IoU decrease.

Map archaeocyathids to porous texture:

- No improvement.
- Likely because less vivid visual contrast than between non-archaeocyathids and red mud.

How should we perform this edit?

- Both perturbations simultaneously and in sequence → Neither produces an improvement over the original model.
- Image sequences in many permutations (non-archaeocyathid to red mud mapping only) → Results match those under the last edit image alone (e.g., editing with sequence A,B,C is similar to editing with image C).
- Different amounts of perturbed pixels → **Perturbing more pixels** produces better performance (e.g., perturbing 1 to 35 to 100% of pixels increases precision from 0.86 to 0.90 to 0.91).

Key takeaways

- Identified model weaknesses** via domain-informed image perturbations.
- Extended an existing model-editing method** to image segmentation to address the identified weaknesses **with no additional labeled data**.
- Editing using all relevant pixels in one image (vs. using multiple images, multiple edits, or fewer pixels) is the **best setting for our application**.
- Future Work: Investigate what properties make an image better for editing.
- Limitations: Extensive experiments performed on Mask R-CNN trained on one rock sample. Novel domains should validate on their own models.