Cichlid Computer Vision Project – Weekly Progress

Week ending Friday, February 14th, 2025

Time Log

Charlie Clark

What progress did you make?

- Attended weekly BioBoost meeting on Monday evening.
- Watched the recording Bree sent about the BioBoost project on Tuesday. • Attended team huddle Thursday afternoon.
- Attended weekly HAAG admin meeting Thursday.
- Attended bi-weekly Bird Audio computational advisor meeting Friday morning.
- Attending bi-weekly Bird CV computational advisor meeting Friday afternoon.
- Ironed out Freeman meeting manager recording download issue.
    - Andrei organized Freeman lab faculty meeting to replace mine (to allow for direct recording downloads); he also organized his own Bird Audio comp advisor meeting to replace Bree's.
        - Deleted my series of Freeman faculty meetings, and asked Bree to delete her comp advisor meeting series for Bird Audio when she gets a chance (to avoid confusion).
        - Emailed faculty and comp advisor to notify them of the change (and clarified that nothing has changed with regard to meeting scheduling).
        - Need to find a replacement meeting manager for Bird CV, as Bina's GTRI Teams license doesn't allow her to create new meetings; reached out to her team mates asking for a volunteer (will assign the role randomly on Monday, if nobody volunteers).
    - Notified admin of the best approach for this issue, as to enable uniformity across teams and labs.
- Helped Eric troubleshoot/resolve his Dropbox access issues.
    - Engaged in a "two-front" conversation with Eric and Curtis from EBB support.
    - Made suggestions to Eric based on my understanding of how the GT Dropbox system works, as well as the information Curtis was telling me.
    - Eventually, we were able to resolve the issue: Eric now has access to the McGrath lab Dropbox file system.
- Continued literature review.

What are you planning on working on next?

- Attend required weekly meetings.
- Attend optional weekly meetings.
- Continue exploring/reviewing the BioBoost and CichlidBowerTracking repos.
- Meet with KQ immediately before weekly BioBoost meeting to go over temporal data with her.
- Attend publication seminar on Tuesday.

Is anything blocking you from getting work done?

- **None**

Researcher

What did you do this week?

(1) I attended/watched the following 3 meetings/recordings:

(a) Cichlid Team Meeting on February 10th. A brainstorm for the BioBoost paper rewrite was conducted. The meeting lasted 1 hr and 41 min.

(b) Bree's Recording on February 11th. I wrote up my notes and questions as a response. The recording was 41 min.

(c) Cichlid Team Working Meeting on February 13th. We discussed image segmentation techniques and identified a missing resource. The meeting lasted 50 min.

(2) I worked on the meeting manager role requirements for the weekly meeting. I posted the meeting recording link to Slack. AI meeting notes are not required due to us not having a faculty advisor. I updated the Teams attendance sheet at Projects / Cichlid-CV / ReID / For Microsoft Planner / Meeting Attendance Tracking.xlsx. I Added the week 6 meeting slides to that same Teams folder and to the Slack message. I will be adding YouTube videos for future meetings now that we know we can make it anonymous.

(3) Last week, I spent a while trying to fix the visualization code to represent the new SORT IoU changes. This week, I spent some more time on it and figured out that the new detections were only in 53 infrared videos. I created a side-by-side visualization of the new and old videos. Overall, even a large decrease in IoU values did not result in too many noticeable changes. If we were continuing with this dataset, I would definitely look more into this, since the results were a bit unexpected for me.

(4) I started looking into finding the file locations, scripts, and training/test splits of each part of the pipeline. As I went along, I created a Lucidspark chart to more easily track down scripts and data for BioBoost.

(5) Based on the project's new direction, I looked into Python packages for image segmentation. I was able to draw contours around fish videos from the Dropbox. I created

5 trials and finally found something that worked for finding fish contours. I created a rough draft to serve as a starting point for segmenting the fish themselves. I shared my code with Eric so that he could start working on the next step as well.
(6) I sent an email to my old advisor and Dr. Lytle to try to track down the email for CS 8903 conversion.

What are you going to do next week?
(1) I need to attend required meetings.
(2) I need to fulfill my meeting manager responsibilities.
(3) I need to finish looking into finding the file locations, scripts, and training/test splits of each part of the pipeline.
(4) I need to start working on segmenting the fish images or finding a way to identify between male and female fish visually.
(5) Other tasks as assigned after the weekly meeting on Monday.

Is anything blocking you from getting work done?
(1) I still need an email from Dr. Lytle for CS 8903. However, I emailed Dr. Lytle and cc'd Bree, as recommended by Bree, and I think she is talking to Dr. Lytle tomorrow. Hopefully it'll be good by next week. :)

Eric Iamarino

What did you do this week?

- Attended Working Meeting for BioBoost
- Watched Cichlid CV Weekly Meeting
- Watched Bree's meeting on BioBoost suggestions
- Spent time tuning KQ's OpenCV code to segment cichlid videos
- Tried using SAM python package to segment cichlid videos
- Fixed issues with Dropbox account
- Adding Week 5 & Week 6 Weekly Reports to website
- Adding Week 6 Weekly Meeting Update to website

What are you going to do next week?

- Attend BioBoost meeting
- Attend publication meeting
- Refine segmentation of cichlid from background tank
- Look for way to further segment and classify

Blockers, things you want to flag, problems, etc.

- None

Abstracts

Charlie Clark

"Geo-Aware Networks for Fine-Grained Recognition", Chu et al. (2019; arXiv pre-print)

- Abstract: "Fine-grained recognition distinguishes among categories with subtle visual differences. In order to differentiate between these challenging visual categories, it is helpful to leverage additional information. Geolocation is a rich source of additional information that can be used to improve fine-grained classification accuracy, but has been understudied. Our contributions to this field are twofold. First, to the best of our knowledge, this is the first paper which systematically examined various ways of incorporating geolocation information into fine-grained image classification through the use of geolocation priors, post-processing or feature modulation. Secondly, to overcome the situation where no fine-grained dataset has complete geolocation information, we release1 two fine-grained datasets with geolocation by providing complementary information to existing popular datasets - iNaturalist and YFCC100M. By leveraging geolocation information we improve top-1 accuracy in iNaturalist from 70.1% to 79.0% for a strong baseline image-only model. Comparing several models, we found that best performance was achieved by a post-processing model that consumed the output of the image-only baseline alongside geolocation. However, for a resource-constrained model (MobileNetV2), performance was better with a feature modulation model that trains jointly over pixels and geolocation: accuracy increased from 59.6% to 72.2%. Our work makes a strong case for incorporating geolocation information in fine-grained recognition models for both server and on-device."
- AI Summary by ChatGPT (4o): "This paper explores the use of geolocation as an auxiliary feature to improve fine-grained image classification, which differentiates between visually similar categories, such as animal or plant species. While previous work has shown that geolocation can help in coarse-grained classification (e.g., distinguishing bridges from mountains), this study systematically examines how raw latitude and longitude data can significantly enhance fine-grained recognition.

The authors propose three methods for integrating geolocation information into classification models:"

- 1. Geolocation Priors: Using Bayesian priors and label whitelisting based on geographic distributions.
- 2. Post-Processing Models: Combining an image classification model's output with a separate geolocation-based model.
- 3. Feature Modulation Models: Incorporating geolocation data directly into image feature learning.
    - "To support their research, they introduce two fine-grained datasets with geolocation data: an extended iNaturalist dataset with geolocation annotations, and a fine-grained evaluation dataset from YFCC100M, focusing on species-level classification."
    - Key Contributions
        - "First systematic study on incorporating raw geolocation data into finegrained image classification."
        - "New datasets with geolocation information, expanding the potential for future research in this area."
        - "Performance improvements: The post-processing approach improved classification accuracy from 70.1% to 79.0% on iNaturalist, while feature modulation boosted accuracy from 59.6% to 72.2% on MobileNetV2."
        - "Insights on model selection: The study found that post-processing models work best for larger networks, while feature modulation is more effective for smaller, resource-constrained models."
    - Future Research Directions:
        - "Generalization to other domains: While the study focused on species recognition, future work could explore other fine-grained tasks, such as vehicle or food classification."
        - "More sophisticated geolocation features: Investigating additional geolocation-based attributes beyond latitude and longitude, such as terrain, climate, or human population density."
        - "Better handling of long-tail distributions: Fine-grained datasets often have class imbalances, and future methods could address this issue more effectively."
        - "Real-world deployment: Testing the models in field applications, such as biodiversity monitoring or automated species identification in conservation efforts."

• Link: http://arxiv.org/abs/1906.01737

Researcher

Hulse, S. V., et al. “Using deep neural networks to model similarity between visual patterns: Application to fish sexual signals.” Ecological Informatics, 2022. https://www.sciencedirect.com/science/articl e/abs/pii/S1574954121002776.

This week, I was looking for papers in Ecological Informatics that addressed fish coloration, just to get an idea of what the current discussion of changing fish coloration was. The evolution of visual patterns in animals, especially in the context of sexual selection, helps us understand how these patterns function during courtship. Researchers applied models called sensory drive and sensory bias to see if the sexual signals of animals, like darter fish, match the visual characteristics of their environments. Using a deep learning tool called VGG19, they found that female darters’ patterns closely resemble their habitats, suggesting camouflage, while there was no clear evidence that male patterns were influenced by their surroundings, highlighting the potential of using advanced image analysis techniques in this area of research. To the researchers’ knowledge, this study was the first to use the Gram matrix to classify images. No future work is discussed by the authors.

Eric Iamarino

Ravi, N., Gabeur, V., Hu, Y.-T., Hu, R., Ryali, C., Ma, T., Khedr, H., Rädle, R., Rolland, C., Gustafson, L., Mintun, E., Pan, J., Alwala, K. V., Carion, N., Wu, C.-Y., Girshick, R., Dollár, P., & Feichtenhofer, C. (2024, October 28). Sam 2: Segment anything in images and videos. arXiv.org. https://doi.org/10.48550/arXiv.2408.00714

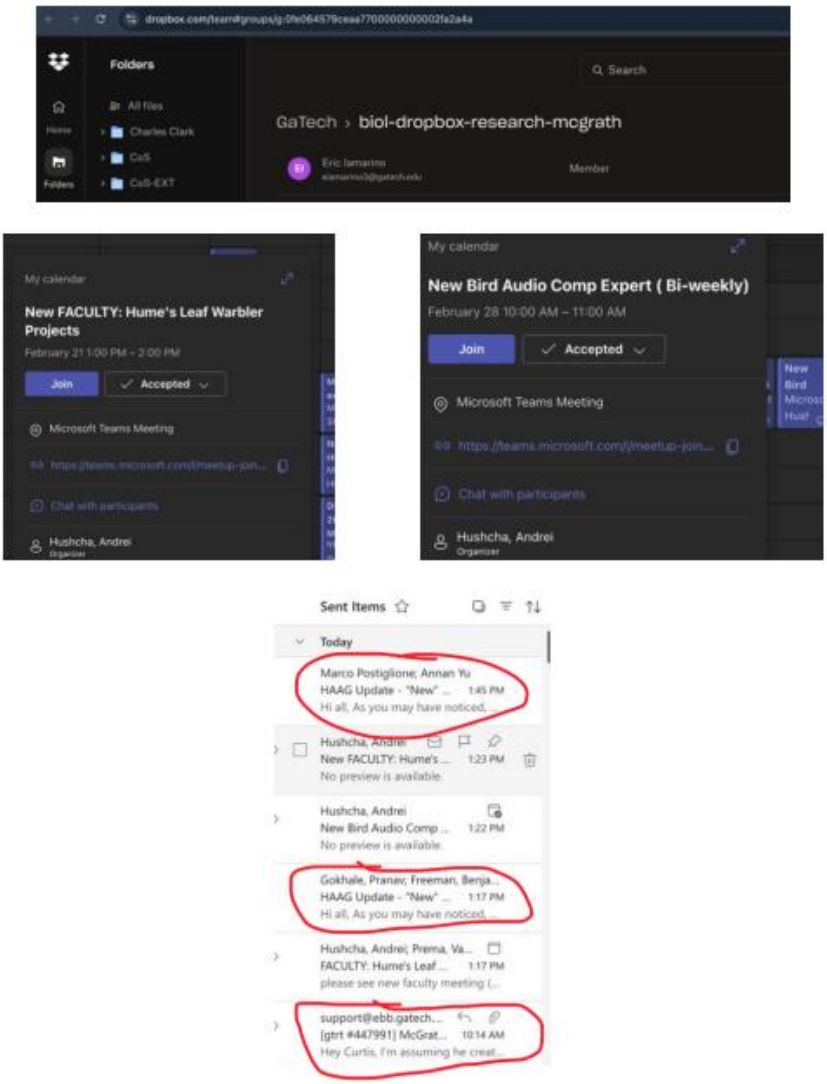
We present Segment Anything Model 2 (SAM 2), a foundation model towards solving promptable visual segmentation in images and videos. We build a data engine, which improves model and data via user interaction, to collect the largest video segmentation dataset to date. Our model is a simple transformer architecture with streaming memory for real-time video processing. SAM 2 trained on our data provides strong performance across a wide range of tasks. In video segmentation, we observe better accuracy, using 3× fewer interactions than prior approaches. In image segmentation, our model is more accurate and 6× faster than the Segment Anything Model (SAM). We believe that our data, model, and insights will serve as a significant milestone for video segmentation and related

perception tasks. We are releasing our main model, dataset, as well as code for model training and our demo

Documentation of Work

Charlie Clark

- Attended weekly BioBoost meeting on Monday evening.
- Watched the recording Bree sent about the BioBoost project on Tuesday.
- Attended team huddle Thursday afternoon.
- Attended weekly HAAG admin meeting Thursday.
- Attended bi-weekly Bird Audio computational advisor meeting Friday morning.
- Attending bi-weekly Bird CV computational advisor meeting Friday afternoon.
- Ironed out Freeman meeting manager recording download issue.
    - Andrei organized Freeman lab faculty meeting to replace mine (to allow for direct recording downloads); he also organized his own Bird Audio comp advisor meeting to replace Bree's.
        - Deleted my series of Freeman faculty meetings, and asked Bree to delete her comp advisor meeting series for Bird Audio when she gets a chance (to avoid confusion).
        - Emailed faculty and comp advisor to notify them of the change (and clarified that nothing has changed with regard to meeting scheduling).
        - Need to find a replacement meeting manager for Bird CV, as Bina's GTRI Teams license doesn't allow her to create new meetings; reached out to her team mates asking for a volunteer (will assign the role randomly on Monday, if nobody volunteers).
    - Notified admin of the best approach for this issue, as to enable uniformity across teams and labs.
- Helped Eric troubleshoot/resolve his Dropbox access issues.
    - Engaged in a "two-front" conversation with Eric and Curtis from EBB support.
    - Made suggestions to Eric based on my understanding of how the GT Dropbox system works, as well as the information Curtis was telling me.
    - Eventually, we were able to resolve the issue: Eric now has access to the McGrath lab Dropbox file system.
- Continue literature review



5. *Results Visualization:* Nothing to visualize this week on my end.

Researcher

For a full list of what I did, see the time log above. The first big thing that I did this week was figuring out what was going on with the new SORT detections when the IoU threshold was lowered. I found that the new detections were only present on 53 infrared videos. This was very odd to me. I created a side-by-side visualization of the new and old IoU detections and posted it to the Bio-Boost Slack channel on 02/11/2025. See Figure 1 for an example screenshot of some of the SORT track differences.

Figure 1: New Detections with Different IoU

After having a meeting with Bree where we brainstormed a new direction for the BioBoost project, I started looking into finding the file locations, scripts, and training/test splits of each part of the pipeline. As I went along, I created a Lucidspark chart to more easily track down scripts and data for BioBoost. See Figure 2 for the current version of the BioBoost resource tracking chart.

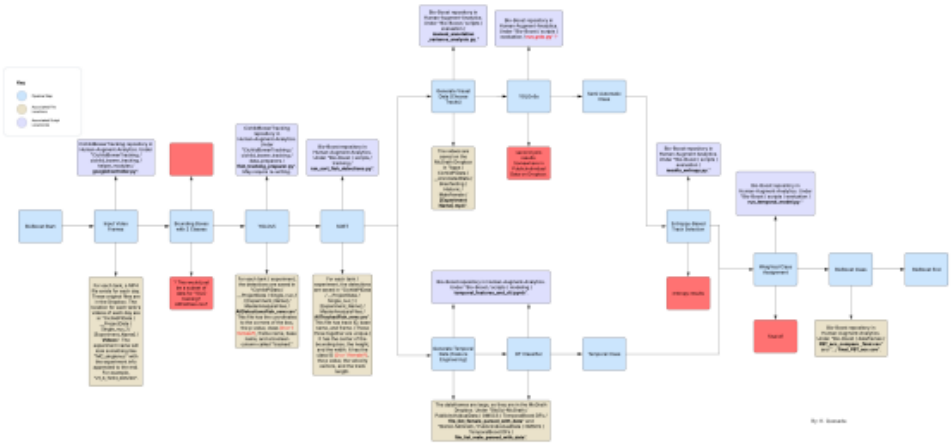


Figure 2: BioBoost Script and Data Tracking Chart

Finally, based on the project's new direction, I looked into Python packages for image segmentation. I was able to draw contours around fish videos from the Dropbox. I created 5 trials and ended up using the following steps for my fish contours: (1) RGB to HSV (2) Gaussian Blur (3) Grayscale (4) Threshold with Edges using .Canny (5) Dilating (6) Eroding + Somewhat Tuned Params (7) Inverting Edges (8) Contouring (9) Filtering Contours (Remove Small and Large Contours that aren't Fish) + Somewhat Tuned (10) Smoothing Contours +

Somewhat Tuned See the main chunk of the final code for this week in Listing 1. A more detailed explanation of how I got to this will follow.

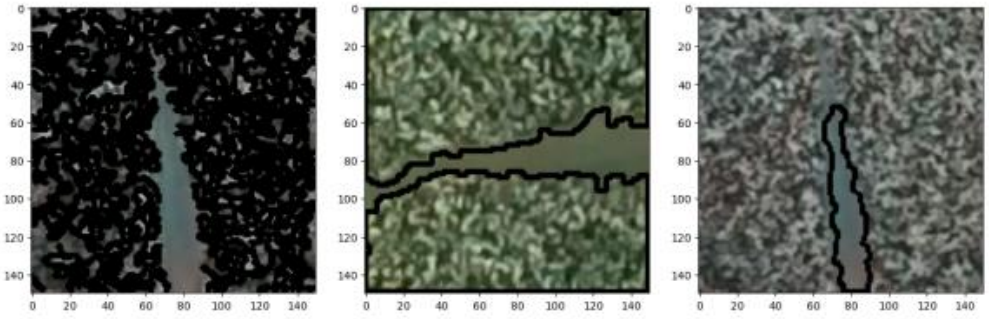
```
# Select Videos and Use OpenCV for Frame Processing Up Here
...
        # Convert Color Space to HSV
        hsv_frame = cv2.cvtColor(frame, cv2.COLOR_RGB2HSV)

        # Apply Blur
        blurred_img = cv2.GaussianBlur(hsv_frame, (3,3), 0)

        # Convert to Gray for Contouring
        gray_frame = cv2.cvtColor(blurred_img, cv2.COLOR_HSV2BGR)
        gray_frame = cv2.cvtColor(gray_frame, cv2.COLOR_BGR2GRAY)

        # Create Thresholding with Edges
        _, threshold_frame = cv2.threshold(gray_frame, 50, 255, cv2.THRESH_BINARY)
        edges = cv2.Canny(gray_frame, 50, 100)

        # Dilate the Edges for Better Contours
        kernel = np.ones((4, 4), np.uint8)
        dilated_edges = cv2.dilate(edges, kernel, iterations=2)
```

```

        # Erode the Edges for Better Contours
        kernel = np.ones((4, 4), np.uint8)
        eroded_dilated_edges = cv2.erode(dilated_edges, kernel, iterations=1)

        # Invert the Edges so that Fish are Contoured instead of Sand
        inverted_edges = cv2.bitwise_not(eroded_dilated_edges)

        # Find the Contours
        contours, _ = cv2.findContours(inverted_edges, cv2.RETR_EXTERNAL, cv2.
    CHAIN_APPROX_SIMPLE)

        # Filter Out Small and Big Contours
        # Removes Noise and Smaller Fish Tail Contours (TODO: BAD)
        contours = [contour for contour in contours if cv2.contourArea(contour) > 300]
        # Removes Tank Detections
        contours = [contour for contour in contours if cv2.contourArea(contour) < 3500]

        # Needed to Save the Contours to Array
        contour_frame = frame.copy()

        # Smooth Contours for Less Jagged Edges
        smoothed_contours = []
        for contour in contours:
            epsilon = 0.004 * cv2.arcLength(contour, True)
            approx_contour = cv2.approxPolyDP(contour, epsilon, True)
            smoothed_contours.append(approx_contour)
...
# Draw Contours and Save and View Videos and Frames Down Here
```

Listing 1: Using OpenCV to Generate Videos of Fish Contours

This was a rough draft since I am new to this, but it is a good starting point. This code was shared with other members of the team for their own usage. I generated videos with the contours added, as well as some individual frames for inspection. In one of my first couple of trials, I converted RGB to HSV, applied Gaussian blur, converted to grayscale, thresholded with edges using .Canny, and contoured. This resulted in everything but the fish being contoured, as can be seen in Figure 3a. In the next trial, I added dilating and eroding prior to contouring, resulting in everything but the fish being contoured, but there were only a couple of large contours instead of a bunch of small ones. This can be

seen in Figure 3b. Finally, I tuned parameters for eroding a bit, inverted the edges, and then added contour filtering and smoothing after contouring. This resulted in contours of the fish themselves, as shown in Figure 3c.



(a) Contour 1 (b) Contour 2 (c) Contour 3

Figure 3: Steps Taken to Find Fish Contours

And that's it! The finished scripts will be uploaded to BioBoost upon completion: https://github.c om/Human-Augment-Analytics/Bio-Boost.

Eric Iamarino

SAM Code:

```
from segment_anything import SamPredictor, sam_model_registry, SamAutomaticMaskGenerator
import cv2
import numpy as np
import matplotlib.pyplot as plt

import torch
import torchvision



"""
References:
- https://www.youtube.com/watch?v=fVeW9a6wltM&t=761s

"""


def show_annotations(data):
if len(data) > 0:
sorted_data = sorted(data, key=(lambda x: x['area']), reverse=True)
ax = plt.gca()
ax.set_autoscale_on(False)
polygons = []
color = []
for ann in sorted_data:
m = ann['segmentation']
img = np.ones((m.shape[0], m.shape[1], 3))
color_mask = np.random.random((1, 3)).tolist()[0]
for i in range(3):
img[:,:,i] = color_mask[i]
ax.imshow(np.dstack((img, m*0.35)))




def main():
# Test videos
videos = [
"MC_singlenuc23_1_Tk33_021220__0001_vid__1748_female.mp4",
"MC_singlenuc23_8_Tk33_031720__0001_vid__495_male.mp4",
"MC_singlenuc24_4_Tk47_030320__0001_vid__454_male_tricky.mp4"
]

# Setup device
device = torch.device("cpu")

# Load SAM model - trying with default/vit_h
sam = sam_model_registry["vit_h"](checkpoint="/Users/ericiamarino/Downloads/sam_vit_h_4b8939.pth") # trying with
default/vit_h
sam.to(device=device)
mask_generator = SamAutomaticMaskGenerator(
model=sam,
points_per_side=32,
```

```
pred_iou_thresh=0.90,
stability_score_thresh=0.96,
crop_n_layers=1,
crop_n_points_downscale_factor=2,
min_mask_region_area=300,
)



print(f"Using device: {device}")

# Process each video
for video in videos:
print(f"Processing video {video}")

# Start video
cap = cv2.VideoCapture(video)

# Choose output file name
output_file_name = ""
if video == "MC_singlenuc23_1_Tk33_021220__0001_vid__1748_female.mp4":
output_file_name = "female_segmented.mp4"
elif video == "MC_singlenuc23_8_Tk33_031720__0001_vid__495_male.mp4":
output_file_name = "male_segmented.mp4"
elif video == "MC_singlenuc24_4_Tk47_030320__0001_vid__454_male_tricky.mp4":
output_file_name = "male_tricky_segmented.mp4"

# Iterate through frames
for frame in range(3):
ret, f = cap.read()
if not ret:
print(f"Could not read frame {frame} from {video}")
break
frame_rgb = cv2.cvtColor(f, cv2.COLOR_BGR2RGB)

# Save frame before as plot
plt.figure(figsize=(10, 10))
plt.imshow(frame_rgb)
plt.axis('off')
plt.savefig(f"{output_file_name}_BEFORE.jpg")
plt.close()

# Use SAM to segment image
masks = mask_generator.generate(frame_rgb)
print(f"Objects detected: {len(masks)}")

# Save plot of detections as image
plt.figure(figsize=(10, 10))
plt.imshow(frame_rgb)
show_annotations(masks)
plt.axis('off')
plt.savefig(f"{output_file_name}_AFTER.jpg")
```

```
plt.close()


cap.release()

print(f"Saved: {output_file_name}")


if __name__ == "__main__":
main()
```

Modifications to Open CV Code:

```
# Tunable parameters
GAUSSIAN_KERNELS = [(3, 3), (5, 5)]
CANNY_LOW_VALS = [50, 60]
CANNY_THRESH_HIGH = 100
MIN_CONTOUR_AREAS = [300, 400]
EPSILON_FACTORS = [0.004, 0.005]

# Fixed parameters
THRESHOLD_VALUE = 50
DILATION_KERNEL_SIZE = (4, 4)
DILATION_ITERATIONS = 2
EROSION_ITERATIONS = 1
MAX_CONTOUR_AREA = 3500
```

Note: Paired these with automated loops that tested the different values Cichlid CV Website for updates: https://sites.gatech.edu/cichlid-computer-vision-project/ BioBoost

Working Meeting: https://gtvaultmy.sharepoint.com/personal/cclark339_gatech_edu/_layouts/15/stream.aspx?id=%2Fpersonal%2Fcclark339%5Fgatech%5Fedu%2FDocuments%2FRecordings%2FBioBoost%20Week%206%20Catch%2Dup%2D20250213%5F133803%2DMeeting%20Recording%2Emp4&referrer=StreamWebApp%2EWeb&referrerScenario=AddressBarCopied%2Eview%2E43ee9245%2Dacfa%2D4a48%2Da7fb%2D58f05a96aa8e