Advancing HAB monitoring with imaging flow cytometry and convolution neural networks

Ryan Tieu, Nick Soares, Maxim Batalin, Dhruvin Naik, Alexis Pasulka

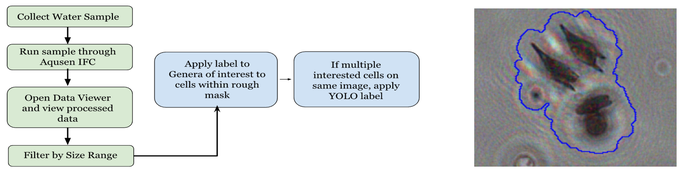
California Polytechnic State University, San Luis Obispo, Biological Sciences Department

Introduction

Harmful algal blooms (HABs) are increasing along the California coast, threatening ecosystems, aquaculture, and public health. Imaging flow cytometry (IFC) improves HAB monitoring by increasing temporal and spatial resolution. This project focused on manually labeling phytoplankton images to validate Aqusens IFC detections and improve AI network performance, along with tuning convolutional neural networks (CNNs) to provide accurate species concentration estimates. Together, these efforts improve early detection of HAB events and expand monitoring capacity along the California coast.

Methods

Image Analysis and Taxa Labelling:



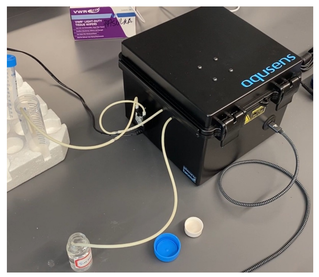
Schematic of workflow for running and analyzing an Aqusens sample (left) and example Aqusens image with a cluster of cells outlined in blue (right).

Phytoplankton Concentration Calculation:

Phytoplankton concentration (C; cells mL^{-1}) was determined from the number of cells counted (N), the number of images analyzed (P), and the volume per image (VF; mL).

$$C = \frac{N}{P \cdot VF}$$

- The Aqusens is a relatively small, portable, high-throughput and cost-effective imaging flow cytometer.
- Images are collected at a rate of 3 frames/sec and segmented into individual objects.
- Digital holograms of each object are constructed using pulses of red, green, and blue LEDs (Gorocs et al 2018).



Results

Comparison between microscopy and imaging flow cytometry for key HAB taxa

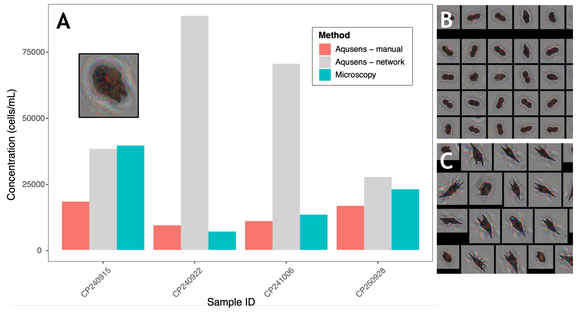


Figure 1. A) *Akashiwo sanguinea* concentration measured by microscopy (teal) and Aqusens manual counts (pink), and Aqusens IFC network detections (grey) with an Aqusens image as an inset. B and C) Aqusens images of *Margalefidinium* and *Ceratium*.

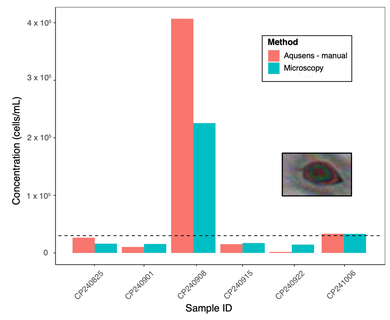


Figure 2. *Prorocentrum spp.* concentration measured by microscopy (teal) and Aqusens manual counts (pink) with an Aqusens image as an inset. The dashed line indicates a reference point for phytoplankton bloom event.

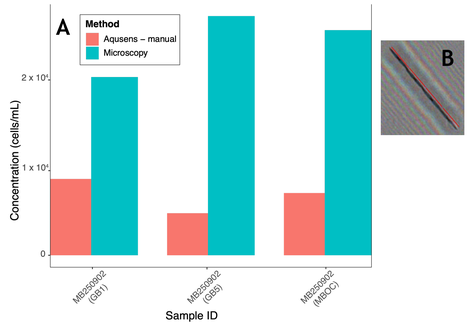


Figure 3. A) *Pseudo-nitzchia* concentration measured by microscopy (teal) and Aqusens manual counts (pink). B) Aqusens image with red lines indicating individual cells in the chain.

- Aqusens IFC manual counts were consistent with microscopy (Figs. 1A, 2) and captured a *Prorocentrum sp.* bloom event (Fig. 2).
- Automated network detections overestimated *A. sanguinea* abundance relative to microscopy and validated image counts (Fig. 1A), in part because the network misclassified cells during large blooms of morphologically similar taxa including *Margalefidinium sp.* and *Ceratium sp.* (Figs 1B, C) on Sept. 22nd and Oct. 6th, respectively.
- Expanding training datasets and improving multiclass models will reduce false positive detections.
- Chains of *Pseudo-nitzschia sp.* complicates quantification of individual cells and reduces agreement between methods (Fig. 3).
- Species-specific correction factors based on average cell length may improve agreement between methods.

Automating HAB Classification with CNNs

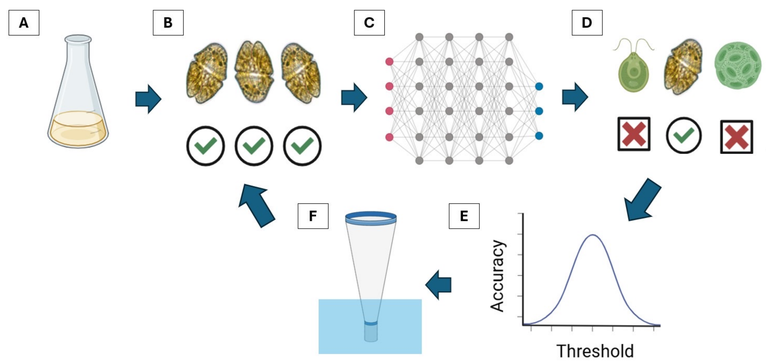


Figure 4. A. sanguinea *CNN development workflow. The first network iteration (Network A) was trained on cultured samples only (A), which were manually labeled (B) and used to train and validate a convolutional neural network (C). The network was then tested on a new set of imagery from marine field samples (D) to obtain performance metrics and to tune the decision threshold parameter (E). The second network iteration (Network B) incorporated field samples (F) into training and repeated steps B through E. Created with images from BioRender.com.*

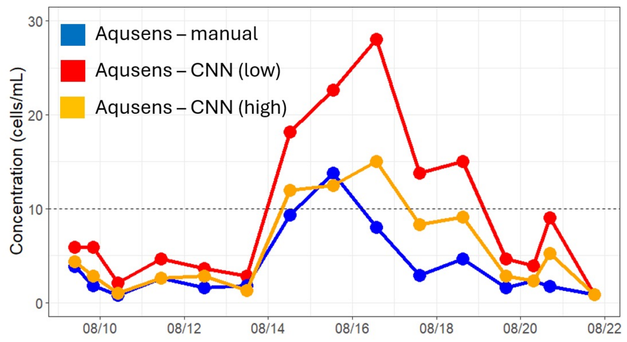


Figure 5. *The Aqusens was deployed at Grassy Bar Oyster Company in August 2026 and collects daily water samples. The same network processed the images twice, once at a high decision threshold (orange) and once at a low decision threshold (red). Images were manually annotated (blue) to determine which threshold quantifies* A. sanguinea *concentration with the highest accuracy.*

- Network development is an iterative process, consisting of multiple iterations of training and testing (Fig. 4). Once validated, networks can be used to detect HAB events (Fig. 5).

Future Directions

- Expand training datasets to improve automated classification accuracy across diverse HAB taxa
- Refine species-specific correction approaches (e.g., chain-forming taxa like *Pseudo-nitzschia*) to improve abundance estimates
- Deploy IFC-based monitoring at new sites to test early detection of HAB events

Acknowledgements

This research was generously supported by the Bill and Linda Frost Fund and the Santa Rosa Creek Foundation. Thank you to Lucendi, Inc. for working closely with us to continue to improve and develop HAB detection networks. We also appreciate the support and feedback from all the PECO lab members.