Algavision: Object Recognition Tool for Phytoplankton Cells in Microscopy Images

Simas Jasiūnas[1,2]*, Giulia Ghedini[1,3], Povilas Treigys[4]

[1] Gulbenkian Institute for Molecular Medicine, Oeiras, Portugal
[2] Instituto Superior Técnico, Lisbon, Portugal
[3] School of Biological Sciences, Monash University, VIC Australia
[4] Institute of Data Science and Digital Technologies, Vilnius University, Vilnius, Lithuania
* Correspondence: *si.jasiunas@gmail.com*

Abstract. Phytoplankton are microbes driving half of global primary production and are widely used model organisms in various studies. Despite advancements in high-throughput methods for acquisition of cell size and abundance data, microscopy remains widely preferred in multi-strain experiments. In this study, an instance segmentation model was trained on a dataset containing 11 strains and integrated into an image processing pipeline. The final version of the pipeline showed superiority over an established thresholding method (Avg. F1 Score 86% vs. 34%). This approach also highlights the potential of clustering unseen strains by prompting a general-purpose segmenter with low-confidence predictions.

Keywords: Cell segmentation, deep learning, microscopy, phytoplankton.

1 Introduction

Phytoplankton are aquatic photosynthetic organisms that sustain a large proportion of primary production [1]. Due to their tractability and fast reproduction, they are widely used in studies regarding ecology and evolution processes [2]. Phytoplankton size is particularly of great interest in such studies given its relationship with resource uptake, respiration and population growth [3]. However, while measuring cell size *per se* is relatively straightforward, tracking changes in cell morphology and cell number of individual strains within diverse communities is still challenging, because current methods cannot easily distinguish different strains due to their similarity. Therefore, microscopy remains a common approach to verify the identity of individual cells, despite being time consuming and prone to human errors.

DOI: https://doi.org/10.15388/LMITT.2026.10

Early segmentation algorithms (e.g. thresholding, watershed) can to some extent automate and speed up identification of cell strains in captured microscopy images, especially, if applied with defined filter sets [4, 5]. However, these image processing pipelines quickly fumble in multi-class segmentation when cells of different strains have similar morphological features – requiring substantial user input to correct misclassifications.

A different way of solving the task to automatically identify and measure cells would be through the use of deep learning models, which have been gaining popularity in the last two decades [6]. These models operate in an end-to-end learning fashion, which extract subtle undefined features of each strain, thus substantially outperforming early segmentation methods on datasets containing multiple classes [7]. However, to our knowledge there is no unified framework or application that allows researchers to use these models for multi-strain phytoplankton research from the get-go, thus demanding particular expertise, computational resources and time allocation for dataset acquisition and model development [8].

The goal of this project is to improve the process of multi-strain phytoplankton cell identification and measurement of their morphological traits (e.g. size, shape, density) in light microscopy images. To achieve that, a state-of-the-art Mask-RCNN instance segmentation model [9] was trained on a newly created dataset, consisting of 11 phytoplankton strains. The acquired model's precision was then increased by implementing a custom post-processing logic, which adjusts the confidence score of each prediction. The final version of the image processing pipeline was shown to be highly superior to the previously established segmentation method. Lastly, low confidence predictions (< 0.1%) produced by the Mask-RCNN were repurposed as prompts for a general-purpose model SAM [10], which generates segmentation masks for novel phytoplankton strains from natural samples.

2 Methods and Materials

All phytoplankton strains were purchased from the Roscoff Culture Collection (France) and were acclimatized in a laboratory in a temperature (~20^{o}C) and light controlled (~50 µmol/m²/s) environment for several months prior to imaging and cultured in sterilized, filtered seawater medium [11]. For imaging, samples from each strain were fixed with Lugol's iodine 1% solution (40:1000 v/v ratio) and loaded onto hemocytometers

(10 µl) or microplates (200 µl). The images were taken with an Olympus IX73 inverted microscope at 400x magnification (Evident Scientific, Evident Europe GmbH, Germany). Acquired images were annotated with the image annotation tool CVAT [12] resulting in a dataset of 86731 annotated cells from 11 phytoplankton strains (Fig. 1).

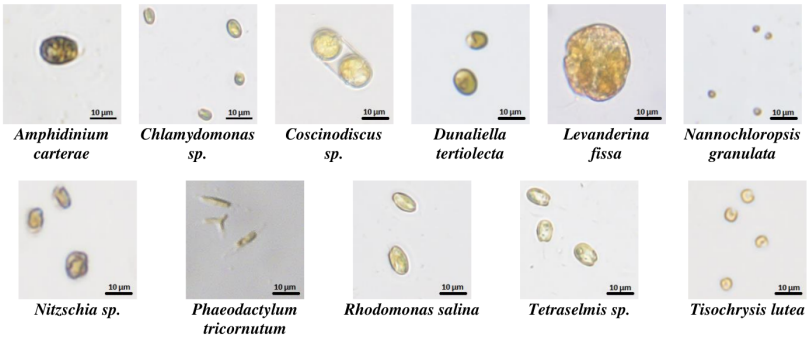

Figure 1. Phytoplankton strains included in the dataset with scale indicated in the lower left corner.

This dataset was used to train, validate and test (80:10:10 split) a Mask-RCNN instance segmentation model, a popular two-stage detector known for its accuracy, which builds up on top of Faster-RCNN by adding a branch for predicting object masks [9, 13]. An implementation of Mask-RCNN model and a checkpoint, that was pre-trained on the COCO dataset [14], was acquired from MMDetection framework's model zoo [15]. The model was trained on the phytoplankton dataset in Google Colab environment on a single L4 GPU (22.5 GB VRAM) for three epochs.

The precision of the trained model was later improved by including a post-processing step, which adjusts the original prediction's confidence score based on the morphological characteristics of each strain: area, major and minor axis, aspect ratio and roundness. Each phytoplankton strain has specific mean (µ) and standard deviation (σ) values of their probability density functions for each of these characteristics, which were acquired by fitting a Log-Normal distribution for each strain in the annotated dataset. The likelihood of the predicted class by the model can be calculated for each characteristic (Eq. 1) and the confidence score can then be adjusted accordingly (Eq. 2).

$$f(x;\mu,\sigma) = \frac{1}{x\sigma\sqrt{2\pi}} exp(-\frac{(ln\,x-\mu)^2}{2\sigma^2})$$

Equation 1. Likelihood estimation. Where: x - prediction's value (area, major, minor, aspect ratio or roundness); μ - mean; σ - standard deviation.

$$adj.\ score = org.score \times \sqrt[5]{\frac{f(area)}{f(area\ m.)} \times \frac{f(major)}{f(major\ m.)} \times \frac{f(minor)}{f(minor\ m.)} \times \frac{f(asp.ratio)}{f(asp.ratio\ m.)} \times \frac{f(round.)}{f(round.\ m.)}}$$

Equation 2. Adjustment of the original confidence score. Each characteristic's likelihood is normalized by dividing it by the maximum possible likelihood (mode value (m.)); therefore, each normalized likelihood value has the same contributive weight for adjusting the original score.

The constructed confidence adjustment logic, together with the trained Mask-RCNN model, was built into an image processing pipeline and was compared to a previously established cell segmentation method based on Yen's thresholding, implemented in Fiji software [16, 17]. The comparison was performed on three phytoplankton community datasets that were acquired after model training. Each dataset was obtained by combining five different strains and taking 50 images per community. F1 score was averaged for all three datasets and was used to compare the accuracy of both approaches (Eq. 3 and 4).

$$F1\ Score = \frac{2TP}{2TP + FP + FN}$$

Equation 3. F1 Score. TP - true positives; FP - false positives; FN - false negatives.

$$Avg.\ F1\ Score = \frac{F1_1 + F1_2 + F1_3}{3}$$

Equation 4. Average F1 score. F1n is the F1 score of each dataset containing the five strains.

Finally, the potential of exploiting low confidence predictions when subjecting images containing unknown strains was explored. In this case, the centers of low confidence segmentation masks (< 0.1%) were used as foreground prompts for the general-purpose ViT-B SAM model. To test the segmentation accuracy of unknown strains a natural sample dataset was created from 50 images containing 354 unknown phytoplankton cells. Segmentation performance was assessed by comparing results before and after SAM prompting. This evaluation included confusion matrices based on detection outcomes and distributions of true positives across their corresponding intersection-over-union (IoU) values.

3 Results

The model's ability to correctly identify and segment 11 phytoplankton strains after training was tested on the test split and was visualized with a confusion matrix (Fig. 2A). As expected, most misclassifications (mis.) occurred between morphologically similar phytoplankton cells, such as *Tisocrhysis* and *Chlamydomonas* (583 mis.) or *Rhodomonas* and *Tetraselmis* (92 mis.). However, the most problematic species by far was the smallest one – *Nannochloropsis*, which was often confused with background objects (debris, etc.) or completely missed (816 mis.). Diatoms – (*Coscinodiscus*, *Nitzschia*, *Phaeodactylum*), were also often confused with background, probably due to their semi-transparent silica shell. Interestingly, the model also occasionally confused morphologically different cells, for example *Chlamydomonas*, which usually has a rounder shape than the more elongated *Phaeodactylum* (39 mis.). To mitigate such occurrences between morphologically different cells a confidence adjustment logic was added (Fig. 2B). These adjustments mainly removed low confidence cases, like *Nannochloropsis* and *Coscinodiscus* being misclassified as debris (161 and 112 mis. less). However, this approach also decreased true positives for some species, mainly *Tetraselmis* and *Chlamydomonas* (55 and 26 mis. more). In total, the model's precision was slightly boosted (+2.5%) at a very small cost of recall (-0.2%).

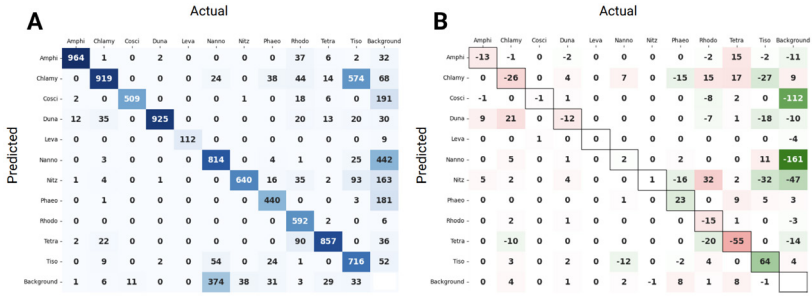

Figure 3. A: Confusion matrix of the prediction results on the test split. Blue shading indicates the number of occurrences. Ground truth and prediction mask overlapping threshold ≥ 0.5 IoU; **B:** Confusion matrix of the change of test split predictions after custom post-processing. Green and red shading indicate improvement and error, respectively. Note that diagonal cells (borders highlighted) are true positives – increases in the number of cells counted are colored in green because that equals improvement.

The constructed image processing pipeline was finally compared to a previously established cell segmentation method based on Fiji macros that measured and counted different strains based on their approximate cell area and shape. Our pipeline mainly struggled with identification and separation of smaller cells that tend to form clusters (*Nannochloropsis*, *Chlamydomonas*), while Fiji macros were majorly dependent on the contrast between the background and the cells, making it a less reliable method, especially in images generated from microplates or containing debris (Fig. 4). Overall, our pipeline's performance (Avg. F1 Score 86.2%) was significantly better than Fiji's (Avg. F1 Score 34.3%) when averaging across the three datasets.

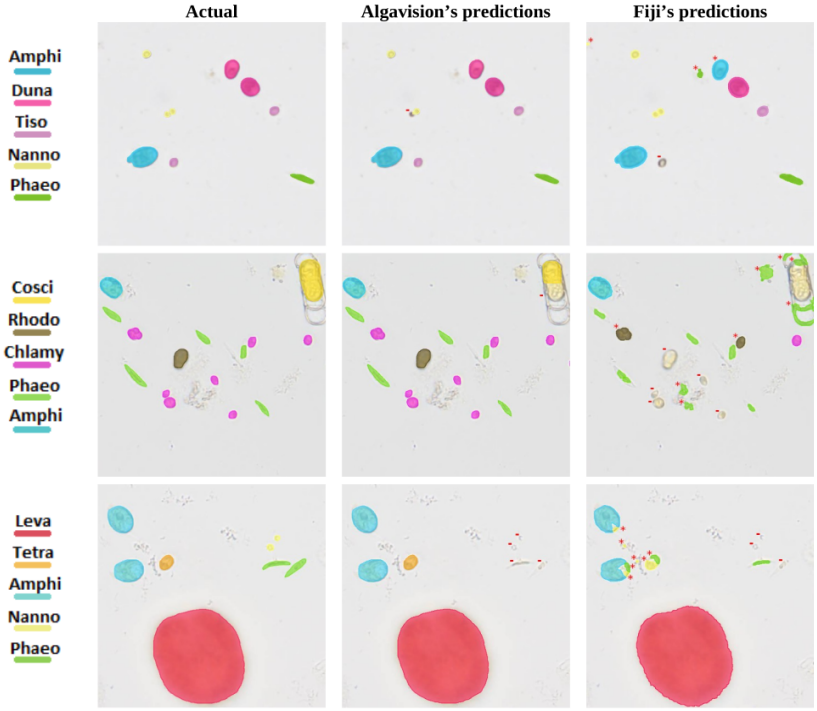

Figure 4. Visual examples of Algavision (our pipeline) and Fiji performance on the three community datasets. Red minuses represent missed cells, while red asterisks represent misclassifications.

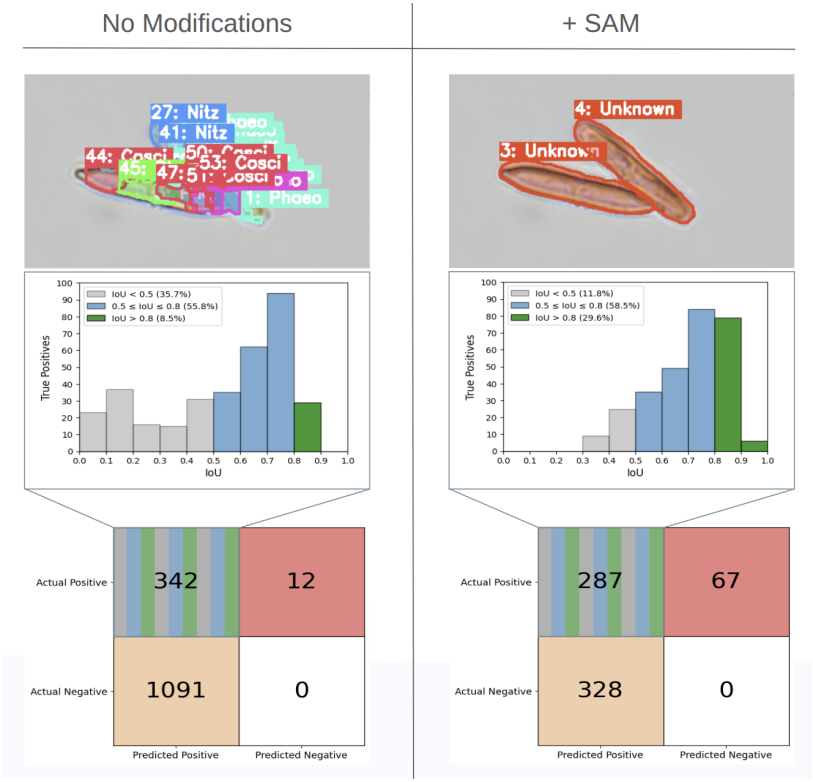

Figure 5. Top: Visual examples of Algavision (our pipeline) trying to identify and segment unknown strains before and after prompting the ViT-B SAM model; **Middle:** Count of true positives against their IoU distributions; **Bottom:** Confusion matrices of both approaches. Ground truth and prediction mask overlapping threshold (to be counted as true positive) > 0 IoU.

Lastly, we tested our pipeline's ability to segment phytoplankton strains that the model was not trained on by using low confidence predictions (< 0.1%) as prompts for the ViT-B SAM model. Even though the Mask-RCNN model did not encounter specific instances in its training set that it is trying to predict, it is still able to generate superficial masks on unknown cells or parts of them due to Mask-RCNN's region proposal. The addition of the ViT-B SAM model mostly increased the generation of highly accurate segmentation masks (> 0.8 IoU: from 8.5% to 29.6%) (Fig. 5). Furthermore,

the false positives were substantially reduced (from 1091 to 328) due to more uniform mask generation, thus increasing the F1 score from 38.3% to 59.2% when trying to segment unknown strains. In the future, such highly accurate masks could be used to facilitate the creation of new datasets or be subjected to clustering by morphological features into a bunch of anticipated strains.

4 Conclusions

In this work, we created Algavision, a wrapper tying Mask-RCNN and SAM models together to extract relevant characteristics of each known or unknown cell and to facilitate research based on phytoplankton communities using light microscopy images. By harnessing the power of deep learning, our image processing pipeline substantially outperformed a previously established thresholding method in all our testing cases (Avg. F1 Score 86.2% vs. 34.3%), making it a more accurate tool in terms of both precision of species identification and quality of the generated masks from which cell size measures are extracted. We also explored the opportunity to repurpose low confidence predictions on unknown species by using them as prompts for SAM, which increased the generation of high quality segmentation masks for unknown species (> 0.8 IoU: from 8.8% to 24.6%), increasing our pipeline's F1 score from 38.3% to 59.2% when predicting unknown strains. We believe that with further improvements, Algavision will become a useful software for the collection of biodiversity data of phytoplankton and other microorganisms.

References

[1] Field CB, Behrenfeld MJ, Randerson JT, Falkowski P. Primary production of the biosphere: integrating terrestrial and oceanic components. Science. 1998;281:237–240. doi:10.1126/science.281.5374.237

[2] Henson SA, Cael BB, Allen SR, Dutkiewicz S. Future phytoplankton diversity in a changing climate. Nat Commun. 2021;12:5372. doi:10.1038/s41467-021-25699-w

[3] Finkel ZV, Beardall J, Flynn KJ, Quigg A, Rees TAV, Raven JA. Phytoplankton in a changing world: cell size and elemental stoichiometry. J Plankton Res. 2010;32:119–137. doi:10.1093/plankt/fbp098

[4] Otsu N. A threshold selection method from gray-level histograms. IEEE Trans Syst Man Cybern. 1979;9:62–66. doi:10.1109/TSMC.1979.4310076

[5] Beucher S, Meyer F. Segmentation: the watershed transformation. In: Dougherty ER, editor. Mathematical morphology in image processing. New York: Marcel Dekker; 1993. p. 433–481.

[6] Tian D, Han Y, Wang B, Guan T, Gu H, Wei W. Review of object instance segmentation based on deep learning. J Electron Imaging. 2021;31:041205. doi:10.1117/1.JEI.31.4.041205
[7] Mahony N, Campbell S, Carvalho A, Harapanahalli S, Velasco-Hernandez G, Krpalkova L, et al. Deep learning vs. traditional computer vision. 2020. doi:10.1007/978-3-030-17795-9
[8] Bachimanchi H, Pinder MIM, Robert C, De Wit P, Havenhand J, Kinnby A, et al. Deep-learning-powered data analysis in plankton ecology. Limnol Oceanogr Lett. 2024;9:324–339. doi:10.1002/lol2.10392
[9] He K, Gkioxari G, Dollár P, Girshick R. Mask R-CNN. In: Proc IEEE Int Conf Comput Vis (ICCV); 2017. p. 2980–2988. doi:10.1109/ICCV.2017.322
[10] Kirillov A, Mintun E, Ravi N, Mao H, Rolland C, Gustafson L, et al. Segment anything [preprint]. arXiv. 2023. doi:10.48550/arXiv.2304.02643
[11] Guillard RRL, Ryther JH. Studies of marine planktonic diatoms: I. Cyclotella nana Hustedt and Detonula confervacea (Cleve) Gran. Can J Microbiol. 1962;8:229–239. doi:10.1139/m62-029a
[12] Sekachev B, Manovich N, Zhiltsov M, Zhavoronkov A, Kalinin D, Hoffman S, et al. Computer Vision Annotation Tool (CVAT) [Internet]. 2020. Available from: https://github.com/cvat-ai/cvat
[13] Ren S, He K, Girshick R, Sun J. Faster R-CNN: towards real-time object detection with region proposal networks. IEEE Trans Pattern Anal Mach Intell. 2016;39:1137–1149. doi:10.1109/TPAMI.2016.2577031
[14] Lin TY, Maire M, Belongie S, Hays J, Perona P, Ramanan D, et al. Microsoft COCO: common objects in context. In: Fleet D, Pajdla T, Schiele B, Tuytelaars T, editors. Comput Vis – ECCV 2014. Cham: Springer; 2014. p. 740–755. doi:10.1007/978-3-319-10602-1_48
[15] Chen K, Wang J, Pang J, Cao Y, Xiong Y, Li X, et al. MMDetection: open MMLab detection toolbox and benchmark [preprint]. arXiv. 2019. doi:10.48550/arXiv.1906.07155
[16] Yen JC, Chang FJ, Chang S. A new criterion for automatic multilevel thresholding. IEEE Trans Image Process. 1995;4:370–378. doi:10.1109/83.366472
[17] Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. Nat Methods. 2012;9:676–682. doi:10.1038/nmeth.2019