Developing a convolutional neural network to classify phytoplankton images collected with an Imaging FlowCytobot along the West Antarctic Peninsula

Schuyler C. Nardelli
Center for Ocean Observing Leadership
Rutgers University
New Brunswick, NJ, USA
nardelli@marine.rutgers.edu

Patrick C. Gray
Marine Lab
Duke University
Beaufort, NC, USA
patrick.c.gray@duke.edu

Oscar Schofield
Center for Ocean Observing Leadership
Rutgers University
New Brunswick, NJ, USA
oscar@marine.rutgers.edu

***Abstract*—High-resolution optical imaging systems are quickly becoming universal tools to characterize and quantify microbial diversity in marine ecosystems. Automated detection systems such as convolutional neural networks (CNN) are often developed to identify the immense number of images collected. The goal of our study was to develop a CNN to classify phytoplankton images collected with an Imaging FlowCytobot for the Palmer Antarctica Long-Term Ecological Research project. A medium complexity CNN was developed using a subset of manually-identified images, resulting in an overall accuracy, recall, and f1-score of 93.8%, 93.7%, and 93.7%, respectively. The f1-score dropped to 46.5% when tested on a new random subset of 10,269 images, likely due to highly imbalanced class distributions, high intraclass variance, and interclass morphological similarities of cells in naturally occurring phytoplankton assemblages. Our model was then used to predict taxonomic classifications of phytoplankton at Palmer Station, Antarctica over 2017-2018 and 2018-2019 summer field seasons. The CNN was generally able to capture important seasonal dynamics such as the shift from large centric diatoms to small pennate diatoms in both seasons, which is thought to be driven by increases in glacial meltwater from January to March. Moving forward, we hope to further increase the accuracy of our model to better characterize coastal phytoplankton communities threatened by rapidly changing environmental conditions.**

Keywords—machine learning, neural network, phytoplankton, polar science

I. Introduction

The West Antarctic Peninsula (WAP) is a highly productive marine ecosystem characterized by large summer phytoplankton blooms that support extensive krill and top predator populations [1]. The WAP is experiencing significant environmental change, threatening this unique and productive ecosystem. One of the fastest warming regions on Earth, WAP winter air temperatures and surface ocean temperatures have increased by 6°C and >1°C, respectively, over the past 50 years [2-4]. In response, 90% of marine glaciers are currently in retreat, the annual ice season has decreased by 92 days over the last 35 years, and there is no longer perennial sea ice in the northern WAP [2], [5].

Ocean warming and melting sea ice have impacted the phytoplankton community, which has implications for the entire food web. Phytoplankton biomass has significantly decreased in the northern WAP, associated with a shift from large-celled diatoms to smaller-celled cryptophytes and mixed flagellates [6]. This shift is concurrent with an increase in low salinity meltwater [7-9]. The increased spatial coverage of low salinity surface waters associated with continued glacial and sea ice melt is predicted to increase the prevalence of smaller-celled phytoplankton communities along the WAP, with important implications for food web structure and energy transfer efficiency [10].

The Palmer Long-Term Ecological Research project (PAL-LTER) was established in 1991 to investigate how changes in sea ice structure the pelagic ecosystem and biogeochemistry along the WAP. The project has previously used High Performance Liquid Chromatography (HPLC) analysis of pigment data to characterize the taxonomic composition of phytoplankton assemblages [11]. This technique uses marker pigments of phytoplankton groups to assess their contribution to the overall abundance. However, HPLC lacks more detailed taxonomic classification and cell size information that is critical to understanding how warming and melting impacts phytoplankton communities along the WAP.

To fill this knowledge gap, in 2017 the PAL-LTER acquired an Imaging FlowCytobot (IFCB; McLane Labs, Falmouth, MA, USA). The IFCB is an automated imaging-in-flow submersible cytometer that uses a combination of video and flow cytometric technology to collect images and measure chlorophyll fluorescence and scattered light for each particle (~10-150 μm) in a 5 mL water sample [12]. These images can be analyzed to determine cell size dynamics, and sorted taxonomically to the genus or species level, thus providing much more detailed organismal information than HPLC methods.

However, the IFCB can generate more than 10,000 high-quality images every hour, which becomes an immense amount of data over the duration of a research cruise or field season. This volume of data makes manual image identification impractical, therefore, these imaging platforms are often complemented by automated detection systems that allow for rapid and precise classification of plankton communities. Currently, there are two typical machine learning approaches for IFCB images: (1) a support vector machine based on a feature selection algorithm (88% overall accuracy with 22 classes; [13]), and (2) random forest (RF) algorithms (~70% overall accuracy depending on the

978-0-7892-3559-0/21/$31.00 ©2021 MTS

model and number of classes, e.g., [14]). Following advancements in the field of computer vision through deep learning [15] the IFCB community is now transitioning to convolutional neural networks (CNNs) for improved accuracy in image classification. CNNs extract features directly from images. Starting with raw imagery and labels, semantically meaningful features are learned as the network trains on a set of images. In theory, extracted features correspond to components of the image relevant to the labels, which makes these models highly accurate and well-suited for image classification tasks.

Since 2017, the PAL-LTER has collected over 10 million images spanning four summer field seasons. The goal of our study was to develop a CNN to sort WAP phytoplankton into taxonomic groups. This would allow for taxonomic classification of entire seasons of collected phytoplankton data in a short amount of time. Additionally, the CNN could be used as a tool to characterize phytoplankton communities in the field in near-real time to inform opportunistic sampling strategies. The combination of the IFCB and a high-accuracy automated classification system would allow the PAL-LTER to learn more about shifts in phytoplankton community and size dynamics associated with rapidly changing environmental conditions.

II. Methods

A. Phytoplankton Image Collection and Processing

IFCB data were collected along the West Antarctic Peninsula over three summer field seasons: 2017-2018, 2018-2019, and 2019-2020. Whole water samples were collected at various depths from both the January cruise along the WAP (Anvers Island in the north to Charcot Island in the south) and from seasonal (November-March) sampling at Palmer Station, Antarctica. 5 mL from each sample was analyzed with the IFCB to acquire images for each phytoplankton cell in the sample. Samples were passed through a 150 µm Nitrex screen prior to analysis to prevent large cells from clogging the IFCB's flow cell. Cells with a major axis length < 25 pixels (7.35 µm) were eliminated from the analysis as the resolution of the images was insufficient to provide clear identification.

Images were processed using methods and software from [13] (https://github.com/hsosik/ifcb-analysis/wiki). Image processing results in a set of 233 features describing each image including fluorescence, scattering intensity, equivalent spherical diameter, area, volume, and other morphometric parameters such as image texture and histogram of oriented gradients.

B. Model Development

Processed images, metadata, and their associated features were uploaded to the web application EcoTaxa (https://ecotaxa.obs-vlfr.fr) [14]. Using EcoTaxa, a subset of 18,699 images was visually inspected and manually classified into 38 living groups (taxonomic resolution ranging from genus to class) and 2 non-living groups (detritus and bubbles), with at least 100 images per group. Samples (images + features) were augmented to increase training sample size via image rotations, flips, gaussian noise, and contrast changes. Features were also randomly multiplied by a factor between 0.8 and 1.2.

After augmentation, a training dataset of 40,000 samples with 1,000 in each class was used to develop a medium complexity CNN (8 convolutional layers and 2 million parameters), and 3,740 unaugmented images, approximately evenly split across classes, were used as a validation dataset. Model precision, recall, and f1-score were calculated for the unmerged data considering all included groups, and for merged data considering only 8 general taxonomic groupings (pennate and centric diatoms, cryptophytes, prasinophytes, mixed flagellates, haptophytes, microzooplankton, and other). The "other" group includes primarily detritus with some bubbles. Precision is defined as true positives divided by the sum of true positives and false positives; it is the proportion of positive identifications that are correct. Recall is defined as true positives divided by the sum of true positives and false negatives; it is the proportion of actual positives that are identified correctly. The f1-score is the harmonic mean of precision and recall. Confusion matrices were also generated showing the percent of manually validated images predicted in each category by the CNN.

C. Model Validation

To test the model, we used it to predict on a random subset of 10,269 new images filtered by cell major axis length > 25 pixels. Additionally, we used EcoTaxa's RF algorithm to predict on the same images, using a maximum of 500 images per group. Predictions from both models were compared to manual identification of the images. Model precision, recall, and f1-score were calculated for unmerged and merged data for both the CNN and RF models, and a confusion matrix was generated for the CNN.

D. Model Application

After training and evaluation, our model was used to predict taxonomic classifications of phytoplankton collected at 0 m from Station B near Palmer Station, Antarctica over the 2017-2018 and 2018-2019 summer field seasons. CNN predictions were compared to manual validation of the images to determine the accuracy of the predicted seasonal trends.

E. Sea Ice Characterization

Sea ice data were calculated using version 3.1 of the GSFC Bootstrap sea ice concentrations. Sea ice duration is the time elapsed between day of advance and day of retreat. All sea ice metrics use the 200 km area south and west of Palmer Station. See [16] for more information.

III. Results

A. Model Accuracy

The overall precision, recall, and f1-score of the model were 93.8%, 93.7%, and 93.7%, respectively. After merging the initial set of 40 classes into the 8 broader taxonomic groups, the precision, recall, and f1-score of the model all increased to 96.5%. Accuracy per group was > 95% for all groups except for microzooplankton (> 80%), mixed flagellates (> 90%), and other (> 90%).

TABLE I. CONFUSION MATRIX FOR BROAD TAXONOMIC GROUPS USING 10,269 NEW, RANDOM IMAGES

True label \ Predicted label	Pennate diatoms (*n*=2788)	Centric diatoms (*n*=601)	Cryptophytes (*n*=1898)	Prasinophytes (*n*=1040)	Mixed flagellates (*n*=2304)	Haptophytes (*n*=16)	Microzooplankton (*n*=803)	Other (*n*=3475)
Pennate diatoms (*n*=1577)	**92.9**	0.8	0.3	0.6	4.8	0.0	0.0	0.7
Centric diatoms (*n*=249)	2.8	**64.3**	5.2	2.4	15.3	0.0	0.0	10.0
Cryptophytes (*n*=2565)	9.4	1.0	**65.0**	4.4	19.8	0.0	0.0	0.5
Prasinophytes (*n*=493)	2.6	1.6	0.4	**39.6**	28.0	0.0	0.0	27.8
Mixed flagellates (*n*=1085)	11.6	1.5	3.9	7.4	**66.2**	0.0	0.3	9.2
Haptophytes (*n*=1)	0.0	0.0	0.0	0.0	0.0	**100**	0.0	0.0
Microzoo-plankton (*n*=6)	0.0	16.7	0.0	0.0	16.7	0.0	**66.7**	0.0
Other (*n*=3475)	26.9	10.9	4.9	18.4	23.8	0.0	0.3	**14.9**

Using the model to predict on the 10,269 new images resulted in unmerged and merged f1-scores of 46.5% and 47.6%, respectively. This is a 12% increase in the unmerged f1-score over EcoTaxa's random forest model (46.5% vs. 41.5%, respectively; [14]). The model predicted most accurately for pennate diatoms (92.9%), and performed moderately well for microzooplankton (66.7%), mixed flagellates (66.2%), cryptophytes (65.0%), and centric diatoms (64.3%; Table 1). Our model was least precise predicting prasinophytes (39.6%) and other cells (14.9%; Table 1). Only one haptophyte was manually identified in the random dataset but was predicted correctly.

B. Phytoplankton Seasonal Succession at Palmer Station

Overall, the CNN captured important seasonal trends in phytoplankton dynamics. In both the 2017-2018 and 2018-2019 seasons, peak phytoplankton biovolume occurred midseason (1 January 2018 and 4 February 2019; Fig. 1). In 2017-2018, the peak was dominated by a mix of cryptophytes, prasinophytes, and mixed flagellates, while in 2017-2018 the peak was dominated by pennate diatoms. The CNN also captured early and late season peaks composed of centric diatoms in 2018-2019 (Fig. 1C-D).

However, there are several discrepancies between methods. In both seasons, but particularly 2017-2018, there were many cells manually identified as "other" that were classified as both mixed flagellates and prasinophytes by the CNN (Fig. 1). In this manner, the CNN appears to overestimate the abundance of these groups. The CNN also underestimated the abundance of cryptophytes, especially during peak biovolume in both years. Importantly, this misclassification of "other" cells also greatly overestimates the phytoplankton biovolume compared to manual validation, causing the seasonal phytoplankton peak in 2017 to appear much higher than for manual validation (Fig. 1A-B).

The CNN also captured interesting seasonal trends in the diatom community. There was less total diatom biovolume in 2017-2018 compared to 2018-2019 (Fig. 2A, 2D). In both seasons, centric diatoms shifted from a dominance of >20 µm

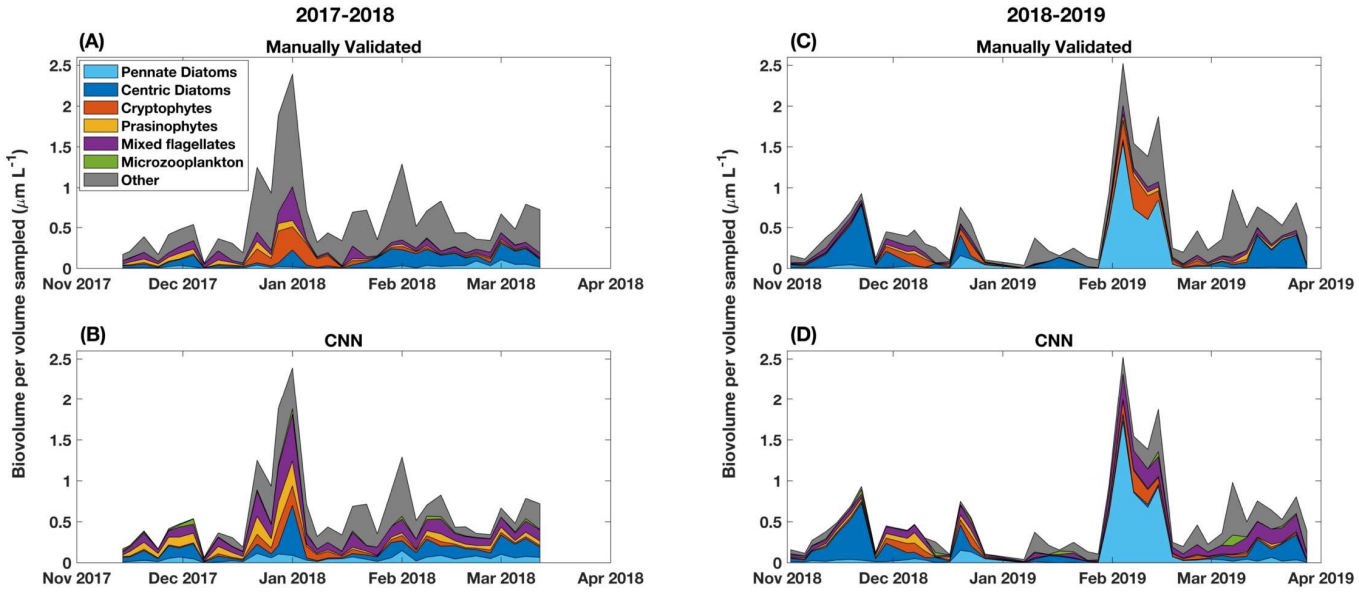

Fig. 1. Methods comparison of phytoplankton seasonal succession for the (A-B) 2017-2018 and (C-D) 2018-2019 summer field seasons, showing biovolume data from (A and C) manual validation and (B and D) CNN predictions.

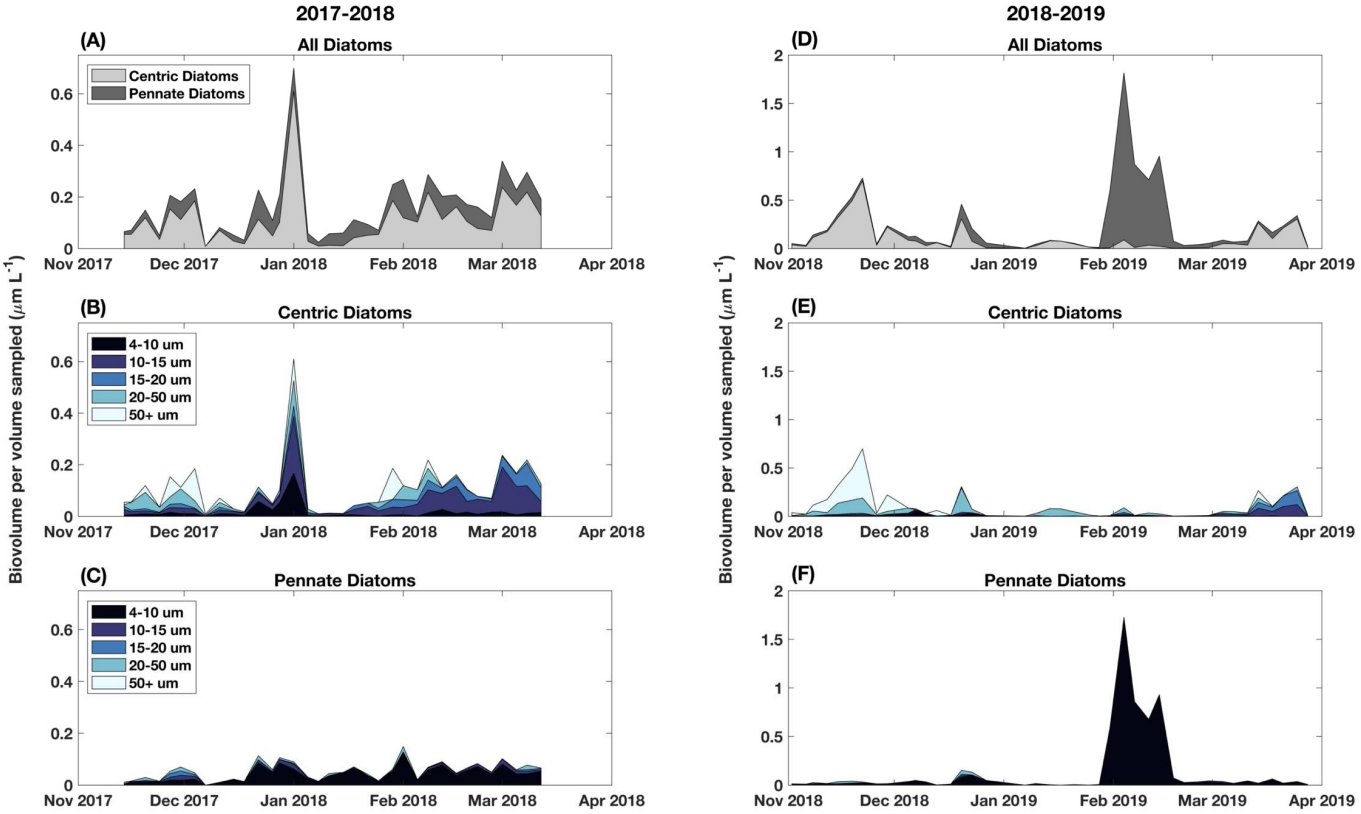

Fig. 2. Diatom seasonal diversity as predicted with the CNN for the (A-C) 2017-2018 and (D-F) 2018-2019 summer field seasons. (A and D) Total biovolume attributed to pennate and centric diatoms. (B and E) Total biovolume attributed to different size classes of centric diatoms. (C and F) Total biovolume attributed to different size classes of pennate diatoms.

cells in November and December, to a dominance of 10-15 µm cells in February and March (Fig. 2B, 2E). Pennate diatoms were consistently dominated by cells < 10 µm, with an increase in biovolume during February and March, especially in 2018-2019 (Fig. 2C, 2F). Both seasons were primarily dominated by centric diatoms, with the notable exception of a large peak in pennate diatom biovolume in 2018 (Fig. 2A, 2D).

C. Sea Ice Dynamics

2017 had lower maximum winter sea ice coverage and shorter sea ice duration than 2018, but a later sea ice retreat (Table 2 and Fig. 3). Sea ice cleared the region rapidly in 2017, dropping from 52% coverage in November, to 12% in December, and 3% in January (Fig. 3). In 2018, the sea ice retreated earlier but coverage stayed higher in the region into the summer, with 24% coverage in November, 17% coverage in December, and 10% coverage in January (Fig. 3).

TABLE II. SEA ICE CHARACTERIZATION

Year	Sea Ice Duration (days)	Date of Sea Ice Retreat
2017	132	December 3
2018	153	November 27

IV. DISCUSSION

A. Model Development: Successes and Challenges

Overall, we achieved the goal of our study: to create a CNN to accurately sort WAP phytoplankton into taxonomic categories. Our overall model achieved an f1-score of 93.7% with an increase to 96.5% for merged taxonomic groupings. This indicates that our phytoplankton imagery data can be successfully and accurately sorted with machine learning techniques, greatly reducing the time spent classifying these images manually. Absolute comparisons to classification algorithms from previous studies is challenging given different numbers of classes, data filtering schemes, and methods for determining what constitutes test data, but in general these metrics compare very favorably to other models. The development of regional and global phytoplankton classifying CNNs presents an opportunity to greatly advance our understanding of plankton diversity and ecology.

However, our model f1-score dropped dramatically from 93.7% during model development to 46.5% during model validation on a new, random dataset with a class distribution representative of that found in natural waters. We believe that this large decrease in model accuracy is a key challenge rarely addressed in the literature. One reason for this decrease is the highly imbalanced class distributions of naturally occurring phytoplankton assemblages compared to our model testing

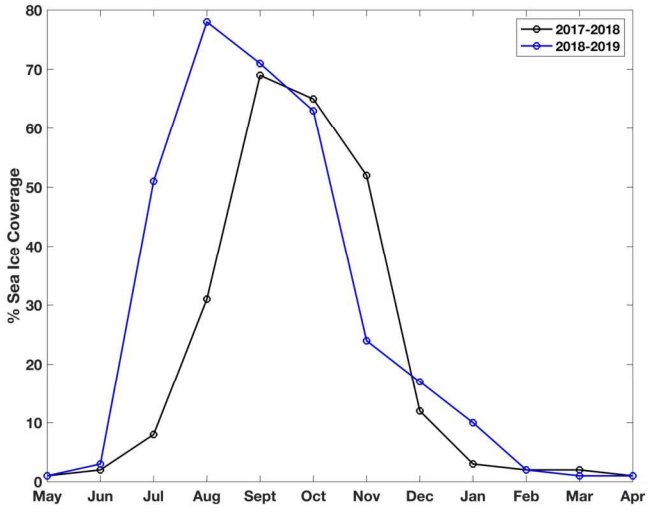

Fig. 3. Percent sea ice coverage in the 200 km area south and west of Palmer Station during the 2017-2018 season (black) and the 2018-2019 season (blue).

dataset (e.g., see *n* values in Table 1). Model categories such as detritus are highly abundant in our dataset, often composing up to 50% of the biovolume in a sample, while other ecologically important groups, such as large, morphologically distinct diatoms including *Corethron penatum* and *Eucampia antarctica* are encountered sporadically in our dataset. A minor misclassification of detritus as a rare class can easily overwhelm that category.

Nearly all previous studies report accuracy for a balanced and curated test dataset rather than a random sample of natural waters. During model development a balanced class distribution is necessary to ensure the model equally weights each category during training. For example, if during model development a single class composed 90% of the training dataset, the model could classify every sample as that class, ignoring all others, and be 90% accurate. The gradient descent optimization algorithm would never learn any other classes. In the few studies that do report accuracy in natural samples, our drop-off is similar (See Table 2 in [13]).

The classes being naturally highly imbalanced creates several model development choices, including whether to exclude, up-sample, or augment low incidence classes, and how specific model classifications should be (e.g., high level classes like diatoms, dinoflagellates, etc. or species level classes like *Thalassiosira* and *Gyrodinium*). We tried to strike a balance in our model setup by eliminating rare classes or merging them into broader groups while keeping groups morphologically distinct to prevent model confusion. However, there remains a degree of high intraclass variance and interclass similarity in morphology that was impossible to eliminate (e.g., 14.9% classification accuracy for "other"; Table 1). This challenge can be addressed on the other end of model development, by filtering samples where model uncertainty is high. The CNN outputs a confidence score (from the Softmax classification layer) for each prediction from 0 to 1 that can be used to filter samples below a certain threshold. While potentially increasing the model accuracy, this could also bias the system against certain classes that are challenging to classify, and thus was not implemented in this work.

Another potential cause of reduced model accuracy is data labelling errors. Theoretically, manual identification of images should be close to perfect, but unfortunately this is not the case. In this work and most others, there is often a bias for training and test data that is easily identifiable by manual validation, which prevents test metrics from translating exactly to the wild. There are also many images with conglomerations of cells including detritus and multiple living species. While these may be manually sorted into a category labelled "multiple" and discarded from the analysis, a CNN might sort these images into the most prominent class present within each image. Additionally, morphologically ambiguous cells may be sorted more accurately by a CNN than by manual identification, as a CNN can mathematically match image attributes to potential groups. One way we attempted to eliminate a portion of these ambiguous cells was to exclude all cells with a major axis length less than 25 pixels (7.35 µm) prior to model training. These small cells are below the quantifiable limit of detection based on instrument resolution, and thus have a high probability of being incorrectly identified. Accurately classifying these smaller cells will likely require techniques other than imaging. The issues of class imbalance can also magnify labelling errors, especially when these errors are within abundant classes such as "detritus".

B. Phytoplankton Seasonal Succession at Palmer Station

Like other studies, we found that following a winter with low sea ice (2017), the phytoplankton community had less diatoms, and more mixed flagellates and cryptophytes, and following a winter with high sea ice (2018), the community was dominated by diatoms (Figs. 1, 3, Table 2) [9], [17]. Following trends found in previous years at Palmer Station [9], we also saw diatoms dominate in the early and late season, and higher cryptophyte concentrations in December and January.

Along the WAP, phytoplankton show strong interannual and regional variability timed with light availability and sea ice retreat. As day length increases in austral spring, solar warming and sea ice melt stabilize the upper water column allowing phytoplankton to remain near the surface in waters with high light availability [18-19]. These conditions initiate large diatom-dominated spring blooms, as we saw in 2018 [20-21]. In 2017, there was 52% sea ice coverage in November, likely inhibiting light penetration and subsequent phytoplankton growth. Dramatic reduction in sea ice coverage between November and December indicates that the ice was rapidly advected out of the region, reducing sea ice melt near Palmer Station and potentially reducing the stability of the upper mixed layer. In 2018, although sea ice retreat is six days later than in 2017, November sea ice coverage is only 24%, allowing adequate light for phytoplankton growth. Additionally, the sea ice lingers into December and January (17% and 10%, respectively), providing a stable environment for growth well into the summer. Matching our results, [22] found that rapid sea ice retreat was associated with lower proportions of centric diatoms during the spring in Ryder Bay, Antarctica (Fig. 2). Sea ice can also hold populations of ice algae, which can seed coastal regions during melting in spring [23]. It is possible that with rapid sea ice advection from the

region in 2017, less ice algae were released to the coastal region near Palmer Station than in 2018 when sea ice lingered and contributed more meltwater.

Despite differences in phytoplankton abundance and community structure between the two years, there were similar trends in the diatom community. The early season was dominated by large centric diatoms > 20 µm timed with sea ice retreat as described above. As both seasons progressed, centric diatoms became smaller (< 20 µm), and the abundance of pennate diatoms < 10 µm increased (Fig. 3). A explanation for this size shift is the increasing amount of glacial meltwater from January to March [24]. Stronger surface stratification due to increased ice melt can reduce nutrients in surface waters, giving an advantage to smaller phytoplankton with high surface-area-to-volume ratios and reduced sinking rates [26]. Additionally, [25] experimentally exposed phytoplankton populations from Potter Cove, Antarctica to low salinity conditions (30 PSU) and found a decline in the abundance of large centric diatoms from ~90% on day 2 to ~0% on day 7, and an increase in abundance of small pennate diatoms from ~0% on day 4 to ~95% on day 8. They attribute these changes to differing osmotic stress tolerances: in large centric diatoms, a decrease in salinity caused cell size increases, compression of chloroplasts, granularization of the protoplasm, and retraction of the cytoplasm, while some small pennate diatoms (e.g., *Fragiliariopsis cylindrus*) may contain genes beneficial for adaptation to extreme environmental conditions in polar oceans and sea ice. Thus, increases in glacial meltwater in late summer could cause diatom communities to become smaller and increasingly dominated by pennate cells as we observed.

C. Conclusions and Next Steps

Our CNN is a step forward for understanding phytoplankton ecology along the WAP. However, there are still improvements to be made before it becomes a long-term tool for the community. As explained above, an important issue to address is class imbalance compounded with labelling errors of abundant classes. One potential way to better represent these undifferentiated classes (e.g., "detritus" or "multiple") is to use unsupervised methods (e.g., non-linear dimensionality reduction, clustering, and manifold learning) to break these classes into several new groups. Defining classes purely via data rather than taxonomy could help models with potentially more easily separable decision boundaries. These techniques could also reduce manually labeled training data needs with semi-supervised classification, and in many cases unsupervised techniques may be sufficient for answering questions about phytoplankton dynamics without any need for supervised classification [27]. Another method could be to use a stage-wise approach, with a one-class-classifier or binary classification to exclude "detritus" and "multiple" images up front to limit the spread of these issues into the full output range which is exacerbated by the prevalence of these classes. In tandem to improving the classification itself, per class uncertainty estimates (sensu [13]) will be critical to unbiased extrapolation from CNN output to ecological dynamics.

With further increases in model accuracy, we hope our model will become a helpful tool for phytoplankton research. Long-term warming and sea ice declines along the WAP are contributing to shifts to smaller and less abundant phytoplankton populations [6], and these trends are likely to continue. Understanding the seasonal and spatial dynamics of phytoplankton diversity is integral to contextualizing how communities will change in the future. Beyond the CNN's ability to rapidly classify entire seasons of collected phytoplankton imagery, it can also be used to characterize phytoplankton communities in near-real time. Getting a snapshot of species and cell size dynamics soon after collecting a sample would aid in opportunistic sampling while still in the field. This would be invaluable, as research time in Antarctica is both limited and expensive.

Lastly, the PAL-LTER is not the only group experiencing these challenges: there is a broad IFCB user community searching for methods to automate sample classification to reduce the need for manual image validation. Various groups are independently creating phytoplankton CNNs and other models for their study sites of interest. We implore the community to begin reporting their model metrics on data with distributions representative of the natural environment, sharing labeled data openly on freely accessible platforms (e.g., EcoTaxa, IFCB Dashboard), and sharing open and reproducible code for processing and model development. As models improve, the community may be able to develop a series of regional models, freely available to download and classify a worker's data, or even a single generalizable model usable for the world oceans. Moving forward towards this vision, it will be critical for oceanographers to collaborate with computer scientists and modelers, incorporating the best computer vision and classification techniques to these datasets to ultimately better understand phytoplankton dynamics in a changing ocean.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation Antarctic Organisms and Ecosystems Program (PLR-1440435) as part of the PAL-LTER program. Thank you to Alison Chase and Sasha Kramer for help with taxonomic identifications, and to Emmett Culhane for helpful discussions regarding the challenges of building a CNN for IFCB data. This work would not be possible without the PAL-LTER field teams who aided in data collection and the Palmer Station and Laurence M. Gould personnel who provided logistical support.

REFERENCES

[1] H. W. Ducklow *et al.*, "West Antarctic Peninsula: An ice-dependent coastal marine ecosystem in transition," *Oceanography*, vol. 26, no. 3, pp. 190–203, 2013.

[2] A. J. Cook *et al.*, "Ocean forcing of glacier retreat in the western Antarctic Peninsula," *Science*, vol. 353, no. 6296, pp. 283–285, 2016.

[3] M. P. Meredith and J. C. King, "Rapid climate change in the ocean west of the Antarctic Peninsula during the second half of the 20th century," *Geophys. Res. Lett.*, vol. 32, L19604, 2005.

[4] J. Turner *et al.*, "Antarctic climate change during the last 50 years," *Int. J. Climatol.*, vol. 25, pp. 279–294, 2005.

[5] S. Stammerjohn, R. Massom, D. Rind, and D. Martinson, "Regions of rapid sea ice change: An inter-hemispheric seasonal comparison," *Geophys. Res. Lett.*, vol. 39, L06501, 2012.

[6] M. Montes-Hugo *et al.*, "Recent changes in phytoplankton communities associated with rapid regional climate change along the Western Antarctic Peninsula," *Science*, vol. 323, pp. 1470–1473, 2009.

[7] M. A. Moline, H. Claustre, T. K. Frazer, O. Schofield, and M. Vernet,

"Alteration of the food web along the Antarctic Peninsula in response to a regional warming trend," *Glob. Chang. Biol.*, vol. 10, pp. 1973–1980, 2004.

[8] C. R. B. Mendes *et al.*, "Shifts in the dominance between diatoms and cryptophytes during three late summers in the Bransfield Strait (Antarctic Peninsula)," *Polar Biol.*, vol. 36, pp. 537–547, 2013.

[9] O. Schofield *et al.*, "Decadal variability in coastal phytoplankton community composition in a changing West Antarctic Peninsula," *Deep Sea Res. Part I*, vol. 124, pp. 42–54, 2017.

[10] S. F. Sailley *et al.*, "Carbon fluxes and pelagic ecosystem dynamics near two western Antarctic Peninsula Adélie penguin colonies: an inverse model approach," *Mar. Ecol. Prog. Ser.*, vol. 492, pp. 253–272, 2013.

[11] W. A. Kozlowski, D. Deutschman, I. Garibotti, C. Trees, and M. Vernet, "An evaluation of the application of CHEMTAX to Antarctic coastal pigment data," *Deep. Res. Part I*, vol. 58, pp. 350–364, 2011.

[12] R. J. Olson and H. M. Sosik, "A submersible imaging-in-flow instrument to analyze nano- and microplankton: Imaging FlowCytobot," *Limnol. Oceanogr. Methods*, vol. 5, pp. 195–203, 2007.

[13] H. M. Sosik and R. J. Olson, "Automated taxonomic classification of phytoplankton sampled with imaging-in-flow cytometry," *Limnol. Oceanogr. Methods*, vol. 5, no. 6, pp. 204–216, 2007.

[14] M. Picheral, C. E. Sanders, and J-O. Irisson, "EcoTaxa, a tool for the taxonomic classification of images," *http://ecotaxa.obs-vlfr.fr*, 2017.

[15] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, 2015.

[16] S. E. Stammerjohn, D. G. Martinson, R. C. Smith, and R. A. Iannuzzi, "Sea ice in the western Antarctic Peninsula region: Spatio-temporal variability from ecological and climate change perspectives," *Deep. Res. Part II*, vol. 55, pp. 2041–2058, 2008.

[17] G. K. Saba *et al.*, "Winter and spring controls on the summer food web of the coastal West Antarctic Peninsula.," *Nat. Commun.*, vol. 5, 4318, 2014.

[18] M. Vernet *et al.*, "Primary production within the sea-ice zone west of the Antarctic Peninsula: I- Sea ice, summer mixed layer, and irradiance," *Deep. Res. Part II*, vol. 55, pp. 2068–2085, 2008.

[19] H. J. Venables, A. Clarke, and M. P. Meredith, "Wintertime controls on summer stratification and productivity at the western Antarctic Peninsula," *Limnol. Ocean.*, vol. 58, no. 3, pp. 1035–1047, 2013.

[20] G. B. Mitchell and O. Holm-Hansen, "Bio-optical properties of Antarctic Peninsula waters: differentiation from temperate ocean models," *Deep Sea Res.*, vol. 38, no. 8–9, pp. 1009–1028, 1991.

[21] B. B. Prézelin, E. E. Hofmann, C. Mengelt, and J. M. Klinck, "The linkage between Upper Circumpolar Deep Water (UCDW) and phytoplankton assemblages on the west Antarctic Peninsula continental shelf," *J. Mar. Res.*, vol. 58, pp. 165–202, 2000.

[22] A. L. Annett, D. S. Carson, X. Crosta, A. Clarke, and R. S. Ganeshram, "Seasonal progression of diatom assemblages in surface waters of Ryder Bay, Antarctica," *Polar Biol.*, vol. 33, pp. 13–29, 2010.

[23] S. F. Ackley and C. W. Sullivan, "Physical controls on the development and characteristics of Antarctic sea ice biological communities - a review and synthesis," *Deep. Res. I*, vol. 41, no. 10, pp. 1583–1604, 1994.

[24] M. P. Meredith *et al.*, "Local- and large-scale drivers of variability in the coastal freshwater budget of the Western Antarctic Peninsula," *J. Geophys. Res. Ocean.*, vol. 126, e2021JC017172, 2021.

[25] M. Hernando *et al.*, "Effects of salinity changes on coastal Antarctic phytoplankton physiology and assemblage composition," *J. Exp. Mar. Bio. Ecol.*, vol. 466, pp. 110–119, 2015.

[26] W. K. W. Li, F. A. McLaughlin, C. Lovejoy, and E. C. Carmack, "Smallest algae thrive as the Arctic ocean freshens," *Science*, vol. 326, p. 539, 2009.

[27] E. Culhane, N. Haentjens, A. P. Chase, P. Gaube, and J. Morrill, "Leveraging unsupervised methods to train image classification networks with fewer labelled inputs: Application to species classification of phytoplankton imagery from an Imaging Flow Cytometer," in *Ocean Sciences Meeting*, 2020.