COMBINING HIGH-THROUGHPUT IMAGING AND AMPLICON SEQUENCING
TO MONITOR EUKARYOTIC PLANKTON

by

Liam MacNeil

Submitted in partial fulfilment of the requirements
for the degree of Master of Science

at

Dalhousie University
Halifax, Nova Scotia
June 2021

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Microbial communities support ocean food webs and respond to the surrounding environment to varying degrees across different time scales. The eukaryotic plankton throughout the oceans are extraordinarily diverse but difficult to monitor using conventional tools. A next generation of ocean observations are possible but remain unrealized to monitor eukaryotic plankton directly from the ocean using high-throughput measurements. In this thesis, I apply digital holography and amplicon sequencing to describe diverse community compositions of micro and mesoplankton. First, I evaluate automatic classification of micro-mesoplankton from seawater and monocultures using a deployable digital in-line holographic microscope and state-of-the-art classification algorithms. Second, I quantify and barcode the micro-mesoplankton community across transects of the Newfoundland Shelf. These results confirm digital in-line holographic microscopes can yield rapid, high-quality plankton images under multiple *in-situ* conditions, that benchmark image recognition tools are highly transferrable to plankton images, and that paired high-throughput amplicon sequencing yields different, although complementary surveys.

LIST OF ABBREVIATINS AND SYMBOLS USED

ASV	Amplicon Sequence Variants
AUC	Area Under the Curve
AUC-PR	AUC Precision-Recall
AZMP	Atlantic Zone Monitoring Program
C_b	Total phytoplankton carbon biomass
C_c	Carbon (pg) per cell
clr	Centered log-ratio
CMOS	Complementary Metal Oxide Semiconductor
CNN	Convolutional Neural Network
CTD	Conductivity-Temperature-Depth system
°C	Degrees Centigrade
DNA	Deoxyribonucleic acid
°N	Degrees North (Latitude)
°W	Degrees West (Longitude)
DBSCAN	Density Based Spatial Clustering with Applications of Noise
DIHM	Digital In-line Holographic Microscope
eDNA	Environmental DNA
GPU	Graphical Processing Unit
QIIME2	Quantitative Insights Into Microbial Ecology
18S rRNA	Eukaryotic small-subunit ribosomal Ribonucleic nucleic acid
TEOS	Thermodynamic Equation of Seawater

ACKNOWLEDGEMENTS

"In life, unlike chess, the game continues after checkmate."

-Isaac Asimov

CHAPTER 1 — INTRODUCTION

1.1. Monitoring Marine Plankton

The sunlit surface ocean (0-200 m) contains a vast array of environments where photosynthetic autotrophs (i.e., phytoplankton) can thrive. Phytoplankton growth is controlled by the delivery of inorganic nutrients to the surface through upwelling, advection, atmospheric inputs and into higher latitudes—in the subpolar and polar regions— the seasonality of light plays an increasingly strong role (Behrenfeld *et al.* 2006; Behrenfeld and Boss, 2014). Under specific environmental conditions, when phytoplankton growth outpaces losses, from top-down (e.g., predation, viral lysis, mixing) and bottom-up (e.g., cellular respiration) processes, blooming events can occur (Behrenfeld and Boss, 2014). The accumulated biomass drives carbon export to the deeper ocean through sinking (Ducklow *et al.* 2001) and predation by microbial consumers (i.e., heterotrophs) links biomass to higher trophic levels (Sherr and Sherr, 1988). Harmful algal blooms can also occur, notably from several groups of dinoflagellates (e.g., *Alexandium, Dinophysis, Gymnodinium*), bringing historically detrimental ecosystem impacts and economic consequences (Hallegraeff, 2010).

Monitoring efforts of planktonic (i.e., drifting) organisms are generating *in-situ* observations of plankton communities to explain and eventually predict their role in primary production and the sequestration of carbon, the cycling of major oceanic elements (e.g., Si, N, P, Fe), the fate of larval-stage fish species, and their response to oceanic climate change (Lombard *et al.* 2019). Observations made *in-situ* will also support more accurate global biogeochemical models for forecasting environmental and

1

biological changes that will shape the oceans in coming decades (Centurioni *et al.* 2019).

Some plankton groups do not simply drift, but actively move throughout their environment (i.e., nektonic), and *in-situ* sample technologies are equally well adapted to sample these groups. Monitoring requires quantitative observations to distinguish in absolute terms, the plankton compositions of different water masses. Remote sensing using satellites have revolutionized quantitative observations of ocean color across the global ocean (Dutkiewicz *et al.* 2020), but they require *in-situ* ground truthing, cannot render detailed taxonomic information, and capture mostly surface layer information in a vertically structured ocean (Dierssen, 2010). Numerous *in-situ* methods exist for quantitative analyses, including pigment markers using chlorophyll fluorescence to infer total photosynthetic biomass (Sauzède *et al.* 2015), and high-performance liquid chromatography to quantify the community structure based on group-specific pigments, including other photosynthetic pigments such as chlorophyll-b/c, carotenoids, and phycobillins (Wright and Jeffery, 2006). Each of these methods, however, quantifies photosynthetic groups and excludes the heterotrophic component, which can be substantial (Duarte *et al.* 2013), and many photoautotrophs regulate pigment concentrations non-linearly in response to temperature, light, and nutrient conditions (e.g., Geider *et al.* 1997), implying the relationship between pigment and cell concentrations is not constant. Although less widely adopted, *in-situ* imaging instruments are steadily being deployed because they can rapidly capture living and non-living particles suspended in the natural environment, along with valuable taxonomic and quantitative information (Benfield *et al.* 2007).

1.2 Plankton Imaging

In microbial ecology, the shape, size, and symmetry of many planktonic eukaryotes contain information about taxonomic identity (Tomas *et al.* 1997) and physiological state (e.g., Irwin *et al.* 2006; Finkel *et al.* 2010). At a community level, the distribution of cell sizes and densities are important dimensions of community structure, and their dynamics provide powerful insights into mechanisms that select for particular groups and size classes (e.g., Fowler *et al.* 2020). Size and morphology also contain information about the pathways of non-living particulates through the planktonic food webs (Trudnowska *et al.* 2021). Digital cameras allow this physical information to be captured and archived for independent verification, further analysis, and reproducibility (Lombard *et al.* 2019). Numerous plankton-focused instruments exist (Table 1.1), each tailored to interrogate a volume of water for objects of a limited size spectra. Although digital cameras are less integrated into ocean sciences as many molecular methods, efforts for global sampling coverage are increasing (Biard *et al.* 2016; Lombard *et al.* 2019). Digital images are arrays of pixels with or without dimensions of color (i.e., RGB). Pixel intensity is the primary image property that defines object size, cross-sectional area, and shape. Depending on image resolution (i.e., pixels per distance), there is a familiar trade-off between object size and throughput: High resolution images usually narrow the acceptable field-of-view, and slow sampling flow rates to afford sharp, focused object detection (Lombard *et al.* 2019).

Table 1.1. Plankton imaging cameras. The chosen image properties derive from reviews in Lombard et al. (2019), Nayak et al. (2021), and original sources: FlowCam (Sieracki et al. 1998), Flow Cytobot (Olson and Sosik, 2007), UVP5 (Picheral et al. 2010), ISIIS (Cowen and Guigand, 2008), Scripps Plankton Camera ([a]Orenstein et al. 2020), Prince William (PW) Sound Plankton camera (Campbell et al. 2020

Camera	Mode	Weight (kg in air)	Deployment	Depth	Total Size-Range (ESD)	Sampling Volume	Max. Cost
HoloSeaS5	Holography	2.6 kg	On Board / Laboratory/ Cast	6000 m	20 - 2000 μm	150 mL min^{-1}	$50,000
FlowCAM Objectives: 2X 4X 10X 20X	Optical	> 22 kg	On Board / Laboratory	-	75-1000 μm 20-300 μm 10-100 μm 2-50 μm	5 mL min^{-1}	$92,500
FlowCytobot	Optical	32 kg	On Board / Laboratory	-	50-150 μm	0.25 mL min^{-1}	$ 158,000
UVP-5	Optical	30 kg	Towed / Cast	3000 m	60-20,000 μm	20-40 L m^{-1}	$39,500
ISIIS Camera	Shadowgraph	NA	Towed	200 m	60-130,000 μm	150 L s^{-1}	$100,000
LISST-Holo2	Holography	10.4 kg		3000 m	25-2500 μm	1800 mL min^{-1}	92,000 €
Scripps Cameras: Micro Mini Macro	Optical	30 kg	Moored	Surface	10-1000 μm 100-5000 μm 500-10000 μm	24 μL s^{-1} 24 mL s^{-1} 4 L s^{-1}	$50,000
PW Sound Camera	Optical	10 kg	Cast	60m	500-10000 μm	3.1 L s^{-1}	NA

1.3 Digital Holographic Microscopy

Holographic microscopes offer a different imaging configuration which enhances the depth of volume that can be acceptably focused (i.e., depth-of-field) without losing image resolution (Jericho and Kreuzer, 2011). Digital holographic microscopes can achieve enhanced depth-of-field using the numerical reconstruction of a wavefront from a coherent light source (i.e., a laser) without an objective lens (Jericho and Kreuzer, 2011). The in-line configuration of holographic microscopes uses a single coherent light source (i.e., point-source) that permits simple, durable, and cost-effective models for *in-situ* deployment (Xu *et al.* 2001; Garcia-Sucerquia *et al.* 2006). The depth-of-field improvement allows in-line microscopes with in-flow designs to effectively image larger volumes for high-throughput sampling, while simultaneously able to reconstruct pixel intensity, amplitude, and phase shift at micrometer scales, recovering information about an object size, shape, and 3-D properties (i.e., refractive index) (Kanka *et al.* 2009; Jericho *et al.* 2012). Together, these traits make holography especially applicable for sampling plankton and particulates in their environment (Nayak *et al.* 2021). Like any new instrumentation, deployment onto ships-of-opportunity for vertical profiles or surface water inflow requires intensive logistical support to monitor data quality, biofouling, and manage the rapid data uptake. If a holographic microscope, or several other digital imaging cameras can be supported to autonomously collect *in-situ* samples, large sums of data can be generated, and data interpretation becomes an equally challenging task.

1.4 Data-Driven Image Classifiers

The rapid access and accumulation of digital observations, especially images, is a common feature in the Natural sciences that totals hundreds of petabytes, requiring transmission rates at hundreds of terabytes per day, and growing quasi-exponentially (Reichstein *et al.* 2019). Data-driven problems have supported the development of sophisticated models for pattern recognition, enabled by Graphical Processing Units (GPUs) that accelerate model training and allow computation-intensive tasks to scale (Strigl *et al.* 2010). For images, features can be encoded numerically through collections of pixels, from which feature detection algorithms can extract, condense, and integrate into hierarchical feature descriptors (Krizhevsky *et al.* 2012; van Noord and Postma, 2017). To date, Convolutional Neural Networks (CNNs), commonly described by its neologism "Deep Learning", have proven to be among the highest performing and most generalizable plankton classifiers (Orenstein and Beijbom, 2017; González *et al.* 2019). Deep learning is an outbranch of machine learning, which has proliferated across most scientific fields due to its seemingly intelligent, task-specific ability to learn patterns across data types (LeCun *et al.* 2015). The learning process can also combine previously learned patterns, leading to empirically robust algorithms that can be repurposed for new problems (Yosinski *et al.* 2014). Many off-the-shelf classifiers have been developed this way, but there is a serious lack in available holographic imagery to enable transferability for holographic plankton images, which are quickly accumulating (Nayak *et al.* 2021). Cultivating a holographic image dataset of labelled plankton objects will be invaluable to support generalizable classifiers that will allow data interpretation to scale and accelerate the information-harvesting from monitoring projects.

1.5 Environmental Metabarcoding

Plankton have other diagnostic traits—also evolutionarily earned—beyond morphology: Environmental DNA (eDNA) can be targeted for specific marker genes that catalogue major clades in the web of life, regardless of size (Hugerth and Andersson, 2017). Assessing community diversity at a molecular level is advantageous because it provides superior taxonomic and phylogenetic resolution to most other methods (Deiner *et al.* 2017). The first published high-throughput sequencing analysis in the marine environment Sogin *et al.* (2006) indicated two distinct deep-sea environments contained nearly two orders of magnitude more bacterial species than previously estimated. Since then, increasingly paralleled sequencing technologies have supported global sampling efforts, including the Tara Oceans project (Bork *et al.* 2015), revealing basin-scale patterns for microbial community composition and structure (e.g., de Vargas *et al.* 2015; Lima-Mendez *et al.* 2015; Faure *et al.* 2019; Gregory *et al.* 2019). The rich sampling depth of high-throughput sequencing has technical costs: The number of biological sequences (i.e., reads) are arbitrarily constrained by the total number of reads that the sequencing instrument produces during a run. The resulting data are fundamentally relative — i.e., compositional (Gloor *et al.* 2017). Their proportions still have meaning, but valid interpretations require statistical methods that satisfy the compositional constraints (Aitchison, 1985). Additionally, eDNA methods are generally applied to bulk DNA extracted from a whole community present in a water sample—which is a much different mode of community survey than imaging methods, which usually focus on intact, living cells (Jian *et al.* 2021).

1.6 Broadening Microbial Community Surveys

Reconciling high-throughput imaging and sequencing, despite their individual strengths, is not clear. Whether they converge on similar patterns observed in the ocean, across geographic boundaries (e.g., hydrographic fronts) for example (Raes *et al.* 2018), is an open question. Clear comparisons between imaging and molecular techniques are needed from real environmental samples. From imaging instruments, community structure can be described with quantitative observations of cell densities and sizes, within a defined size spectrum, while the marker genes can validate taxonomy and indicate what fraction of the community was missed, if any, from the images. Each method has strengths and limitations, and no method yet presented can cover the entire range of planktonic communities, but in combination, imaging and eDNA can reveal community structure and composition, respectively. This thesis addresses several challenges in community surveys: First, developing a labelled image database to train several classification algorithms targeting marine plankton, allowing more efficient interpretations of future field deployments. Second, to quantify cell density and biovolume for total biomass using the digital images from a spatial transect and interpret community distributions across an environmental gradient. Third, combining eDNA and imaging to confirm taxonomic composition of the spatial dataset, and compile a larger, more finely annotated dataset, for future use.

1.7 Thesis Structure

This thesis is structured into two distinct data chapters. Both chapters analyze community composition with a digital in-line holographic microscope to classify and quantify micro-mesoplankton. The overarching goals were to improve monitoring efforts of the micro-mesoplankton through better image classification algorithms, and to combine imaging and marker gene analysis for *in-situ* plankton distributions.

Chapter 1 focuses on imaging analysis; the pronouns are in third person to acknowledge my co-authors. This chapter uses hologram reconstruction and object detection software on taxonomically identified cultures and environmental samples imaged from bulk volumes. Greater detail of numerical reconstruction and downstream hologram processing are provided in Chapter 2. Plankton were identified manually, and a set of deep learning algorithms were developed to automatically classify objects into 19 groups. This chapter's goal is a proof-of-principle: The object detection pipeline can produce sharp, focused plankton images from a range of groups, and many can be reliably classified automatically with a small number of training examples.

In Chapter 3, paired samples for imaging and DNA filtration were taken onboard a 2019 research cruise and compared for community structure and composition. Quantitative observations are derived from imaging samples, including phytoplankton biomass from basic volume-to-carbon scaling laws. The resulting community analysis is supported by high-frequency environmental data collected in tandem. Together, the biological and environmental observations present latitudinal and cross-shelf patterns in community composition and water column structure.

Finally, in Chapter 4, the overall findings are discussed. This chapter contains possibilities for improving quantitative imaging profiles with digital holographic microscopes, the importance of choosing careful deployment practices, and adequate tools for technical and computational requirements. A broad outlook for combining imaging and molecular methods in a next generation of ocean observations is also discussed.

CHAPTER 2 — PLANKTON CLASSIFICATION WITH HIGH-THROUGHPUT SUBMERSIBLE HOLOGRAPHIC MICROSCOPY AND TRANSFER LEARNING[1]

2.1 Abstract

Plankton are foundational to marine food webs and an important feature for characterizing ocean health. Recent developments in quantitative imaging devices provide in-flow high-throughput sampling from bulk volumes— opening new ecological challenges exploring plankton variation and diversity, alongside technical hurdles to automate classification from large datasets. However, a limited number of deployable imaging instruments have been coupled with the most prominent classification algorithms— effectively limiting the extraction of curated observations from field deployments. Holography offers relatively simple coherent microscopy designs with non-intrusive 3-D image information, and rapid frame rates that support data-driven plankton imaging tasks. Classification benchmarks across different domains have been set with transfer learning approaches, focused on repurposing pre-trained, state-of-the-art deep learning models as classifiers to learn new image features without protracted model training times. Combining the data production of holography, digital image processing, and computer vision could improve *in-situ* monitoring of plankton communities and contribute to sampling the diversity of the plankton. Here we use a light and portable digital in-line holographic microscope (The HoloSea) with maximum optical resolution of 1.5 μm, intensity-based object detection through a volume, and four different pre-

trained convolutional neural networks to classify > 3800 micro-mesoplankton (> 20 μm)

images across 19 classes. The maximum classifier performance was quickly achieved for

each convolutional neural network during training and reached F1-scores > 89%. Taking

classification further, we show that off-the-shelf classifiers perform strongly across every

decision threshold for ranking a majority of the plankton classes. These results show

compelling baselines for classifying holographic plankton images, both rare and plentiful,

including several dinoflagellate and diatom groups. These results also support a broader

potential for deployable holographic microscopes to sample diverse microbial eukaryotic

communities, and its use for high-throughput plankton monitoring.

2.2 Background

Plankton are an integral component of the global ocean. Plankton abundance and

composition can be coupled to environmental conditions and yield important insights into

aquatic food webs (e.g., Hays *et al.* 2005; Irwin *et al.* 2015). Often hugely diverse and

occupying numerous trophic modes in surface ocean ecosystems, classifying plankton

composed in a water mass is challenging, error prone, and a bottleneck of time and costs.

Recent developments in imaging instruments allow biological contents to be visualized

directly from bulk volumes at high image resolution, without disintegrating cell structures

(Benfield *et al.* 2007). Imaging instruments have used a variety of optical methods

including flow cytometry (Olson and Sosik, 2007), shadowgraphs (Cowen and Guigand,

2008), holography (Garcia-Sucerquia *et al.* 2006), among others. Several such devices

have imaged plankton size classes that collectively encompass autotrophs and

heterotrophs, spanning four orders of magnitude in size from 2 μm-10 cm (Benfield *et al.*

2007; Zetsche *et al.* 2014; Lombard *et al.* 2019). The high sampling frequency from digital imaging also opens new ecological challenges exploring microbial eukaryotic diversity (Colin *et al.* 2017), alongside technical challenges to automate classification from spatial and temporally dense datasets (e.g., Greer *et al.* 2015; Biard *et al.* 2016).

Digital holography is based on the diffracted light field created by interference from objects in a sample which is illuminated by a coherent light (e.g., a laser): That interference pattern is recorded by a digital sensor and composes a hologram (Schnars and Jüptner, 2002). Since their inception (Gabor, 1948), holographic microscopes have been applied widely at micrometre scales to observe, for example, particle distributions (Sheng *et al.* 2006), coral mucus production (Zetsche *et al.* 2016), and to differentiate cancerous pancreatic cells from healthy ones (Kemper *et al.* 2006). Holographic microscopes have advanced considerably with improving computational techniques for digital reconstruction and focus enhancement (Rivenson *et al.* 2019). Digital in-line holographic microscopy (DIHM) with a point-source laser is a simple, lens-free implementation of Gabor-style holography that can capture a 3-D sample using a common path optical configuration, whereby both reference and interfered light waves copropagate and are recorded by a digital camera (Kreuzer and Jericho, 2007). DIHM has several advantages for biological studies including a simple design with a larger depth of field than conventional light microscopy, allowing rapid imaging of larger volumes and 3-D numerical refocusing with no required staining of cells (Xu *et al.* 2001; Jericho and Kreuzer, 2011). Due to its simplicity, DIHM can easily be incorporated into various cell imaging configurations including amplitude and phase images (Jericho *et al.* 2012) and to date, numerous studies have used holography to image marine plankton (Hobson *et al.*

1997; Malkiel *et al.* 1999; Sun *et al.* 2008; Zetsche *et al.* 2014; Rotermund *et al.* 2015). There is increasing interest to use its advantages towards automating classification of plankton and particulates from water samples (e.g., Göröcs *et al.* 2018; Bianco *et al.* 2020; Guo *et al.* 2021). A review of holographic microscopes for aquatic imaging can be found in Nayak *et al.* (2021).

Plankton exhibit substantial morphological variation within and between major groups, are often imaged at different orientations, appear partially occluded, or damaged. Extracting features from plankton images originally relied on handcrafted feature descriptors, which are label-free and train classifiers like support vector machines or random forest efficiently (Gorsky *et al.* 2010). But detecting features based on predefined traits rapidly reaches its limits. Instead, deep learning algorithms have gained popularity for their state-of the-art performance and, at least in part, because they require no domain specific knowledge or impose descriptors for pattern recognition, rather features are learned during training (LeCun *et al.* 2015). Deep learning involves representing features at increasing levels of abstraction and for image tasks, the most successful models have been convolutional neural networks (CNNs): A layered neural network architecture, with layers equating to depth, and where convolutions substitute as feature extractors (Schmidhuber, 2015). These CNNs learn features through sequential layers connected to the local receptive field of the previous layer and the weights learned by each kernel (Schmidhuber, 2015). For plankton, CNNs have improved the classification stage of automation efforts (Dai *et al.* 2016; Luo *et al.* 2018). But the natural imbalance in plankton datasets and frequent drifts in class distributions ([a]González *et al.* 2017) render

accuracy benchmarks for performance biased towards majority classes and poor evaluation metrics (e.g., Corrêa *et al.* 2017; Dunker *et al.* 2018).

Achieving state-of-the-art classification at scale often requires large training datasets for CNNs, but generic features can be extracted from pre-trained models and repurposed— termed transfer learning— such that CNNs have a baseline that can recognize features unspecific to any image, similar to Gabor filters or color blobs (Yosinski *et al.* 2014). Transfer learning has achieved classification benchmarks equivalent to traditional feature descriptors (e.g., Sharif Razavian *et al.* 2014). Large plankton image datasets do exist— some containing several million labelled images across hundreds of classes (e.g., Orenstein *et al.* 2015)— but there is a current lack of easily deployable plankton imaging devices capable of rapidly sampling several litres. Holographic microscopy combined with computer vision, could bridge high throughput *in-situ* data production with increasingly automated classification and enumeration of major plankton groups.

The purpose of this study is to show whether species of micro-mesoplankton can be detected in-focus from volumetric samples, classified with deep learning algorithms, and to evaluate classifiers with threshold-independent metrics— which, to date, are rarely considered for imbalanced plankton classification tasks.

2.3 Methods

2.3.1 The HoloSea: Submersible Digital In-Line Holographic Microscope (DIHM)

General DIHM designs for biological applications are reviewed in Garcia-Sucerquia *et al.* (2006) and Xu *et al.* (2001). A similar submersible DIHM, the 4-Deep

HoloSea S5[2] (92×351 mm, 2.6 kg), first introduced by Walcutt *et al.* (2020), was used

here to image plankton cells. Its principal advantage is a simple lensless, in-flow

configuration with 0.1 mL per frame and high frame rates ($> 20 \text{ s}^{-1}$) that support a

maximum flow rate $> 130 \text{ mL min}^{-1}$. Housed in an aluminum alloy casing, the HoloSea

uses a solid-state laser (405 nm) coupled to a single mode fiber optic cable acting as a

point source to emit spherical light waves through a sapphire window. As light waves

travel through the sampled volume, both the waves scattered by objects and reference

waves copropagate until they interfere at the plane of the monochrome camera sensor

(7.4 μm pixel) to form an interference pattern (i.e., a 2048×2048 hologram). The camera

is aligned 54mm away from the point source and recorded holograms are stored as PNG

images for further numerical reconstruction and analyses.

2.3.2 Numerical Hologram Reconstructions

Hologram reconstruction from point-source holography was first proposed by

Kreuzer *et al.* (1992), and its principles are well described (Xu *et al.* 2001; Jericho and

Kreuzer, 2011; Jericho *et al.* 2012). The workflow from reconstruction to object focusing

are shown in Figure 2.1. In order to recover the information about objects within

holograms at the specific focal distance from the point source, wave front intensity was

digitally reconstructed based on a Helmholtz-Kirchhoff transformation (Kanka *et al.*

2009) in 4-Deep Octopus software[3]. Each hologram was reconstructed at multiple z-

distances from the point source using a 50 μm step size through the sample volume. To

---

[2] http://4-deep.com/
[3] http://4-deep.com/products/octopus-software/

detect regions of interest (ROIs) in each reconstructed plane, we used 4-Deep Stingray software[4] with a globally adaptive threshold algorithm based on Otsu (1979). During the detection step, ROIs could also be discriminated based on their size; for our purposes, we defined a range of two orders of magnitude (20-2000 μm) to encompass micro-mesoplankton. Detected ROIs were clustered together across multiple z-planes based on the Euclidian distances between their centroids using the Density Based Spatial Clustering with Applications of Noise (DBSCAN) algorithm (Ester *et al.* 1996). Each resultant cluster contained the same ROI tracked at multiple consecutive z-planes within the volume. To identify the plane containing an in-focus object within each cluster, we used Vollath's F4 autocorrelative algorithm (Vollath, 1987)— the object with the highest correlation score between pixels was then stored in our database and the rest of objects within the cluster were discarded.
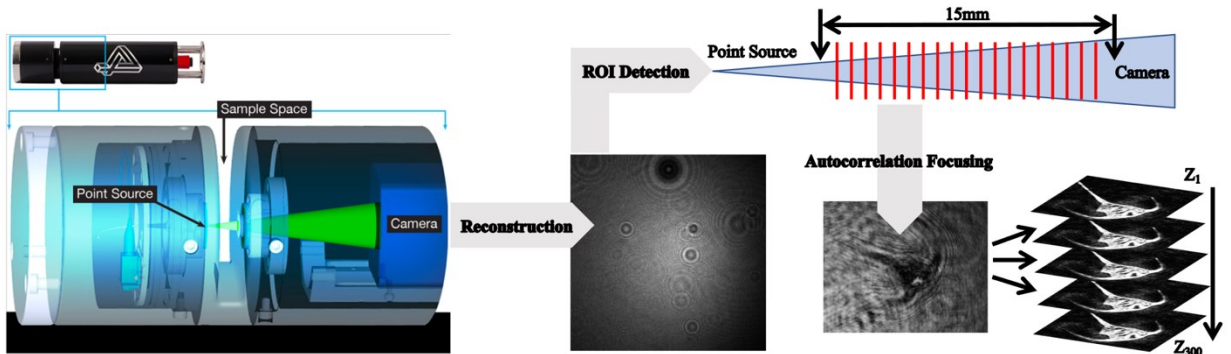
Figure 2.1. The workflow for imaging, detecting, and selecting in-focus objects. Volumes are recorded in the microscopes sample space and the interference pattern is reconstructed to create a hologram. Plankton objects are first detected as ROIs across 300 reconstructed planes (i.e., z-distances) of a hologram corresponding to the 15mm sample space. The plane containing an in-focus object is calculated via autocorrelation and Vollath's F4 algorithm.

2.3.3 Holographic Image Dataset

The plankton for our experiments (Table 2.1) included monocultures grown in artificial seawater and 500mL surface (1 m) water samples from Bedford Basin compass buoy station (44° 41'37" N, 63° 38'25" W). Monoculture samples were grown under f/2 nutrient replete and the recommended temperature and light conditions (Bigelow Laboratories, Maine, USA). Samples were pumped through the sample chamber using a peristaltic pump and recorded at 10fps. The resulting image dataset was augmented by rotating each image horizontally, vertically, and translated to enlarge the number of training images, and hence the learnable features threefold (Krizhevsky, 2012). All images were scaled to 128×128 pixels preserving the aspect ratio of the source images. Classes were randomly split approximately 50:10:40 for training, validation, and testing, respectively. Training and validation samples were divided into five stratified k-folds, where each fold retains the proportion of classes in the original training set (Pedregosa *et al.* 2011). We included a "noise" class to filter holographic artefacts (Garcia-Sucerquia *et al.* 2005).

Table 2.1. Taxa identity, size ranges, and total number of images. Cell sizes are taken from apical cell length measurements, using 25 examples for each class.

Class	Taxonomic Group	Size (μm)	Strain	Examples
Alexandrium tamarense	Dinoflagellate	20-80	CCMP1771	201
Ceratium fusus	Dinoflagellate	50-350	Environmental	56
Ceratium lineatum	Dinoflagellate	80-230	Environmental	44
Ceratium longpipes	Dinoflagellate	200-340	CCMP1770	378
Ceratium sp.	Dinoflagellate	140-230	Environmental	64
Chaetoceros socialis	Diatom	40-360	CCMP3263	102
Chaetoceros straight	Diatom	30-120	CCMP215	325
Chaetoceros sp.	Diatom	30-430	CCMP1690	114
Crustacean	Animal	180-640	Environmental	13
Dictyocha speculum	Silicoflagellate	30-105	CCMP1381	185
Melosira octagona	Diatom	80-460	CCMP483	173
Noise	Artefact	—	—	150
Parvicorbicula socialis	Choanoflagellate	25-85	Environmental	36
Prorocentrum micans	Dinoflagellate	30-120	CCMP688	1074
Pseudo-nitchzia arctica	Diatom	35-150	CCMP1309	33
Rhizosoenia setigera	Diatom	200-530	CCMP1330	306
Rods	Morphological	60-280	—	396
Skeletonema costatum	Diatom	60-130	CCMP2092	157
Tintinnid	Ciliate	90-310	Environmental	20

2.3.4 Convolutional Neural Networks (CNNs)

The plankton detected in our holograms were classified with four different CNNs: VGG16 (Simonyan and Zisserman, 2014), InceptionV3 (Szegedy *et al.* 2015), ResNet50V2 (He *et al.* 2016), and Xception (Chollet, 2017). In terms of model depth, VGG16 is the shallowest, InceptionV3 and ResNet50 are near equal, while Xception is the deepest. Each uses convolutions as feature extractors but with different model architecture (See Table A.1 in appendix). Due to the modest size of our plankton dataset, we used a transfer learning approach where each model was pre-trained on ~ 1.4 M images binned into over 1,000 classes from the ImageNet dataset (Deng *et al.* 2009). Pre-trained models have already learned generalizable features from the ImageNet dataset—

20

which includes animals, sports objects, computers, and other classes very different from plankton— that provides a powerful baseline for feature recognition (Yosinski *et al.* 2014). Classification was implemented in the Python deep learning toolbox Keras (Chollet, 2015), which is accessible as a core component of the Tensorflow package (Abadi *et al.* 2016).

Each model was applied in two different ways, first as a feature extractor by only retraining the deepest model layers to preserve the pre-tuned weights (Yosinski *et al.* 2014), and secondly by maintaining the first 10-20 layers and retraining the remaining layers. The second method was exploratory and involved freezing the first 10 layers in VGG16, and the first 20 layers for the other deeper models, which have presumably already learned generic features. We used dropout for each method at a probability of 0.3 to prevent overfitting (Srivastava *et al.* 2014) and added a Softmax classifier to transform the fully connected vector into a probability distribution specific to 19 classes (Janocha and Czarnecki, 2017). Our images were preprocessed according to each CNNs requirements (Chollet, 2015), and the greyscale color channel was repeated for each colored channel (i.e., RGB) that the models observed from ImageNet.

Prediction bias from our class imbalances, where the most abundant class was nearly three times greater than the least abundant, was offset by maintaining class proportions during training using stratified k-folds (Yadav and Shukla, 2016). Combining the predictions on the validation and test sets from each fold, for each model, created an ensemble of networks to evaluate prediction variance (Hansen and Salamon, 1990). Training was repeated for 20 epochs for each fold, where an epoch represents an entire pass of the training set. Training specifications included a batch size of 32, and

momentum values of 0.9 in batch normalization layers of ResNet50, InceptionV3, and Xception (Ioffe and Szegedy, 2015). The learning algorithm minimized the log loss (cross-entropy) function through backpropagation using the Adam optimizer (Kingma and Ba, 2014)— the learning rate was set at 0.01 and reduced by a factor of 10 if the loss function failed to improve by $1e^{-3}$ after five epochs. Holographic reconstructions, object detection and classification were implemented in the NVIDIA CUDA GPU toolkit (Nickolls *et al.* 2008) using a NVIDIA GeForce GTX960 GPU with 16GB of RAM.
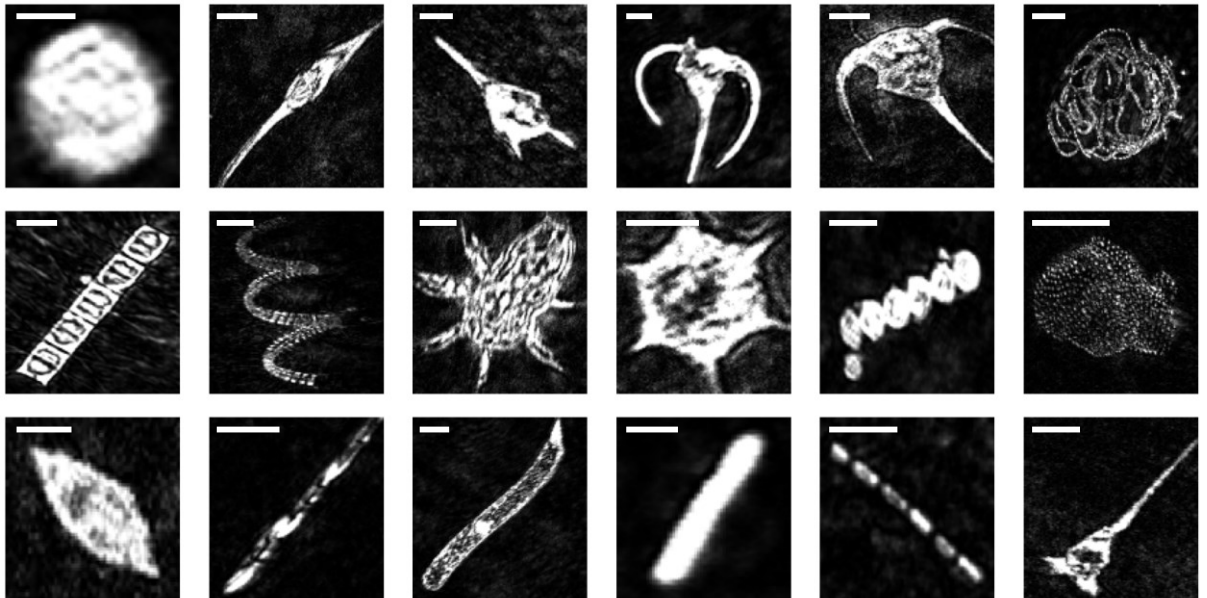
Figure 2.2. Amplitude images reconstructed and detected from specific focal planes for each plankton class. From top left to lower right: *Alexandrium tamarense*, *Ceratium fusus*, *Ceratium lineatum*, *Ceratium longpipes*, *Ceratium* sp., *Chaetoceros socialis*, *Chaetoceros* straight, *Chaetoceros* sp., Crustacean, *Dictyocha speculum*, *Melosira octagona*, *Parvicorbicula socialis*, *Prorocentrum micans*, *Pseudo-nitchzia arctica*, *Rhizosolenia setigera*, Rods, *Skeletonema costatum*, Tintinnid. All images are segmented to 128×128 pixels and scale bars represent 50 μm.

2.3.5 Validation Measures

Classification performance was evaluated using three broad families of metrics: Thresholding, probabilistic, and ranked. To extend each metric to our multi-label problem, we binarized classes (one vs. all) to mimic multiple binary classification tasks. Thresholding measures are estimated from the quantity of true positives (tp), true negatives (tn), false positives (fp), and false negatives (fn) observed during training and testing. These measures assume matching class distributions between training and test sets, which we satisfied in each stratified fold. Accuracy is simply defined by the total proportion of correct predictions, whereas precision is defined by the proportion of correctly predicted positives (tp) to all predicted positives ($tp + fp$), also known as the predictive positive value (1).

$$Precision = \frac{tp}{tp + fp} \quad (1)$$

The recall defines the proportion of correctly predicted positives (tp) to all positive examples ($tp + fn$), it is equivalent to the true positive rate (2).

$$Recall\ (Sensitivity) = \frac{tp}{tp + fn} \quad (2)$$

The balanced score between precision and recall can be represented by the F1-score, calculated using a harmonic mean (3) (Ferri *et al.* 2009).

$$F1 = 2 * \frac{(Precision * Recall)}{(Precision + Recall)} \quad (3)$$

Ecologically meaningful plankton classifiers predict few false positives and a high proportion of true positives across all classes (Faillettaz *et al.* 2016). This priority favors precision, because because high precision scores imply few false positives, and the F1-score as the relative balance between precision and recall, as such, high F1-score contains

fewer false positives and false negatives across all labels (Faillettaz *et al.* 2016).

Although both metrics are more sensitive than accuracy to the performance of minority classes, each only summarizes classifier performance at a single decision threshold: The predicted probability of an image belonging to a class is converted to a label only when it surpasses a fixed, and often arbitrarily defined threshold (Tharwat, 2018). To overcome this, we generated precision-recall curves at every decision threshold to visualize their trade-off— in other words, the relationship between the fraction of correctly predicted true positives (predictive positive value) and the true positive rate (Davis and Goadrich, 2006). Precision-recall curves are robust for imbalanced classification because they are unaffected by the increasing true negatives after labels are binarized (Saito and Rehmsmeier, 2015). To summarize classifier performance for each class across every decision threshold, we computed the average precision of each class (4), where R_n and P_n are recall and precision at the nth threshold, respectively (Pedregosa *et al.* 2011).

$$Average\ Precision = \sum_n (R_n - R_{n-1})\, P_n \ (4)$$

Average precision is analogous to a non-linear interpolation of the area under each precision-recall curve (AUC-PR) (Boyd *et al.* 2013)— as a rank measure, the AUC is closely related to statistical separability between classes (Davis and Goadrich, 2006; Ferri *et al.* 2009). For a specific class, the performance baseline when evaluating AUC-PR is defined by the ratio of positives (P) to negatives (N) in the test set $y = \frac{P}{P+N}$, and is equal to the probability of a positive example being correctly classified over a negative example (Saito and Rehmsmeier, 2015). The baseline is therefore different for each class.

2.4 Results

2.4.1 Holographic Data

In total, > 17,000 holograms comprising > 70 GB of data were produced from our samples. Reconstructed by Octopus software, holograms had the highest intensity in the central axis which attenuated at the hologram edges (Figure 2.1). Hologram intensity was reconstructed in the order of eight milliseconds for a 2048×2048 hologram. The numerical holograms reconstruction, ROI clustering, and autofocusing that compose our multi-stage detection steps generated 3826 in-focus plankton objects from 19 classes (Figure 2.2). In total, the full workflow amounted to approximately 44 hours of computational time dominated by in-focus detection (> 95%), and the remainder by classification. Six classes were generated from the environmental samples including *C. fusus*, *C. lineatum*, *Ceratium* sp., Crustaceans, *P. socilais*, and Tintinnids. The remaining classes derived from monoculture and represented individual plankton species. In total, the environmental classes were less abundant than classes derived from pure cultures. The size of plankton objects ranged from 20-640 μm, with the majority smaller than 200 μm and belonging to microplankton (Table 2.1). The classes proved highly imbalanced with the greatest difference between mesoplankton Crustaceans containing 13 images, and the microplankton dinoflagellate *P. micans* containing 1074 images (Table 2.1). After augmentation, the CNN training data contained 7215 samples which when subdivided into stratified folds contained 5772 images for training and 1443 images for validation.

2.4.2 Overall Classification

The classification source code is publicly available on Github (MacNeil, 2020). The feature extraction and retraining methods produced indistinguishable performance results across classification metrics, so we will consider only the feature extraction results here. For feature extraction, the overall classification performance based on the accuracy, precision, recall, and F1-score are reported in Table 2.2. The InceptionV3 model achieved the lowest precision values at 83% and F1-score of 81%. All the remaining three models performed comparably reaching precision scores > 88%, and F1-scores > 87%. The Xception model consistently outperformed every other to achieve precision and F1-scores of 89%. The underlying classification performance for each taxon is described below by their AUC-PR. Each model clearly achieved maximum precision, recall, and F1-scores quickly— in five or fewer epochs— while the mean and standard deviation for predictions across epochs was generally low (< 2.5 %). The log loss error showed similar model behaviour overall, with error minima in fewer than five epochs and Xception obtaining the lowest error.

Table 2.2. Average performance and standard deviation of each model across folds for each threshold metric on the test set.

| | | Threshold Metrics (%) | | | |
	Model	Accuracy	Precision	Recall	F1-Score
Feature	VGG16	88.2 ± 1.2	88.4 ± 1.5	88.1 ± 0.9	87.8 ± 1.0
Extraction	InceptionV3	82.2 ± 1.8	83.7 ± 2.4	81.1 ± 2.2	81.7 ± 1.4
	ResNet50V2	88.2 ± 1.1	88.6 ± 1.3	88.1 ± 07	87.9 ± 0.9
	Xception	90.1 ± 1.6	89.8 ± 0.9	90.7 ± 0.4	89.8 ± 0.7

2.4.3 Taxa-Level Classification

The AUC-PR values for each class are reported in Table 2.3. The precision-recall curves for each model (See Figures A.4-7 in appendix) broadly showed that the highest AUC-PR values and therefore the 11 highest ranked classes included the dinoflagellates *A. tamaranse* and all four *Ceratium* taxa, along with diatoms for all three *Chaetoceros* taxa and *M. octagona,* the silicoflagelle *D. speculum.*, and our noise class. Both Xception and ResNet50 ranked the rarest class of Crustaceans highly. As the best classifier, Xception even ranked rare taxa *C. lineatum* (0.91) and Crustacean (0.86) higher than the *Chaetoceros* straight morphotype (Figure 2.2), despite containing less than a seventh of examples. Classification performance deteriorated for the remaining seven taxa to ranking only marginally better than random for the choanoflagellate *P. socialis*, and the diatoms including *P. arctica.* and *S. costatum*, as well as the Rods morphotype and the ciliate Tintinnids. The dinoflagellate *P. micans* was the only taxa that was unanimously ranked worse than random in each model— that is, AUC-PR values below their class baseline— despite it containing nearly three times as many examples as the next most abundant class. No clear difference in classification performance occurred between size classes.

Table 2.3. Area under the precision-recall curves calculated using average precision for each class.

Class	AUC-PR			
	VGG16	InceptionV3	ResNet50V2	Xception
Alexandrium tamarense	0.97	0.85	0.96	0.98
Ceratium fusus	0.88	0.55	0.78	0.89
Ceratium lineatum[a]	0.76	0.60	0.76	0.91
Ceratium longpipes	0.97	0.93	0.98	0.99
Ceratium sp.	0.79	0.59	0.85	0.92
Chaetoceros socialis	0.98	0.96	0.99	0.99
Chaetoceros straight	0.80	0.61	0.77	0.84
Chaetoceros sp.	0.93	0.83	0.96	0.98
Crustacean[a]	0.56	0.30	0.84	0.86
Dictyocha speculum	0.98	0.88	0.97	0.99
Melosira octagona	0.98	0.92	0.97	0.98
Noise	0.96	0.87	0.94	0.98
Parvicorbicula socialis[a]	0.01	0.01	0.01	0.01
Prorocentrum micans[a]	0.19	0.21	0.20	0.16
Pseudo-nitchzia arctica[a]	0.02	0.01	0.03	0.03
Rhizosoenia setigera	0.05	0.04	0.04	0.04
Rods	0.19	0.13	0.14	0.14
Skeletonema costatum	0.06	0.07	0.05	0.06
Tintinnid[a]	0.01	0.01	0.01	0.02

[a]indicate rare classes with < 25 examples in the training set.

2.5 Discussion

This work demonstrates the usefulness of DIHM equipped with a workflow for volumetric hologram reconstruction, objection detection and autofocusing to classify plankton images using off-the-shelf CNNs. In general, plankton size did not obviously affect classification, but the sharpest images and most resolvable features were ranked higher, except for the dinoflagellate *A. tamarense*, which was likely well recognized as the only visually circular species in the dataset. In the highly ranked dinoflagellates, apical and antapical horns in *C. fusus* and *C. lineatum* and the spines in *C. longpipes* and

Ceratium sp. resolved clearly and were conspicuous features. The dinoflagellate *P. micans* was poorly resolved and classified, it is possible that the small cell size (< 100 μm) limited any detection of its thecal plates or small (< 10 μm) apical spine (Hasle and Syvertsen, 1997). For the diatoms, the chained *C. socialis*, *Chaetoceros* sp., and *M. octagona* were all distinct from each other with colonies, spirals, and straight chains that likely contributed to their reliable classification. More broadly, many chained objects showed a discernible interstitial space between cells which was especially distinct in *Chaetoceros* and *S. costatum* (Figure 2.2), although the setae of the *Chaetoceros* classes was rarely visible. Among the poorly ranked diatoms *S. costatum, R. setigera*, and *P. arctica* all lacked further morphological definition. Similar to the WHOI plankton dataset (Orenstein *et al.* 2015), the small sized choanoflagellate *P. socialis* only displayed colonies of flame bulbs and the silica loricae and flagellum cannot be seen— likely explaining its unanimously poor ranking by each CNN.

The complex morphology of plankton also presents a problem of image scale: The features available for detection in this study were limited to those that remained after objects were segmented to 128×128 scale. These image sizes are different from the ImageNet images used to train each CNN— VGG16 and ResNet50 were trained on 224×224 images and InceptionV3 and Xception were trained on 229×229 images. This suggests encouraging transferability to our holographic plankton images. Although scaling effectively normalizes the wide variety of features and explicitly retains scale invariant features, imaged plankton features can obviously vary with size, and therefore scale invariant features only partially describe the spatial composition of any object (Gluckman, 2006). Segmenting objects at multiple scales could capture scale-variant

features, but examples of scale-variant detection are less common. Artist attribution is an example of a complex classification task where multi-scale images (256, 512, 1024, 2048 pixels) systematically improved CNN predictions using both coarse and fine grain features of digitized artworks belonging to the Rijksmuseum, at the Netherlands State Museum (van Noord and Postma, 2017). But currently, multi-scale CNNs lose scale invariant features that otherwise emerge during scaling and augmentation, and these features are not guaranteed to emerge during convolutional feature extraction. Further research on scale-variant feature detection could overcome this limitation and help identify the diversity of plankton features that are more or less resolvable at different scales.

Holography has certain technical challenges for capturing high-quality plankton features, owing first to the need for numerical reconstruction of a sample volume, followed by object detection and autofocusing. In assessing the HoloSea, Walcutt *et al.* (2020) observed two notable biases underlying particle size and density estimates, including the attenuated light intensity from the point source, both radially and axially across the sample volume and secondly, that foreground objects inevitably shade the volume background. Although this study is concerned with classification, both biases are present in this study. Several modifications offered by Walcutt *et al.* (2020) apply here: Adjusting the point source-to-camera distance to expand sample space illumination and create a more uniform light intensity, scaling object detection probability based on pixel intensity, and local adaptive thresholding to improve ROIs detection consistency at the dimmed hologram edges— as opposed to the fixed, global thresholding algorithm used here. Because objects are less likely to be detected at the hologram edges, only a fraction

of the particle field is consistently imaged. The total volume imaged, calculated as the product of the number of holograms and the volume of each hologram (maximally 0.1 mL), should be corrected by the actual illuminated proportion of the sample volume: For the HoloSea, Walcutt *et al.* (2020) empirically derived the working image volume at 0.063mL per hologram. The digital corrections are likely simpler and should be implemented in future quantitative assessments; unless the increasing ability of deep learning algorithms in holographic reconstruction, enhancing depth-of-field and autofocusing can outperform instrument-specific corrections (Rivenson *et al.* 2019). Nonetheless, holography opens new opportunities for high-throughput volumetric image analysis and the robust modular casings of DIHM— which operate in the abyssopelagic zone (~6000m) (Bochdansky *et al.* 2013) and High Arctic springs (Jericho *et al.* 2010)— make for versatile instruments to deploy in oceanic environments.

Classification tasks for almost every image domain have greatly improved with transfer learning (Weiss *et al.* 2016), including for plankton (Orenstein and Beijbom, 2017). With a transfer learning approach, our results show good classification performance for multiple groups of abundant micro and mesoplankton— encompassing the size spectra (5-50 μm) that microbial eukaryotic diversity peaks (de Vargas *et al.* 2015). Classification performance was also high for several rare taxa including Crustaceans, *C. fusus*, *C. lineatum*, and *Ceratium* sp., all of which contained fewer than 50 training examples. Publicly shared datasets like ImageNet have been central for classification benchmarks, increasing training examples for a wider recognition of features within and across imaging modes and minimizing the imbalance of class distributions in small and large datasets (Kornblith *et al.* 2018). For plankton, open access

datasets such as the In Situ Ichthyoplankton Imaging System (Cowen and Guigand, 2008)

dataset shared through the Kaggle's National Data Science Bowl competition, and the

WHOI dataset captured by the Imaging Flow Cytobot (Olsen and Sosik, 2007) are

important starting points. But both image modes are quite different from holographic

images: To improve transferability of feature recognition, an open database specific to

the holographic domain could promote wider use and shrink the gap between its high-

throughput image production and analyses. To that end, the holographic plankton images

used here will be publicly available in the Cell Image library (See Availability of data

and materials).

Although the primary concern of this work is detection and classification from

holographic images, generalizing classifiers to unseen plankton populations remains

challenging ([a]Gonzalez *et al.* 2017). Plankton vary widely and are invariably observed

unevenly. However rare plankton classes can be important and removing them from

datasets (e.g., Faillettaz *et al.* 2016; Luo *et al.* 2018) is not desirable if imaging

instruments are to be maximally effective in sampling the plankton community. Ballast

water quality testing, for example, relies on presence-absence of rare, invasive taxa

(Casas-Monroy *et al.* 2015). The proper classifier evaluation is in performance on the

original imbalanced datasets, not how certain performance measures can be tuned by

synthetically manipulating class balances (Provost, 2000). As an alternative, optimizing

decision thresholds in precision-recall curves for each class has seen revived interest, and

benefits from bypassing the generated biases in common oversampling methods (Collell

et al. 2018). For evaluating classifiers of imbalanced plankton datasets, we encourage

wider use of ranking metrics like AUC-PR, which summarize the trade-offs of any

particular metric at every decision threshold and appear rarely used in plankton classification tasks (e.g., Lumini and Nanni, 2019; Pastore *et al.* 2020).

In machine learning, quantification is increasingly separated from classification as a different, and altogether more challenging learning task; several quantification approaches are reviewed in [a]González *et al.* (2017). For *in-situ* plankton imaging systems, classification algorithms do not account for shifting class distributions across samples, false positive rates acquired during model training, and because most plankton studies aim to estimate total group abundance across observations in space, or through time, the learning problem then becomes at the level of the sample, not the individual image ([b]González *et al.* 2017). Although any classifiers false positives can be corrected for (e.g., Briseño-Avena *et al.* 2020), a generalizable classifier would contain robust sample-level error, not at the taxon level ([b]González *et al.* 2017). The features learned by CNNs for classification, similar to those described here, can be used for plankton quantification. González *et al.* (2019) input high-level features from pre-trained CNNs into quantification algorithms to estimate plankton prevalence throughout more than six years of the Martha's Vineyard time series collected by the Imaging Flow Cytobot and showed high correspondence— even approaching perfect— between probabilistic quantifiers and ground-truth estimates even in rare taxa (< 1 mL^{-1}). These results are encouraging that even imperfect quantifiers can deliver biologically meaningful estimates of a wide range of plankton.

2.6 Conclusion

This work integrates a simple and deployable high-throughput holographic microscope with autofocused object detection and state-of-the-art deep learning classifiers. The combined high-throughput sampling and digital image processing of the HoloSea shows its ability to produce and reconstruct sharp images of important plankton groups from both culture and environmental samples, although some further optical corrections are desirable. Classifying a wide-ranging plankton classes, both rare and abundant, the pre-trained CNNs showed compelling baselines through rapid learning and complex feature recognition despite the starkly different holographic image domain. Overall, this ensemble of tools for holographic plankton images can confidently separate and classify the majority of our micro-mesoplankton classes. With the exception of a small dinoflagellate and choanoflagellate with poorly resolved features, classification performance was unaffected by plankton size.

Holographic microscopes are well suited for volumetric sampling in aquatic ecosystems and the relatively simple in-line microscope configurations, comparable to the model used here, can be modified for robust designs to deploy in harsh environments. These advantages allow in-line holographic microscopes to be towed, attached to conventional CTD rosettes, or stationed *in situ* for continuous monitoring. Moreover, the recent achievements in holographic reconstruction and image processing allow micrometer resolution from high-throughput instruments. Achieving real-time data interpretation remains unfeasible, but the rapid sampling capacity of holography leaves automatic classification, although improved, an outstanding challenge.

We contribute a publicly available dataset to improve CNN transferability and enhance benchmarks for plankton classification. The improvements in holographic hardware and digital capacity argues for wider use in aquatic microbial ecology and more broadly, its high-throughput potential and data-rich images warrants wider adoption in cell imaging tasks.

CHAPTER 3 — COMBINING DIGITAL HOLOGRAPHY AND AMPLICON SEQUENCING TO DESCRIBE EUKARYOTIC PLANKTON ACROSS THE NEWFOUNDLAND SHELF

3.1 Introduction

The Grand Banks of Newfoundland (Figure 3.1) are a group of continental embankments where two major wind-driven Western Boundary Currents—the Labrador Current and the North Atlantic Current—converge in the Northwest Atlantic Ocean (Richardson, 2001). The Labrador Current carries cold, dense water southward from the Arctic, and the North Atlantic Current is the northward branch point of the Gulf Stream, carrying warmer less dense waters northward (Richardson, 2001). The Grand Banks represent the foremost component of the Newfoundland-Labrador Shelves, containing at least 14 designated "Ecologically and Biologically Significant Areas" (DFO, 2019) that support one of the most seasonally productive regions in the Northwest Atlantic (Henson *et al.* 2009). Along the banks, two designated areas are at contrast: The Northeast Slope deepens gradually and is subject to a majority of Arctic water influence, whereas the Southeast shelf is an ancient sandy plateau, remaining uniquely shallow (< 90 m), and subject to intensive vertical mixing between the North Atlantic Current and the Labrador Current (Han *et al.* 2008). The Southeast Shoal region, located along the shelf break of the SE Grand Banks (East of 51 °W and South of 45 °N), is a historically productive ecosystem of highly dense benthic communities and seasonally dynamic phytoplankton blooms (Fuller and Myers, 2004). At both the Northeast Slope and the SE Grand Banks, and on the Newfoundland Shelf generally, mixed layer depth exerts strong controls on surface productivity (e.g., Harrison *et al.* 2013); it is broadly defined as the portion above the deepest mixing point which can exchange heat and gases with the atmosphere,

propagates acoustic sound, and delivers deep, nutrient-rich waters to the surface (Kara *et al.* 2003). Incorporating mixed layer depth estimates into biological community surveys can help interpret physical and biological patterns from the water column structure.

Quantitative observations of plankton communities have been collected using a diversity of techniques: Net tows for manual cell counts, pigment analysis, particulate organic carbon and nitrogen content (POC and PON), flow cytometry for cell density of smaller phytoplankton and bacterioplankton, among others (Lombard *et al.* 2019). Imaging techniques are widening with the numerous optical methods (Table 1.1) that record quantitative information in 2-D and 3-D images: From imaging approaches, standard ecological observations can be obtained including cell density (cells L^{-1}), diversity, and species behaviour (Lombard *et al.* 2019). Cell size can also be captured as an indicator of functional diversity (e.g., Dutkiewicz *et al.* 2020) with unintrusive, in-flow sampling designs. Together, the sampling capacity and stored digital information in plankton image data has tremendous potential for quantitative microbial ecology, especially for morphologically diverse eukaryotes (Benfield *et al.* 2007). But any single instrument can only record a fraction of the eukaryotic size spectrum (Lombard *et al.* 2019); plus, high-throughput instruments like digital in-line holographic microscopes (DIHM) have lower image resolution compared to light microscopy (a trade-off for enhanced depth-of-field)— although genus-level resolution is often achievable (Sun *et al.* 2008; Bianco *et al.* 2020; Guo *et al.* 2021). In this gap, revealing the microbial diversity in a water sample, high-throughput sequencing technologies has been revolutionary (Di Bella *et al.* 2013).

In the marine environment, both cellular and multiple sources of extracellular DNA can be filtered from water samples, termed environmental DNA (eDNA), and sequenced using newly available high-throughput technologies (Deiner *et al.* 2017). For marine environments, high-throughput sequencing of microbial eDNA has matured swiftly, revealing previously unexpected eukaryotic diversity (e.g., Stoeck *et al.* 2010), global patterns of community structure (Lima-Mendez *et al.* 2015)— emphasizing biotic controls of community composition— and apparent species richness > 150K in surface oceans (de Vargas *et al.* 2015). But indiscriminate diversity surveys are often not desirable, and thus for targeting specific groups, marker genes have emerged as a powerful molecular fingerprint across environmental samples, known as metabarcoding (Deiner *et al.* 2017). Here, evolution is our guide: The small subunit (SSU) of the ribosomal RNA (rRNA) gene is highly conserved (Woese and Fox, 1977), and bears hypervariable regions (V1-V9) that differ among prokaryotic (16S) and eukaryotic (18S) life (Eickbush and Eickbush, 2007). Despite its benefits, metabarcoding by high-throughput sequencing produces compositional data, i.e., sequence (read) counts that do not reflect absolute abundances (Gloor *et al.* 2017). The decoupling of reads from absolute cell counts derives from the nature of sequencing instruments—because read counts are an artefact of the machine (Gloor *et al.* 2017)— and from cells themselves: Many taxa contain several copies of the SSU rRNA gene (Zhu *et al.* 2005; Sargent *et al.* 2016; Gong and Marchetti, 2019). Compositional analysis is not new (Aitchison, 1982), but still, it can only identify proportional differences, not absolute. In 1896, Karl Pearson noted that these constraints on compositional data lead to spurious correlations in

biological data (Pearson, 1897). In essence, quantitative observations are needed to support metabarcoding analysis.

Surveying plankton density and biomass can complement molecular surveys by connecting magnitudes of abundance to diversity. For microbial eukaryotes, the size structure via biovolume (μm^3) is a widely reported proxy for biomass that warrants a brief explanation. Size structure of the microbial community is an important property of ocean food webs, indicating carbon content through cell volume (Menden-Deuer and Lessard, 2000), and associating plankton taxa to geometric shapes that can explain differing cell buoyancy and contributions to the biological pump (e.g., Tréguer *et al.* 2018), or reflect evolutionary fitness (Ryabov *et al.* 2021). The size diversity of phytoplankton alone is enormous, spanning nine orders of magnitude in cell volume (Finkel *et al.* 2010). Here, basic digital information is extracted from 2-D holographic plankton images to estimate biovolume using simple geometric models (Hillebrand *et al.* 1999). Although simple geometric models oversimplify eukaryotic morphology (e.g., Sun and Liu, 2003), intricate cell measurements have high degrees of human error and many cell structures (e.g., spines or setae) contribute little to overall carbon content (Menden-Deuer and Lessard, 2000). For this reason, biovolume has been widely estimated by extrapolating linear measurements to simple geometries (e.g., Álvarez *et al.* 2012; Sacca, 2017; Hryick *et al.* 2019). Lastly, biovolume can be rendered into biomass as carbon content (pg C Cell^{-1}) through a volume-to-carbon scaling relationship (power function) (Menden-Deuer and Lessard, 2000). Although the volume estimates and scaling laws are simple, they are surprisingly robust (e.g., Jakobsen *et al.* 2015), and consistent with

modern 3-D confocal microscopy for some abundant microphytoplankton (Mcnair *et al.* 2021).

This chapter compares DIHM, marker gene metabarcoding, and standard physicochemical data to describe eukaryotic community composition and structure within the Grand Banks of Newfoundland. The two spatial transects considered here represent contrasting physical and biological environments that are captured by our paired techniques.

3.2 Methods

3.2.1 Study Area

The data collected here is from a large section of the Newfoundland Shelf during the Atlantic Zone Monitoring Program (AZMP) 2019 cruise (*RRS James Cook; cruise JC190;* November 25-December 4, 2019). The AZMP has conducted oceanographic and biological sampling across the shelf regions in the Northwest Atlantic Ocean on seasonal and interannual scales since 1998 (Therriault et al. 1998). The data analysis focuses on two transects: The Bonavista Banks and Southeast (SE) Grand Banks. The Bonavista Banks transect contains 14 stations across > 400 km (48.7-50.3 °N) located along the Northeast Slope. The SE Grand Banks contained 15 stations across > 550 km (46.5-42 °N) extending across the shelf onto the Southeast Slope (i.e., Southeast Shoal). The distance between stations varied (Figure 3.1) — both the largest and smallest occurring on the SE Grand Banks— with averages of 97 km and 66 km at SE Grand Banks and Bonavista Banks respectively.
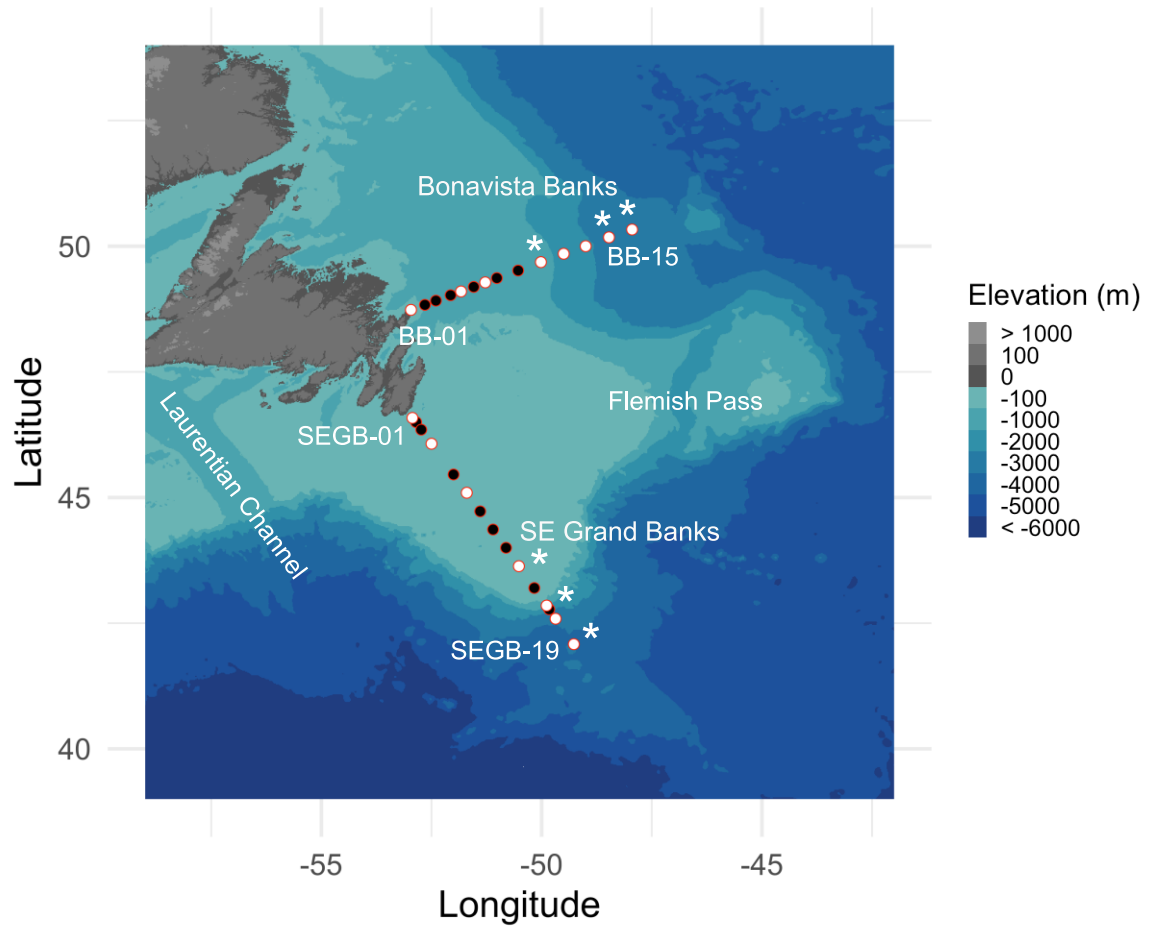
Figure 3.1. The Bonavista Banks (BB) and SE Grand Banks (SEGB) transects with sampling stations depicted. Every station (white and black) contains a full water column profile, white dots indicate Niskin bottle samples for DNA and imaging at 5m, and the asterisks indicate additional Niskin samples from 20 and 50m. Isobaths are illustrated from the slope to the off-shelf regions. Note that the ship did not sample every original station, but original station names are kept here, thus SEGB-19 was the 15th station sampled at the SE Grand Banks.

3.2.2 Oceanographic Physicochemical Data

Full water column profiles were obtained by lowering a conductivity-temperature-depth (CTD) rosette equipped with a calibrated fluorometer and a dissolved oxygen sensor (Seabird SBE-9+), from surface waters to 10 m above the seafloor. All CTD files are generously provided by Department of Fisheries and Oceans (Fisheries and Oceans

Canada, 2020) and were handled in R using the oce package ([a]Kelley, 2018). No nutrient

data was made available from these casts. The CTD files were filtered in a standard

fashion: CTD casts were trimmed for downcasts to remove upcast anomalies, interpolated

using the default method at 1 m increments, smoothed with local (Boxcar) averaging and

gridded into sections (Kelley, 2018). No aberrant measurements were flagged in the CTD

casts. The available CTD variables included physicochemical data for temperature (°C),

absolute salinity (g kg^{-1}), oxygen (mg L^{-1}), and chlorophyll-a fluorescence (µg L^{-1}). Since

seawater density (S_a; kg m^{-3}) is a function of heat content and dissolved salt

concentrations, *in-situ* seawater density calculations incorporated each stations latitude,

longitude, temperature, and absolute salinity [5]. The physicochemical data corresponding

to the Niskin samples (Table 3.1) were extracted and averaged within three depth bins (5-

6, 20-21, 50-51) to define average conditions at 5, 20, and 50 m. These values were

considered to correspond to our Niskin bottle samples for community composition, cell

densities and size structure.

The depth of mixing in the water column controls nutrients delivery to surface

plankton communities. Two complementary methods were used to estimate mixed layer

depth across the Bonavista Banks and SE Grand Banks transects: First, a conventional

criterion cut-off defining the vertical depth where a combined change of 1 °C and 0.125

kg m^{-3} equals the mixed layer depth; second, using a derivative method, where the

greatest rate change in temperature across a depth ($\frac{\Delta T}{\Delta z}$) indicates a thermocline, where the

mixed layer is above this point.

---

[5] Seawater density is a function of temperature and absolute salinity (S_a). S_a (g kg^{-1}) is preferred over
practical salinity (psu) in accordance with the Thermodynamic Equation of Seawater (TEOS-10) protocols;
absolute salinity incorporates conductivity and empirically derived location-dependent seawater ion ratios.
Details can be found in [a-b]Kelley (2018).

3.2.3 Sample Collection: Paired Imaging and Filtration

Seawater was collected using 20 L Niskin bottles attached to the CTD rosette. The full dataset includes 51 water samples from 34 stations (See appendix Figure C.1) emptied directly into 10 L cubitainers and sampled for imaging and DNA filtration. This full dataset was sequenced for the 18S rRNA gene V4 variable region and is described below, but only 27 samples belonging to 15 stations along the Bonavista Banks and SE Grand Banks were analyzed for the paired imaging samples (Table 3.1). A minimum of 2 L sample volume was allocated for paired imaging of micro and mesoplankton and filtration of eDNA. First, seawater was pumped at 150 mL min^{-1} using a peristaltic pump (Fisherbrand™ GP1000) into a digital in-line holographic microscope (DIHM), the HoloSeaS5[6] (92×351 mm, 2.6 kg) at 10 frames s^{-1}. The lower size limit of the object detection algorithm was 20 μm (MacNeil *et al.* 2021). Directly after imaging, samples were pumped through sterile tubing (Masterflex®) and filtered onto 10 μm polycarbonate Isopore Membrane filters (Millipore, United States). Each filter was handled with ethanol-sterilized tweezers and promptly stored at -80 °C.

Table 3.1. All 15 stations, ordered from on-shelf to off-shelf, on the SE Grand Banks (SEGB) and Bonavista Banks (BB) analyzed for discrete water samples, with complete coverage for metabarcoding and partial coverage for imaging samples. The grey columns indicate the stations sampled at 5, 20, and 50 m, with volume sampled, total holograms, and estimated volume imaged per depth.

---

[6] http://4-deep.com/

Station	SEGB-01	SEGB-05	SEGB-08	SEGB-12	SEGB-15	SEGB-17	SEGB-19	BB-01	BB-06	BB-08	BB-11	BB-12	BB-13	BB-14	BB-15
Lat (°N), Lon (°W) Time Depth (m)	46.58, -52.93 Day 5	46.07, -52.5 Day 5	45.09, -51.7 Night 5	43.63, -50.51 Night 5 20 50	42.85, -49.88 Day 5 20 50	42.58, -49.68 Night 5	42.08, -49.27 Night 5 20 50	48.3 -52.96 Day 5	49.1, -51.83 Night 5	49.28, -51.28 Night 5	49.68, -50.01 Day 5 20 50	49.85, -49.5 Night 5	50, -49 Night 5	50.17, -48.47 Night 5 20 50	50.33, -47.95 Day 5 20 50
CTD	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Community Composition	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cell Density	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Size Structure	Yes	Yes	Yes	Yes	Yes	Yes	Yes	NA	NA	NA	NA	NA	NA	NA	NA
Holograms (5, 20, 50m)	1701	1694	1884	1667 1946 2177	1866 1799 1895	2204	2146 3793 1585	2455	1893	2739	2326 2608 1896	3116	2623	1998 3024 2000	2293 2206 2953
Volume (L) (5, 20, 50m)	2.3	2.2	2.6	2.1 2.0 2.2	2.2 2.0 2.2	2.4	2.4 2.4 2.1	2.6	2	2.7	2.3 2.5 2.1	2.7	2.7	2.2 2.7 2.7	2.7 2.6 2.0
Volume Imaged (mL) (5, 20, 50m)	107	106	119	110 112 137	117 113 119	139	135 239 105	155	119	173	147 164 119	196	165	126 190 126	144 139 186

3.2.4 DNA Extraction and Illumina MiSeq Sequencing

DNA was extracted from the 10 μm polycarbonate filters using the DNeasy Plant Mini Kit (Qiagen, Germany) according to the manufacturer's instructions, with the following modifications. Extractions began by adding 50 of lysozyme (5 mg mL^{-1}) (Fisher BioReagents, United Kingdom) to each filter, then vortexing on high for 30 s, followed by a five-minute incubation at room temperature. Then 400 μL of the DNeasy Plant Mini Kit lysis buffer AP1 was added to each filter plus 45 μL of proteinase K (20 mg mL^{-1}) (Fisher BioReagents, United Kingdom). The samples were then incubated at 52 °C with shaking (300 rpm) for one hour. Once completed, 4 μL of RNase A (Qiagen, Germany) was added to the filters, vortexed and a 10-minute incubation— tubes were inverted twice during incubation to homogenize sample contents. Afterwards, according to the manufacturers protocol, 130 μL of DNeasy Plant Mini Kit Buffer P3 was added to the lysate, then incubated for five minutes on ice, centrifuged for five minutes (20,000 g), and the lysate was finally pipetted onto a spin column (Qiagen, Germany) to be centrifuged (20,000 g) for two minutes. The remaining steps to isolate and elute the DNA contents followed the manufacturer's protocol. The final DNA elution included 100 μL of elution buffer. The DNA concentrations and purity were measured with a NanoDrop 2000 (Thermo Scientific, United States). The final DNA aliquot (27 μL) was stored at -80 °C until further analysis

Sequencing by Illimina MiSeq and data processing followed the Microbiome Amplicon Sequencing Workflow (Comeau *et al.* 2017) within the QIIME2 v. 2020.8 (Quantitative Insights Into Microbial Ecology) platform (Bolyen *et al.* 2019). Samples were amplified using dual-indexing Illumina fusion primers that targeted the V4 508 bp

region of the eukaryotic 18S rRNA gene (E572F- CYGCGGTAATTCCAGCTC; E1009R- AYGGTATCTRATCRTCTTYG) (Comeau *et al.* 2011). To briefly describe the sequencing preprocessing steps according to Comeau *et al.* (2017): Paired-end sequencing reads were inspected for high-quality read pairs using FastQC (v. 0.11.8) (Andrews, 2010), primers were removed using Cutadapt (v. 2.10) (Martin, 2011), forward and reverse reads were then stitched, and any low-quality or chimeric reads are filtered out using the default options (Comeau *et al.* 2017) of VSEARCH (v. 2.7.0) (Rognes *et al.* 2016). To derive single nucleotide resolution between high-quality reads and determine Amplicon Sequence Variants (ASVs, Callahan *et al.* 2017), Deblur (Amir *et al.* 2017) was used to subtract erroneous reads based an upper-bound error profile from the sequence-to-sequence Hamming distances of neighbouring reads. During this step, we also positively filtered for 18S sequences using the PR2 18S rRNA database (Guillou *et al.* 2012). As an initial sensitivity test, three trim lengths were compared (350 375, 400 bp) for removing low-quality reads and the impact on community composition (Mohsen *et al.* 2019) using an in-house toolkit[7].

The representative ASVs were taxonomically classified using the multinomial naïve-Bayes QIIME2 q2-feature-classifier plugin (Bokulich *et al.* 2018). The classifier was pre-trained in-house on the SILVA (v.138.1) database (Quast *et al.* 2012), specific for our 18S V4 region (E572F-E1009R). The classified ASVs were then filtered in two ways: (1) Rare ASVs were removed if they were less than 0.1% of mean sample depth in accordance with Illumina's estimation of bleed-through errors during sequencing, (2) all ASVs only classified to the rank of kingdom were not analyzed further. The remaining

---

[7] https://github.com/dianahaider/q2-comp

ASVs were then exported from a QIIME artefact to a BIOM table and imported into R (v. 3.6.2) (R Core Team, 2019) as a Phyloseq object (McMurdie and Holmes, 2013) for further analysis.

3.2.5 Biodiversity Analysis

The extensive taxonomic profiles of ASVs from the 18S rRNA gene V4 region provided a powerful tool for detecting biological diversity within and between compositional samples. Here, richness per sample at the taxonomic rank of genus was estimated as a surrogate for α-diversity, using the R package breakaway ([a]Willis and Bunge, 2015). Briefly, the breakaway model estimates unobserved genera by fitting a non-linear regression to consecutive taxa frequency ratios (singletons, doubletons, tripletons, etc.) and predicts the number of unobserved taxa (i.e., The number of groups with a frequency of zero). An intuitive explanation for this model is that if relatively few taxa were rarely observed, then most of the diversity was probably sampled, thus standard errors for unobserved taxa will be small and the observed richness is reliable, and vice-versa. This method adjusts observed genus richness by estimating unobserved taxa and calculating model error, in this way, the uncertainty of observed richness can be visualized accounting for the sequence depth. Most usefully, this adjustment allows more reliable comparison between samples (Willis, 2019).

Differences in community composition between samples, known as β-diversity, was also assessed. Unlike α-diversity, β-diversity requires statistical distances to be calculated directly onto compositional data, requiring ASV counts to be transformed into log-ratios from a constrained simplex into an unconstrained (Euclidean) space of real

numbers (Aitchison, 1986). To prevent undefined log-ratios, a widely used pseudo-count of one was added to each zero (Quinn *et al.* 2019; Silverman *et al.* 2020). All ASV counts were transformed using a centered log-ratio (clr), which is a transformation of a composition $x = (x_1, \dots, x_i, \dots, x_D)$ defined as:

$$clr(x) = \left[ln \frac{x_1}{g(x)}, \dots, ln \frac{x_i}{g(x)}, \dots, ln \frac{x_D}{g(x)} \right] \quad (1)$$

where each ASV (x) is divided by the geometric mean of all ASVs in a sample given by $g(x) = \sqrt[D]{x_i \dots x_D}$. The clr-transform is a widely used compositional transformation because it uses a sample-wide reference (Quinn *et al.* 2019). β-diversity was estimated by computing a principal component analysis of the Aitchison distances between samples and the statistical significance of clusters was tested using a permutational analysis of variance (Anderson, 2001) in the vegan package (Oksanen *et al.* 2013).

3.2.6 Digital in-line holographic imaging

The paired imaging and filtration of each water sample provides quantitative profiles to match the taxonomic profiles from the 18S metabarcoding. The imaging samples were analyzed from surface (5 m) waters for 15 stations, seven from the SE Grand Banks and eight from Bonavista Banks, of which three shelf break stations along each transect were also analyzed at 20 and 50 m. The specifications and theoretical background for the HoloSea are described in Chapter 2 (MacNeil *et al.* 2021). The biological Regions of Interest (ROIs) in each reconstructed plane are filtered for in-focus objects using 4-Deep Stingray software[8], and these objects compose our quantitative

---

[8] http://4-deep.com/products/stingray-software/

profiles. Every in-focus image was identified to the lowest possible taxonomic rank (Tomas *et al.* 1997). Cell concentrations were calculated for each sample based on the fraction of volume imaged from the total volume calculated by multiplying the number holograms by the effective volume per hologram (0.063 mL) of the HoloSea (Walcutt *et al.* 2020). A linear regression compared estimated cell concentrations with the binned fluorescence values for both transects, using a Model II regression for uncertainty associated with all measurements (Laws and Archie, 1981).
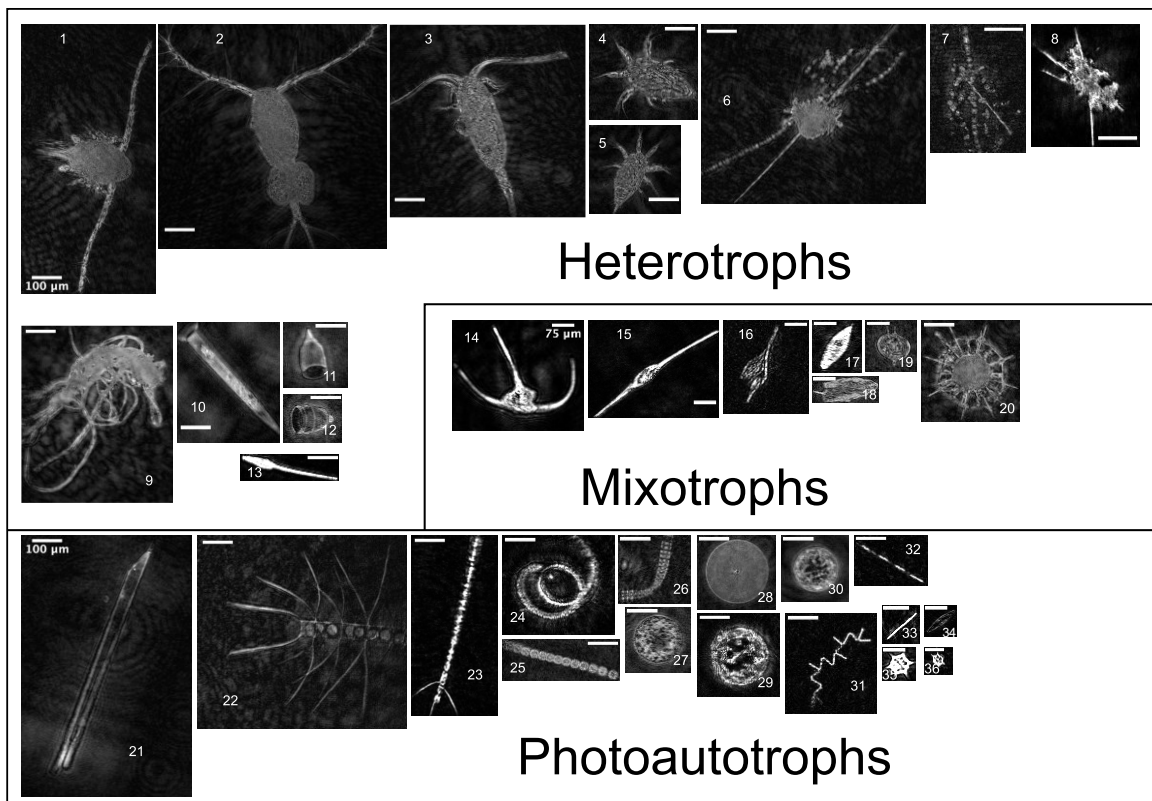
Figure 3.2. The collection of taxa observed on the Grand Banks. Groups are broadly divided into their trophic level with scale bars for each level, labelled at the first image. The heterotrophs included adult copepods (1-3), larval nauplii (4-5), the osmotrophic phylum Labyrinthulomycetes (6-8), the Amoebozoa *Platyamoeba* (9). tintinnids (10-12), and appendicularians (13). The mixotrophic dinoflagellates (14-19) included several genera of *Tripos* (14-16), with *Tripos fusus* (15), *Tripos lineatum* (16), *Gyrodinium* (17), *Prorocentrum* (18), and *Protoperidinium* (19). The heterotrophic radiolarian Acantharia (20) commonly bears photosynthetic symbionts, creating a mixotrophic nutrition (Decelle

and Not, 2015). The photosynthetic autotrophs included diatoms (21-35), and the silicoflagellates genera *Dictyocha* (35-36). The diatom genera included *Proboscia* (21), *Chaetoceros* (22-24), plus taxonomically unresolved chain-forming groups (25-26), centric groups (27-30), *Thalassionema* (31), *Pseudo-nitzschia* (32), rod-shaped groups (33), and *Nitzschia* (34).

3.2.6 Eukaryotic Phytoplankton Biovolume

Phytoplankton biovolume was estimated from the holographic images to quantify the plankton community structure and as a proxy for total biomass. Biovolume estimates were restricted to the phytoplankton (diatoms, dinoflagellates) on the SE Grand Banks stations, totalling 3203 objects. The other phytoplankton belonging to the Ochrophytes, were excluded due to scarcity. Phytoplankton were selected because they contain established shape-specific formulas (Hillebrand *et al.* 1999; Sun and Liu, 2003), and as described below, the SE Grand Banks provided the most examples. Similar procedures exist for the copepods (e.g., Hrycik *et al.* 2019), however they appear exceptionally inconsistent (Karnan *et al.* 2017) and likely skewed by complex, peripheral structures (e.g., prorosome).

The biovolume calculation extrapolated simple linear cell measurements (length, width, diameter) from image contour approximation in OpenCV (Bradski, 2000), to a volumetric quantity (μm^3) assuming rotational symmetry. The assigned shapes and volume formulas are based on classifications by Álvarez *et al.* (2012) and listed in Table 3.2. This extrapolation method has shown high consistency on 2-D images taken by the FlowCam plankton camera (e.g., Álvarez *et al.* 2012; Álvarez *et al.* 2014; Hrycik *et al.* 2019). Each volume estimate was converted to carbon content per cell (pg) based on an allometric volume-to-carbon scaling relationship (Menden-Deuer and Lessard, 2000). When scaling volume-to-carbon, cell carbon (C_c; pg C Cell^{-1}) is given by:

$$C_c = b \log (V_c) + a \quad (2)$$

Where cell volume (V_c) is scaled by the coefficients a and b. For diatoms $a = 0.288$ and $b = 0.881$. For non-diatoms $= 0.216$ and $b = 0.939$ — here this includes the dinoflagellates. The scaling coefficients account for empirical differences between diatoms and other major protists (single-celled eukaryotes), attributed to their characteristically large vacuoles that create a non-linear relationship between cell volume and carbon density (Menden-Deuer and Lessard, 2000). Phytoplankton carbon biomass (C_b) is calculated by summing cell abundance (n_i) with C_C described above (Jakobsen *et al.* 2015):

$$C_b = \sum_{n=1}^{N} n_i C_C \quad (3)$$

Table 3.2. The taxonomic grouping of phytoplankton from the Grand Banks, including the assigned simplified shape and volume formula for biovolume calculation. The W variable indicates cell width, L indicates cell length, and D indicates cell diameter.

Taxonomic Group	Genera	Shapes for Biovolume	Volume Formula
Centric	NA	Sphere	$V = \dfrac{\pi}{6} \times W^3$
Centric diatom	*Chaetoceros*	Cylinder	$V = \dfrac{\pi}{4} \times W^2 \times L$
	Proboscia	Cylinder	$V = \dfrac{\pi}{4} \times W^2 \times L$
Chains	NA	Cylinder	$V = \dfrac{\pi}{4} \times W^2 \times L$
Pennate diatom	*Pseudo-nitzschia*	Cylinder	$V = \dfrac{\pi}{4} \times W^2 \times L$
	Nitzschia	Cylinder	$V = \dfrac{\pi}{4} \times W^2 \times L$
Diatom	*Thalassionema*	Rectangular Box	$V = W \times L \times D$
Dinoflagellate	*Tripos*	Ellipsoid	$V = \dfrac{\pi}{4} \times L^2 \times W$
	Tripos fusus	Fusiform	$V = \dfrac{\pi}{6} \times W^2 \times \dfrac{L}{2}$
	Tripos lineatum	Ellipsoid	$V = \dfrac{\pi}{4} \times L^2 \times W$
	Protoperidinium	Ellipsoid	$V = \dfrac{\pi}{4} \times L^2 \times W$
	Prorocentrum	Ellipsoid	$V = \dfrac{\pi}{4} \times L^2 \times W$
	Gyrodinium	Ellipsoid	$V = \dfrac{\pi}{4} \times L^2 \times W$
Rods	NA	Cylinder	$V = \dfrac{\pi}{4} \times W^2 \times L$

3.4 Results

3.4.1 Oceanographic Conditions

The physical and biological properties differed markedly between Bonavista Banks and SE Grand Banks transects (Figure 3.3). The Bonavista Banks section is horizontally structured with cold, mixed, and dense surface waters and low ambient fluorescence < 1.25 µg L^{-1} (Figure 3.3 A). The surface waters become increasingly warm (> 4 °C) and saltier (> 34 g kg^{-1}) towards the shelf break of the Northeast Slope (> -50 °W) (Figure 3.3 B). Conversely, the SE Grand Banks is both vertically and horizontally structured with a stratified water column on the shelf containing warmer, less saline, less dense surface waters and a localized bloom indicated by relatively higher fluorescence (> 4 µg L^{-1}) at the shelf break (> -51 °W)— i.e., Southeast Shoal (Figure 3.3 C). The higher fluorescence corresponds with intrusions of colder, marginally saltier, but overall denser water masses (Figure 3.3 C-D), where upward transport is evident. Both transects had the highest dissolved oxygen content (> 8 mg L^{-1}) along the coast above 50 m at the Bonavista Banks and below 50 m at the SE Grand Bank and. The SE Grand Bank slope also exhibits a highly oxygenated pocket co-incident with the productive shelf break.
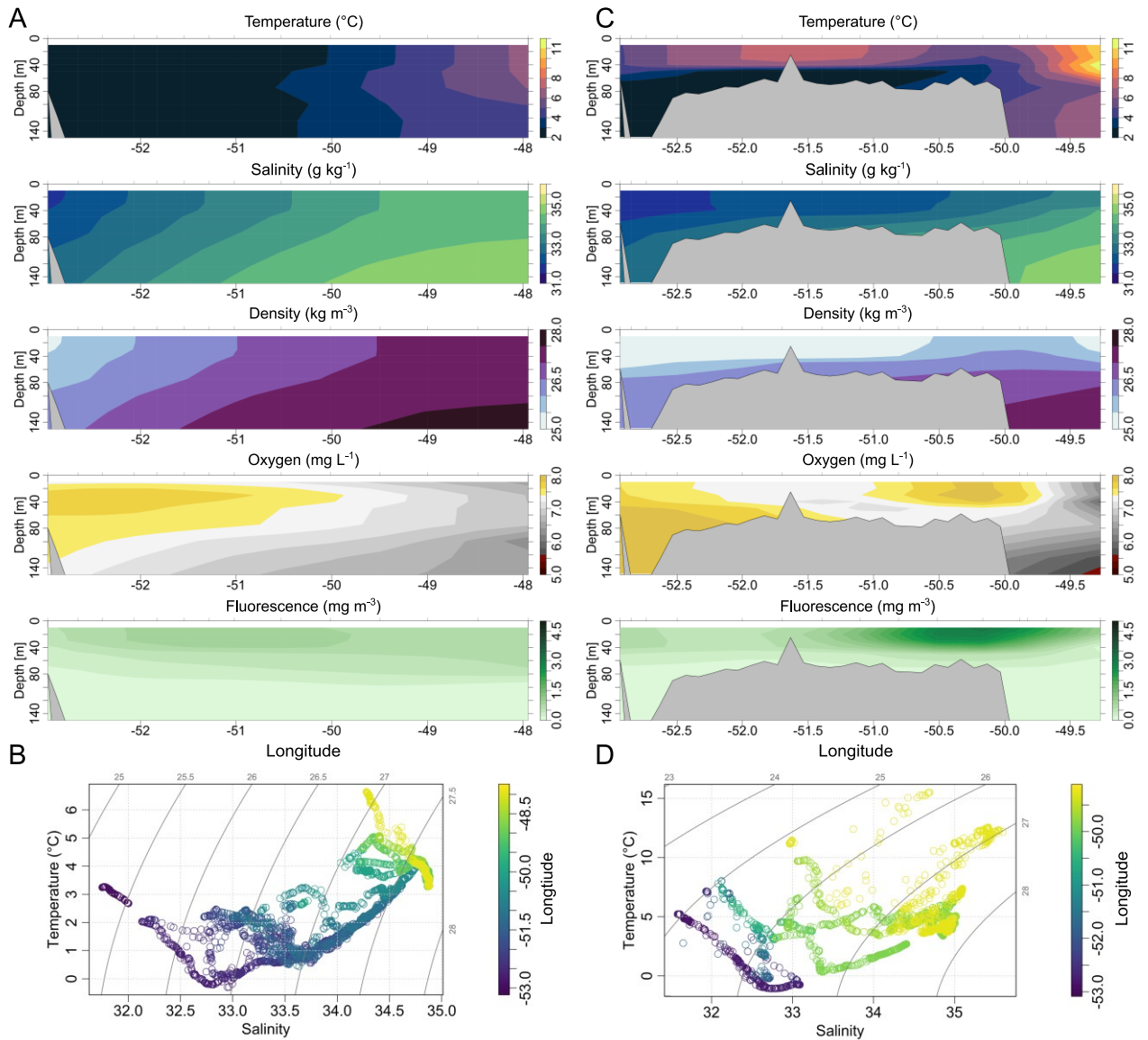
Figure 3.3. The 150 m section plots for A-B) Bonavista Banks and C-D) SE Grand Banks with corresponding temperature-salinity diagrams across the longitudinal gradient of the shelf. See Figure C.1 in appendix for 1000 m section plots. The location of each profile cast is indicated as tick marks on the top x-axis of each section plot. The isopycnals on the temperature-salinity diagrams indicate constant density. Section plot colormaps were taken from cmocean, which correct long-standing visual biases in classic oceanography jet colormaps that create arbitrary maximum and minimums within the color palette, regardless of the underlying data (Crameri et al. 2020).

The estimated depth of the mixed layer varied considerably between the criterion

and derivative method: The derivative method indicates substantial variability between

profiles on the dynamic Bonavista Banks and the maximum depth of the mixed layer was estimated < -50 °W, past the shelf break (Figure C.2 in appendix). However, both methods estimated relatively similar patterns showing a deeper mixed layer on the Bonavista Banks (< 100-150 m) than on the SE Grand Banks (< 45 m). The maximum mixed layer depth estimated by the derivative method at the Bonavista Banks is potentially spurious, but the well-mixed surface layer does indicate a potentially deepening mixed layer (Figure 3.3 A). On the SE Grand Banks, the mixed layer depth estimates shallow at the shelf break, (> -51 °W) where denser, colder water disrupts stratified water column. Together, the water column profiles indicate a shelf break upwelling event, likely induced by a hydrographic front at SE Grand Banks edge, where deeper, colder, denser waters— often nutrient rich— rise to the surface layer.

3.4.2 Plankton Community Composition and Diversity

The sequencing reads from the Bonavista Banks and the SE Grand Banks were a subset from the full dataset across the Newfoundland Shelf. This full dataset of 51 samples produced > 571 K sequencing reads with 181 to 84,984 sequences per sample (median = 5846). Denoising retained > 337 K sequencing reads ranging widely between 181 to 53,894 sequencing reads per sample (median = 3147). Filtration for rare and kingdom-level classifications removed an additional 5557 reads and 1269 rare taxa, resulting in a final > 332 K sequencing reads belonging to 248 taxa. The dataset was zero-inflated, with > 75% sparsity (zeros). Focusing on samples taken at the Bonavista Banks and SE Grand Banks, sparsity was 71 and 81%, respectively, and sequences per sample ranged 181 to 26,299 (median = 5977). The frequency of singletons was generally

low (< 1%) for both transects. The shallow sequencing depth with 181 reads (SEGB-05) was excluded from the subsequent analysis.

The first sensitivity test for the sequencing library, comparing the raw (unadjusted and untransformed) α and β-diversity for each quality-score trim length (350, 375, 400 bp) showed no large or systematic differences. Thus, only the 350-trim length results are presented here. For the Bonavista Banks and SE Grand Banks, the α-diversity adjusted for unobserved taxa estimated minimal differences from raw observed values, however the standard errors show systematic underestimates in the observed values (Figure 3.4). In total, richness ranged from 8-49 genera per sample (mean = 27), with the largest standard error suggesting an underestimate of < 10 genera. Between the Bonavista Banks and SE Grand Banks, average estimated richness is 32 and 20, respectively. Maximum richness at the Bonavista Banks occurred towards the gradual shelf break (BB-11, BB-12), and off-shelf at SE Grand Banks (SEGB-19). Several weak correlations were observed between richness and the physicochemical data: Richness was negatively correlated with temperature (R^2 = -0.3) and positively correlated with oxygen (R^2 = 0.49) at the Bonavista Banks. With opposing correlation coefficient signs, richness positively correlated with temperature (R^2 = 0.34) and negatively correlated with oxygen (R^2 = -0.29) at the SE Grand Banks.
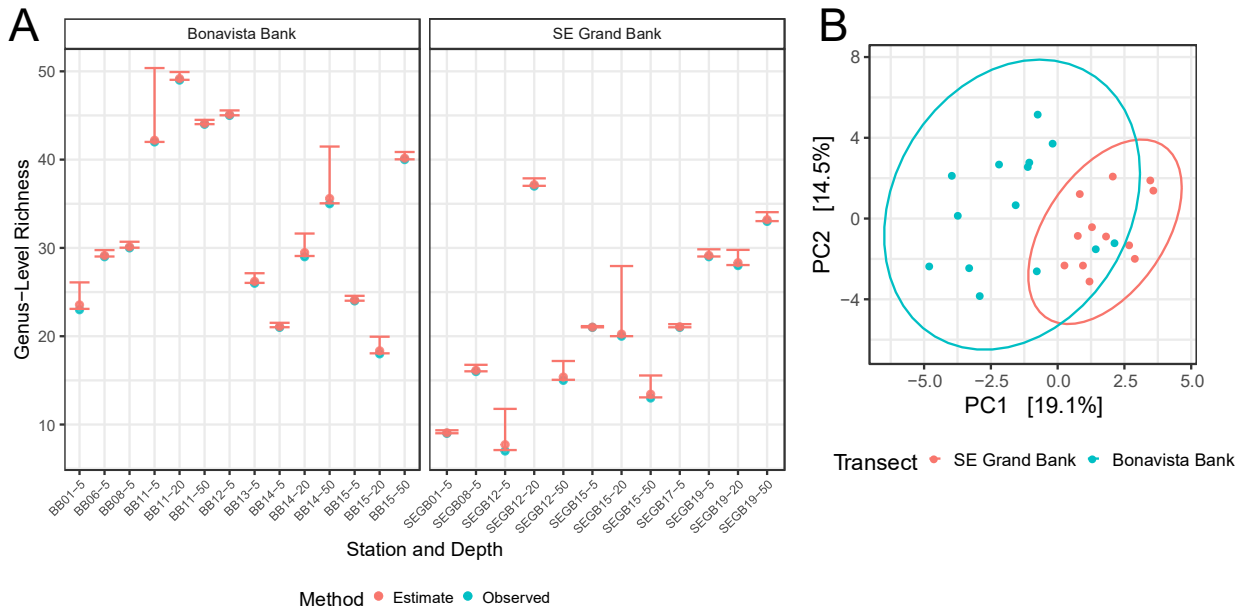
Figure 3.4. The 18S rRNA-derived A) α-diversity and B) β-diversity for both spatial transects. The error bars in A) represent standard errors of the breakaway model. The x-axis is oriented to show the stations across each shelf-gradient by longitude and depth. The distance measure used in the B) is the Aitchison's distance— a compositionally-valid Euclidean distance of the *clr*-transformed ASV dataset.

In total, photoautotrophic taxa composed ~45% of the identified groups, followed by ~30% mixotrophic dinoflagellates, and ~25% heterotrophs. In terms of sequence proportions, the heterotrophic copepod order Cyclopoida, within the phylum Arthropoda, dominated the full dataset (Figure 3.5). The Cyclopoida groups were not classified to a higher taxonomic resolution (i.e., lower taxonomic rank), but four copepod species were identified in the order Calanoida (*Calocalanus curtus*, *Centropages hamatus*, *Paracalanus parvus*, and *Temora longicornis*) at low proportions (< 1%). The copepod-dominated arthropods generally composed > 50% of sequence proportions at and off the SE Grand Banks shelf break (Figure 3.5). Among the diatom and dinoflagellate phytoplankton, dinoflagellates contribute to overall proportions between 2-70% on both the Bonavista Banks and SE Grand Banks, mostly belonging to four genera: *Tripos,*

Protoperidinium, Biechelaria, and *Pelagodinium*. Diatoms contribute fewer sequence

reads to overall proportions, between < 1-10%, consisting mostly of ASVs classified as

Chaetoceros, *Guinardia*, *Pseudo-nitzchia*, and *Thalassiosira*. The differences in

community composition between Bonavista Banks and SE Grand Banks, the *β*-diversity,

reveals a cluster for the SE Grand Banks regardless of depth; the three Bonavista Banks

samples that co-cluster with the SE Grand Banks stations were all characterized by low-

richness, arthropod dominated compositions (Figure 3.4). But these results were not

statistically significant (*p* = 0.165) using a permutational analysis of variance.
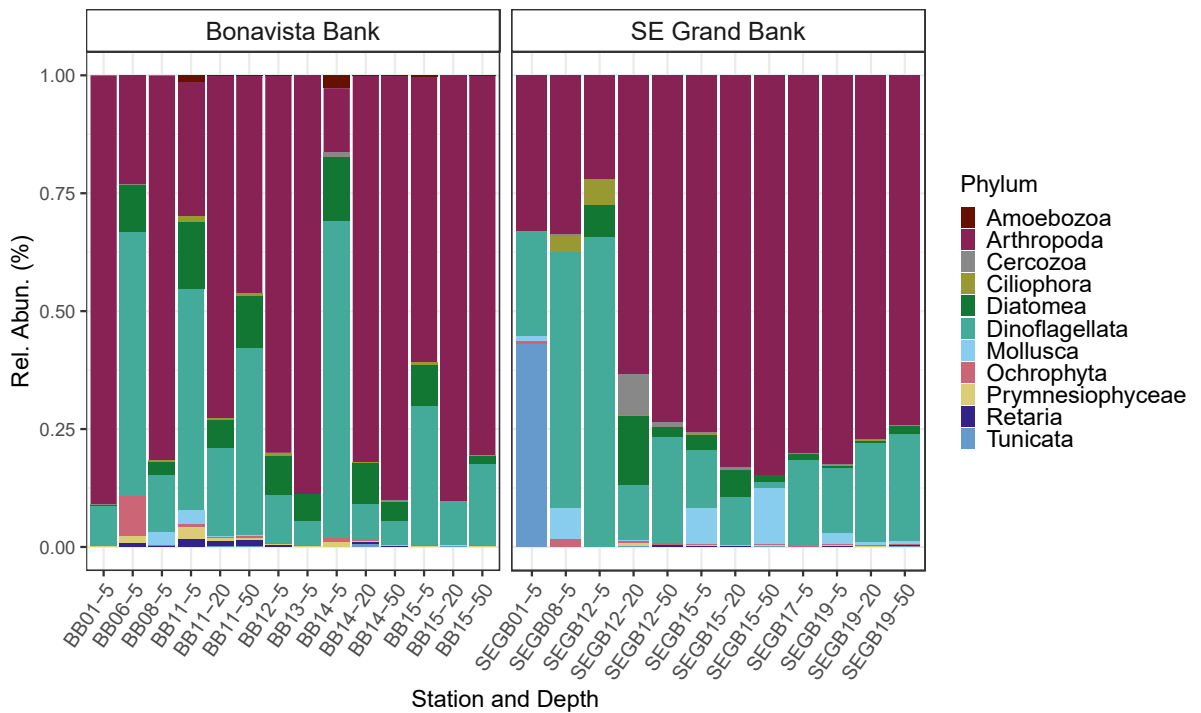
Figure 3.5. ASV proportions at the phylum-level from Bonavista Banks and SE Grand Banks transects. The x-axis is oriented to show the stations across each shelf-gradient by longitude and depth.

3.4.3 Imaging Composition, Concentration, and Biovolume

In total, > 55K holograms were analyzed from Bonavista Banks and SE Grand Banks samples and > 105K objects were detected. The volume imaged of these samples, a fraction of the total volume, ranged from 108-196 mL. The bulk of every sample was composed of small (< 50 µm) non-living particles and imaging artefacts, with > 4600 objects taxonomically identified into 26 groups, ranging from phylum (e.g., Labyrinthulomycetes) to species-level (e.g., *Tripos fusus*), and large non-living particulates > 50 µm. A collage of the groups is shown in Figure 3.2 and catalogued externally[9]. The phytoplankton were overwhelmingly diatoms and dominated by chain-forming, centric, and rod-shaped cells, for which higher taxonomic resolution was challenging. For the chains, three genera could be distinguished in multiple forms (*Chaetoceros, Pseudo-nitzchia,* and *Thalassionema*), all of which were detected in the paired DNA samples. Most chain-forming and centric diatoms remained only morphologically identified, lacking diagnostic features, and many likely belong to the genera *Guinardia* and *Thalassiosira* detected in paired DNA samples. The rod-shaped diatoms were composed of the genera *Proboscia* and *Nitzchia*, but most (> 90%) were only morphologically identified, although much too large to belong to any dinoflagellate. The dinoflagellates were more physically defined, with four genera (*Gyrodinium, Prorocentrum, Protoperidium*, and *Tripos*). *Tripos fusus* and *Tripos lineatum* were the only groups identified at a species level, however in the matching DNA samples, *Tripos fusus* and *Tripos lineatum* were undetected, and *Tripos tenuis* was the only identified

---

[9] https://liammacneil.github.io/Holo-Plankton/

species of the *Tripos* genera. *Tripos* contained the largest proportions of genus-level reads with a majority unassigned to any species.

The heterotrophic taxa included the zooplankton, tintinnids, and amoebozoans. The zooplankton were composed almost entirely of copepods, binned into the phylum Arthropoda for comparison with DNA samples. The tintinnids contained at least two genera *Salpingella* and *Condonella* (Dolan *et al.* 2012), but similarly, these were binned into the phylum Ciliophora for comparison. The Amoebazoans were generally unidentifiable below the phylum level, only distinguishable based on their large, amorphous cell bodies with near-uniform pixel intensity across the glycocalyx cell lining (Figure 3.2). The phylum Labyrinthulomycetes are a special case: They were initially grouped with non-living particulates and only identified in images after DNA sequencing revealed their presence as the species *Thraustochytrid aurantiochytrium*. The Labyrinthulomycetes are increasingly recognized as an important group ecologically and promising for biotechnologies— they are described further in the discussion (section 3.5.2).
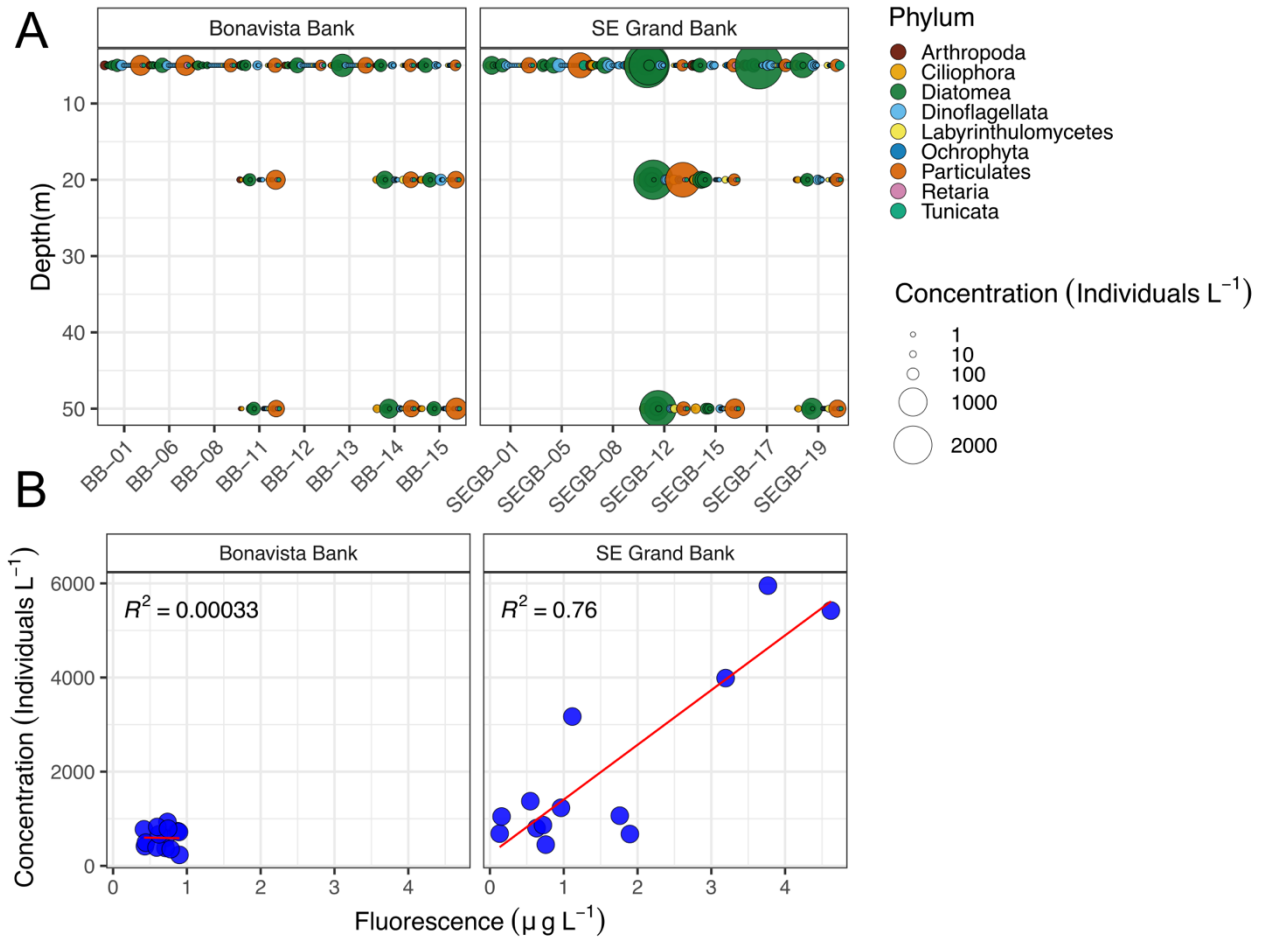
Figure 3.6. The imaging-based A) plankton concentrations oriented by longitude, from on to off-shelf stations. B) Correlation with the corresponding fluorescence measurements from both transects.

A scatterplot of fluorescence against cell concentrations (Figure 3.6) and total carbon biomass (Figure 3.7 B) with linear correlation coefficients (R^2) are included. The cell concentrations reflected the low fluorescence at the Bonavista Banks with no obvious correlation, whereas a relatively high correlation coefficient ($R^2 = 0.76$) was observed on the SE Grand Banks. In general, higher cell concentrations clustered on the shelf break of the SE Grand Banks, supporting the presence of an upwelling event that stimulated high concentrations of centric (> 350 L^{-1}), chain-forming (> 650 L^{-1}) and rod-shaped (> 5000

L^{-1}) diatoms. Relatively high concentrations of the dinoflagellate genera *Tripos* (> 80 L^{-1}), Tintinnids (> 200 L^{-1}), and Copepods (> 75 L^{-1}) were also observed at the shelf break.

Cell biovolume ranged more than four orders of magnitude from 1.13×10^3 - 3.86×10^8 µm^3 and carbon content per cell (C_c, eq. 2) ranged from 6.2 - 17.4 pg C — both minimum and maximum values were chain-forming diatoms. The diversity of cell sizes, representing the functional diversity, indicates microplankton (20-200 µm) comprised > 70% of the imaged cells, and that the less abundant mesoplankton (> 200 µm) were represented by mostly rod-shaped and chain-forming cells. Commensurate with the prevalence of chain-forming and elongate diatoms, > 90% of the shapes were cylinders and the mean aspect ratio (Cell width: length) was 0.24 (standard deviation ± 0.22). Although no pattern of biovolume emerged across SE Grand Banks, total phytoplankton carbon biomass (eq. 3) per litre peaked at the shelf break (Figure 3.7) and positively correlated with fluorescence (R^2 = 0.66).
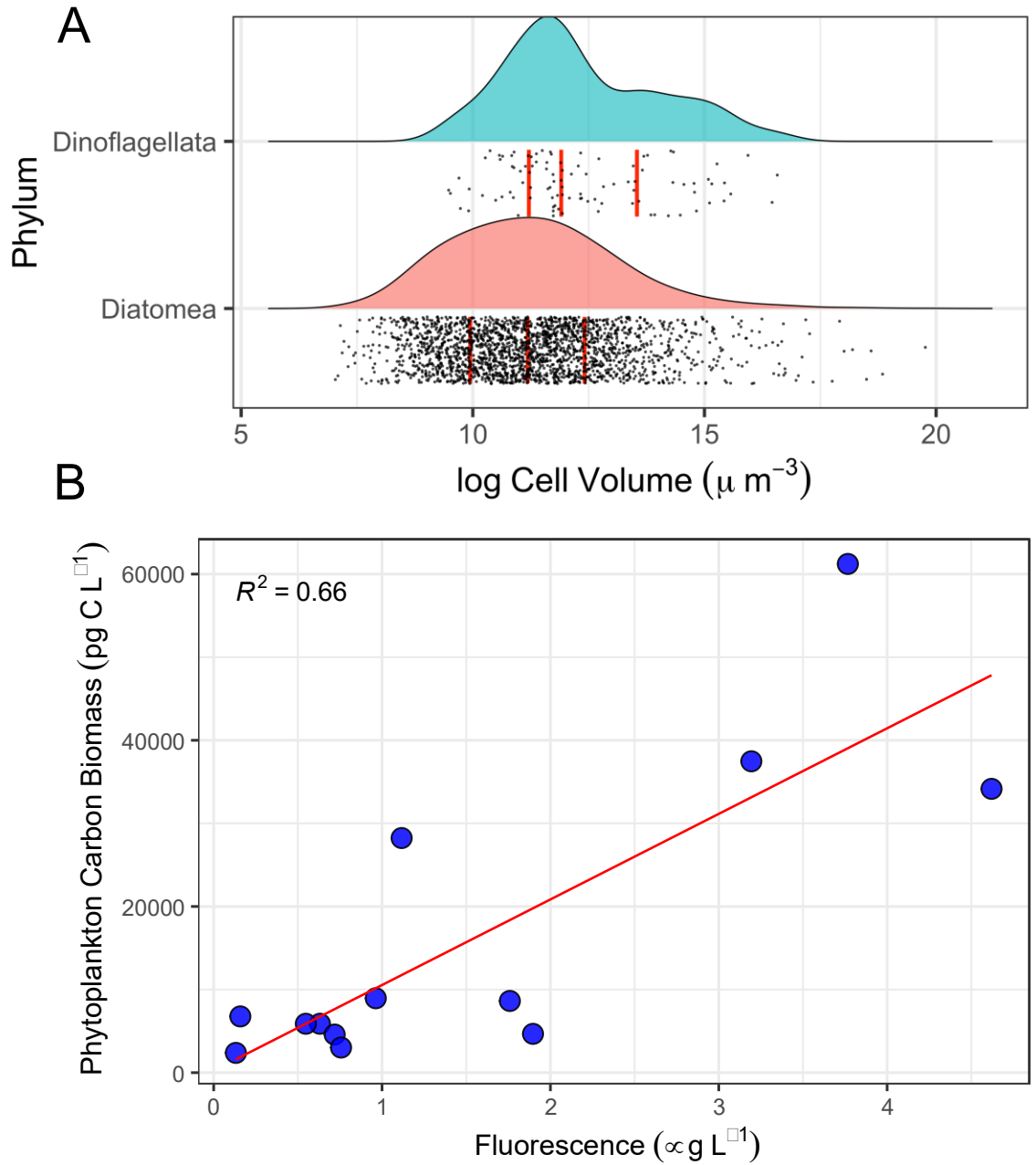
Figure 3.7. The A) log-scaled distribution of estimated biovolume for the phytoplankton, with vertical red lines in the point cloud indicating mean with upper and lower quartiles. B) The correlation coefficient of estimated phytoplankton biomass against fluorescence.

3.5 Discussion

3.5.1 Oceanographic Physicochemical Data

The combined profiles of physicochemical data, taxonomy from a eukaryotic marker gene, and quantitative imaging present clear differences between the northern and southern shelf-gradients on the Grand Banks of Newfoundland. The physical conditions at the Bonavista Banks reflected the dominant influence of Labrador Current with cold, well mixed surface layers containing relatively low productivity in fluorescence and imaging samples. The maximum mixed layer depth (< 800 m) was estimated at Bonavista Banks, which appears abnormal compared to every other profile. However, the southwest sector of the subarctic Atlantic (55-60 ° N, 56-45 ° W) contains the deepest average mixed layer depth in the global ocean, where winter conditions create mixed layers above 300-600 m (Kara *et al.* 2003; Harrison *et al.* 2013; Holte *et al.* 2017); so, our estimate might not be so far afield. In comparison, the SE Grand Banks combines a shallow and dynamic environment with submarine canyons that cut into the Southern slope and shape rich benthic habitats and pelagic productivity (Fuller and Myers, 2004). The SE Grand Banks is often denoted as the Tail of the Grand Banks, and the steep slope creates a geographic boundary between the North Atlantic Current and the Labrador Current, in other words, a hydrographic front between distinct water masses (Fratantoni and McCartney, 2010). This separation between cold, fresher shelf waters and warmer, saltier open ocean waters is a ubiquitous feature of the Northwest Atlantic Ocean (Sverdrup, 1942). The results here are a snapshot of a diverse microbial community partitioned by the SE Grand Banks environmental gradient (Figure 3.3, 3.6), containing relatively low productivity on-shelf, punctuated by higher productivity at the shelf break dominated by

diatoms, likely stimulated by upwelling of deep, cold, nutrient rich waters off the continental shelf (> -50 °W).

3.5.2 Community Composition and Diversity

Based on the dominance of diatoms in the imaged-derived community composition, and the positive correlation between cell concentrations and phytoplankton biomass with chlorophyll fluorescence, the DNA-derived community compositions clearly decoupled from absolute abundances. There are several technical and biological factors that skew compositional data from the underlying absolute abundances (discussed in the introduction, section 3.1). In this case, the dominant proportions of Arthropods in the DNA sequences are caused by the single filtration (10 µm) post-imaging, where the dominance of the DNA extracted from the large (> 500 µm), multicellular arthropods become over-represented during the amplification stage of DNA sequencing. If instead, after imaging, water samples were first filtered through a larger pore size, e.g., > 250-500 µm as per conventional net tows (Djurhuus *et al.* 2018), many Arthropods (e.g., Krill) would be separated, and the eukaryotic compositions could more accurately represent the true community within the constrained sampling capacity of the sequencing instrument. However, this will not eliminate the fundamental differences in sampling levels between imaging and metabarcoding. The images collected here are mostly living cells, but both cellular and extracellular DNA exist in the bulk DNA pools of seawater, thus it cannot be determined what fraction of a sequenced sample is living metabolizing cells, dead cells, dormant cysts, or detritus (Torti *et al,* 2015). There is also persistent uncertainty in assigning ASVs as biological organisms when reference databases, including the PR^2 18S

rRNA database used here, remain incomplete and biased for organisms of historical

importance (del Campo *et al.* 2014; Keeling *et al.* 2014). Furthermore, in marine

metabarcoding studies, statistical testing for diversity differences should be interpreted

cautiously: Inferring differences in microbial diversity ([b]Willis *et al.* 2015; 2017), or the

relative contribution of one group to overall differences in composition (e.g., differential

abundance testing *sensu stricto* Gloor *et al.* 2017), have been overwhelmingly developed

for human microbiome samples. There are many reasons behind this, but primarily,

clinical treatments allow strict control and treatment populations where effect sizes are

likely to be larger, and thus for ocean samples, robust testing for genuine differences is

likely verifiable only through multi-and-inter seasonal sampling. The time-series data in

the companion AZMP samples on the Scotian Shelf, in the Bedford Basin time-series

(Halifax, NS), or the ferry-based sampling in the Strait of Georgia (BC), are better

candidates for this approach.

Despite its limitations, the metabarcoding analysis benefitted the overall

community assessment. Principally, it provided a far deeper sampling of eukaryotic

genera than its paired imaging samples. Only 12 genera were identified across all

imaging samples, whereas 78 unique genera were identified in the metabarcoding

samples; genus-level richness was four times larger than the total imaging richness in a

single surface water sample (BB-11) on the Bonavista Banks (Figure 3.4). Even though

the imaging samples recovered quantitative profiles of the abundant micro-mesoplankton,

it was far from detecting the broader micro-mesoplankton community. There are

important microbial eukaryotic groups missed in the imaging samples but present in rare

(< 2%) proportions in DNA samples. These include protist phyla Cercozoa, which were

unidentifiable probably due to their simple morphology and small cell size (< 50 µm).

Also, the phyla Prymnesiophyceae was missed, which included important Haptophyte

groups of coccolithophorids (*Emiliania huxleyi*) that are widely abundant phytoplankton

in the North Atlantic Ocean (Bolaños *et al.* 2020), and non-coccolithophorid genera

Phaeocystis, which forms the densest colonial blooms in the North Atlantic and Southern

Oceans (Vogt *et al.* 2020). The Haptophytes were likely missed in imaging samples

because they are usually smaller cells, within the nano-microplankton (2-50 µm) range,

without conspicuous morphological structures detectable by the DIHM. The

metabarcoding also helped taxonomically validate the groups that were detected in

imaging samples. The Labyrinthulomycetes are the clearest example for improved

taxonomy as they were missed due to human error but detected with species-level

resolution in paired DNA samples. Their absence in the paired imaging samples directed

the re-assessment of detected objects which lead to 10 identified specimens. The

Labyrinthulomycetes are especially relevant in marine environments because they

contain epibiotic protists capable of degrading organic matter through osmotrophy and

creating hotspots of microbial diversity that enhance surface-to-deep ocean carbon

transfer (Bochdansky *et al.* 2017). The Labyrinthulomycetes identified here belong to the

order Thraustochytrids (Raghukumar, 2002), a dominant marine group also detected in

the photic zone during the Tara Oceans expedition (de Vargas *et al.* 2015; Pan *et al.*

2017)— although the Tara expedition samples largely omit the Northwest Atlantic

Ocean, where they have been detected at high densities in the bathypelagic zone

(Bochdansky *et al.* 2017). The Thraustochytrids express a unique lipid metabolism

production of abundant triacylglycerols that are highly valuable in dietary and medical research (Morabito *et al.* 2019).

3.5.3 Quantitative Plankton Imaging

The plankton images captured a variety of micro-mesoplankton (20-2000 μm) and broadly recorded the shelf-gradient and latitudinal differences in productivity. The error admitted by comparing the averaged values from the continuous CTD sensors to the discrete Niskin samples are likely small due to the consistency in CTD rosette sampling regimes (Lombard *et al.* 2019). The CTD fluorescence measurements only detects chlorophyll-bearing phytoplankton, meaning the non-phytoplankton (i.e., Copepods, Tintinnids, Amoeba, Labyrinthulomycetes) included in our cell concentrations further detached its correlation from fluorescence. A measurement of total particulate organic carbon would offer more appropriate comparisons to imaging-derived biomass, but this was not available. By comparison, our quantitative estimates are generally consistent with autumn survey data of the Grand Banks, showing lower productivity post-autumn bloom (September-October) (Head and Pepin, 2010; Harrison *et al.* 2013). At the SE Grand Banks shelf break, the observed fluorescence peak is lower than the Autumn seasonal peak (> 5 μg L^{-1}) (DFO, 2020). Similarly, the heterotrophic Appendicularia genera (*Oikopleura*) detected in both DNA and image samples were scarce (< 25 cells L^{-1}), reflecting low Autumn abundance (Pepin *et al.* 2011). Copepods are also presumed to be in low abundance (> 75 L^{-1}), although survey data from late autumn or early winter are rare for the Newfoundland Shelf (e.g., Head and Pepin, 2010). The notable exception to the heterotrophs was the lack of Euphausiids (krill), which are too large > 1 cm for

detection in our images and were not detected in any DNA samples yet are a widely abundant arthropod considered integral to the regional food web (Head and Pepin, 2010; DFO, 2020).

3.6 Conclusion

Quantitative imaging by digital in-line holography captured abundant micro and mesoplankton from every trophic level across two spatial transects structured by latitude, and cross-shelf gradients. The total dataset contained > 1 TB and > 5 GB of sequencing data. The productive SE Grand Banks revealed the presence of the hydrographic front between two diverging major Western Boundary Currents, stimulating high diatom concentrations in late autumn. Using the pipeline for hologram reconstruction and object detection outlined in Chapter 2 (MacNeil *et al.* 2021), routine determination of phytoplankton biomass can be achieved faster than conventional light microscopy using basic digital image analysis for linear dimensions. However, the full plankton community cannot be captured by a single instrument, and we have highlighted some trade-offs for a paired metabarcoding analysis—yielding unmatched taxonomic identification, but fundamentally compositional data decoupled from the underlying absolute abundances. Currently, matching a sequencing read, i.e., an ASV, from a metabarcoding survey to a particular image is challenging given the potentially large differences in sample sources. Matching individual 18S rRNA barcodes to taxa imaged with scanning electron microscopy has been achieved for cultured strains (e.g., Luddington *et al.* 2012), but the diversity in environmental samples makes scaling this method prohibitive. Digital imaging instruments are increasingly scaling to survey abundant plankton groups, and

pairing imaging-sequencing technologies can recover community composition, functional diversity, and species behaviours (e.g., Mars Brisbin *et al.* 2020). Combined with standard environmental sensors from a CTD-rosette, paired imaging-sequencing sampling can support broader community assessments and monitoring projects and are expected to become more precise with increasingly streamlined and automated tools.

CHAPTER 4 — CONCLUSIONS

4.1 Thesis Summary

The pace and scale of environmental shifts throughout the global ocean calls for greater understanding and more accurate predictions of planktonic ecosystems upholding ocean food webs (Lombard *et al.* 2019). Scalable sampling technologies and adequate computational tools are needed. In this thesis, I present classification and quantification of diverse micro and mesoplankton using a deployable digital in-line holographic microscope and state-of-the-art computational, molecular, and statistical tools. Chapter 2 demonstrates automatic classification of major plankton groups imaged holographically and detected in-focus using a pipeline based on pixel-intensity, along with model training and evaluation techniques constrained by imbalances in plankton abundances. This work shows high reliability in automatic classification for most plankton groups and considers improvements to future monitoring efforts by bridging developments in quantification machine learning. This work also provides an open access dataset and source code to encourage reproducibility and to augment transferability of recognition algorithms to plankton.

Chapter 3 combines digital holographic imaging with high throughput metabarcoding to describe micro and mesoplankton community composition across a productive continental shelf gradient. This analysis incorporates quantitative cell counts and size estimates paired with sample diversity for the micro-mesoplankton community. The size estimates from the quantitative imagery used basic image analysis tools to automatically extract linear cell dimensions for extrapolation into cell volume and biomass based on empirical cell size relationships. The metabarcoding analysis paired to

the image data provided compositionally-valid inferences of diversity within and between samples—greatly expanding the community survey of rare eukaryotes not amenable to image detection. The environmental conditions corresponding to all biological observations was characterized by high-frequency physicochemical data recorded by CTD casts.

It must be considered that although the physicochemical data contains a high degree of spatial coverage and vertical resolution, the biological samples amounted to 65L of seawater in the upper 50m of the water column. This is clearly limiting but the wealth of imaging and sequencing data is an encouraging prospect for information mining of small samples. Although the bulk sample volumes were limited to < 3 L by the cruise water budget, these volumes are substantially greater than the pipetted subsamples collected for conventional light microscopy (5-100 mL), which seriously underestimate species richness in the oceans (Rodríguez-Ramos *et al.* 2014; Cermeño *et al.* 2014). Imaging discrete, known volumes also permitted the quantitative estimates for cell concentrations, size-structure, and planktonic biomass. Overall, digital imagery is a considerable advancement for information storage, automated learning algorithms, and reproducibility. These strengths give digital imaging, especially the high-throughput 2-3D information of holography, a much greater potential to scale, through mining smaller samples for meaningful information or through improved classification and quantification techniques.

4.2 Some Limitations of, and Potential for Digital Holographic Microscopes

The taxonomic identity and biovolume estimations from the holographic images are affected by the in-flow sampling design, where plankton are usually imaged at their most hydrodynamically resistant plane (Moberg and Sosik, 2012). This results in some images with a small cross-sectional area at the imaging plane, and consequently carry less information about cell volume when extrapolating linear dimensions. Furthermore, three idealized assumptions are made for calculating biovolume and the cellular carbon content: First, that simple, shape-specific models represent the average cellular volume; second, that any cells third dimension is in proportion with the preceding two (Mcnair *et al.* 2021) and third, that carbon content is constant and driven by cell volume (Leblanc *et al.* 2012). Clearly, these assumptions oversimplify the stoichiometric differences between major plankton groups and the truly complex phytoplankton shapes (Figure 3.2), but often peripheral structures contribute little to overall carbon content — most notably in thecate dinoflagellates (Menden-Deuer and Lessard, 2000). So, it remains that biovolume-to-carbon scaling relationships with simple geometric models are robust (Menden-Deuer and Lessard, 2000; Álvarez *et al.* 2012; Álvarez *et al.* 2014) and still advocated as standard practice (e.g., Jakobsen *et al.* 2015; Mcnair *et al.* 2021).

The error associated with extrapolation from 2D images can be avoided with holography, using the 3D information encoded by the phase-shift of the point-source laser reconstructed simultaneously with amplitude images— where the degree of shift corresponds to the size of the cells third dimension (Monaldi *et al.* 2015; Ling *et al.* 2020). For DIHM with a single point source for the interfered and reference waves, this approach has proven accurate for estimating cell volume on red blood cell smear slides

compared to the gold-standard hematology analyzers (Ling *et al.* 2020). But smear slides have a shallow depth-of-field that require fewer reconstructed z-planes compared to bulk seawater samples. Further work is needed to evaluate phase-shift volume estimates from complex plankton images before scaling onto big environmental datasets. The precise and rapid phase-shift reconstruction from holography, without the time-consuming constraints of confocal microscopy (e.g., Roselli *et al.* 2015) or error-prone manual measurements, is a prime candidate for machine learning algorithms to derive new, more precise volume-carbon relationships in the future.

4.3 A Next Generation of Ocean Observations

Deployable imaging instruments will be among the new tools for generating ocean observations. The costs of implementing and supporting high-performance imaging instruments with technical expertise remains an impediment for wider adoption in ocean science. If adopted, numerous quantitative applications are possible, including phytoplankton abundance and community composition to ground-truth satellite observations and constrain global biogeochemical models (Lombard *et al.* 2019), or, to map predator-prey distributions and interactions for fish-larvae survivorship (Axler *et al.* 2020). Several novel species behaviours have recently been identified using *in-situ* imaging: Pseudopodial feeding strategy in acantharians from the East China Sea, likely missed previously due to destructive sampling by conventional plankton nets (Mars Brisbin *et al.* 2020); the frequent parasitization of the cosmopolitan copepod *Oithona* at the Scripps Pier, in the Pacific Ocean ([a]Orenstein *et al.* 2020), and the peak concentrations in the dinoflagellate *Cystodinium* cysts under non-optimal conditions for

either parasitic mixotrophy or autotrophy— suggesting its behaviour and feeding mode have been oversimplified with conventional techniques (Tapics *et al.* 2021). Imaging instruments can also be used to investigate size-structure patterns as an indicator of planktonic ecosystems functional diversity (Fowler *et al.* 2020), which is a deeply evolutionary trait that affects ocean food webs, biogeochemistry and evidently is selected by a host of physical (e.g., inertial and viscous forces) and biological (e.g., grazing, viral lysis) factors that structure ocean water masses (Vogel, 1981; Ryabov *et al.* 2021). Deconstructing the biologically relevant information from plankton imaging will require increasingly sophisticated tools.

Decreasing the time between collection and interpretation of the *in-situ* plankton datasets growing from numerous imaging types and environments is a major challenge. More labelled datasets are becoming open access to advance techniques for automatic recognition of both broad functional types and detailed taxonomic inventories (Benfield *et al.* 2007; Lombard *et al.* 2019). This thesis has resulted in >20K labelled holographic plankton objects from environmental and monoculture samples. These improvements support supervised machine learning techniques for automatically classifying objects, which have reproduced *in-situ* plankton distributions and abundances for a pre-defined number of categories in the Ligurian current (Faillettaz *et al.* 2016), Martha's Vineyard (González *et al.* 2019), California Current (Briseño-Avena *et al.* 2020), monitoring stations in the Northwest Atlantic and East Sound, Washington (Guo *et al.* 2021). Where labelled image databases are available, as discussed in Chapter 2 (MacNeil *et al.* 2021), future monitoring efforts based on supervised learning should adjust classifiers by false positive rates or use quantification algorithms to evaluate model error at the sample level

and adjust for out-of-distribution biases (i.e., dataset drift) that befall environmental samples ([b]González *et al.* 2017; González *et al.* 2019). When the distribution of taxa in a new sample changes from that learned by a computer model, creating dataset drift, community composition estimates will become skewed, and yet this is guaranteed to occur and is often a cue of interest ([b]Orenstein *et al.* 2020). Some plankton-relevant quantification algorithms have become available in R and Python[10], and it remains an active research focus across many disciplines (Fiksel *et al.* 2021).

4.4 Stationary and Mobile Sampling Platforms: The Bedford Basin (NS) and Strait of Georgia (BC)

Imaging instruments that can be cast or deployed autonomously onto stationary and mobile platforms can drastically expand the coverage of observations in time and space. Since April of 2016—disrupted from March 2020-2021 due to COVID-19 restrictions— weekly surface water samples (0.5-1 L) have been collected from the Bedford Basin compass buoy station and holographically imaged on a benchtop setup. I have operated the sampling since May 2019, increasing sampling volume to consistent weekly 1 L volumes which has more than doubled the hologram acquisitions collected weekly through 2016-2018, typically ~3000 holograms per sample. In total, this time-series has accumulated 165 weekly samples and >130K holograms (~0.5 TB) containing millions of unlabelled objects 20-2000 μm.

A similar version of the HoloSea instrument was deployed from February-November 2020 aboard the Queen of Alberni ferry, during crossings of the Strait of

---

[10] https://github.com/bertocast/quantification

Georgia, British Colombia. Deployment was conducted by Ocean Networks Canada, integrating the microscope into the ferry pump and valve control system, where water inflows into a mounted chamber containing the microscope which images water passing through its open sampling space. This setup ideally immerses the instrument in cool water to prevent condensation, although we have since learned this has costs in frequent sediment accumulation and mussel growth. However, this sampling routine produced imaging datasets with both spatial and temporal dimensions autonomously, initiating the microscopes hologram recording with vessel speeds > 5 knots through an ethernet connection transmitting >350 Mbps (10 fps) to a network-attached storage system. Additional technical requirements, as in most digital imaging projects, included graphical imaging cards in GPU-enabled laptops or computers, an integrated visual C++ environment (e.g., Microsoft Visual C++ 2010 x 64), and software licensing usually as a sentinel HASP key. Autonomous sampling has accumulated > 300K holograms (~2 TB) but has experienced technical challenges in reducing instrument condensation and maintaining weekly cleaning schedules during COVID-19 restrictions, leading to deteriorating image quality into the summer and autumn months. Nonetheless, > 45K holograms from at least 10 transects in March-April, each >70 km, appear amenable to object detection and quantitative analysis. Altogether, this has required extensive engineering consultation and is an excellent case study for the challenges and technical maintenance currently required to support autonomous mobile sampling efforts.
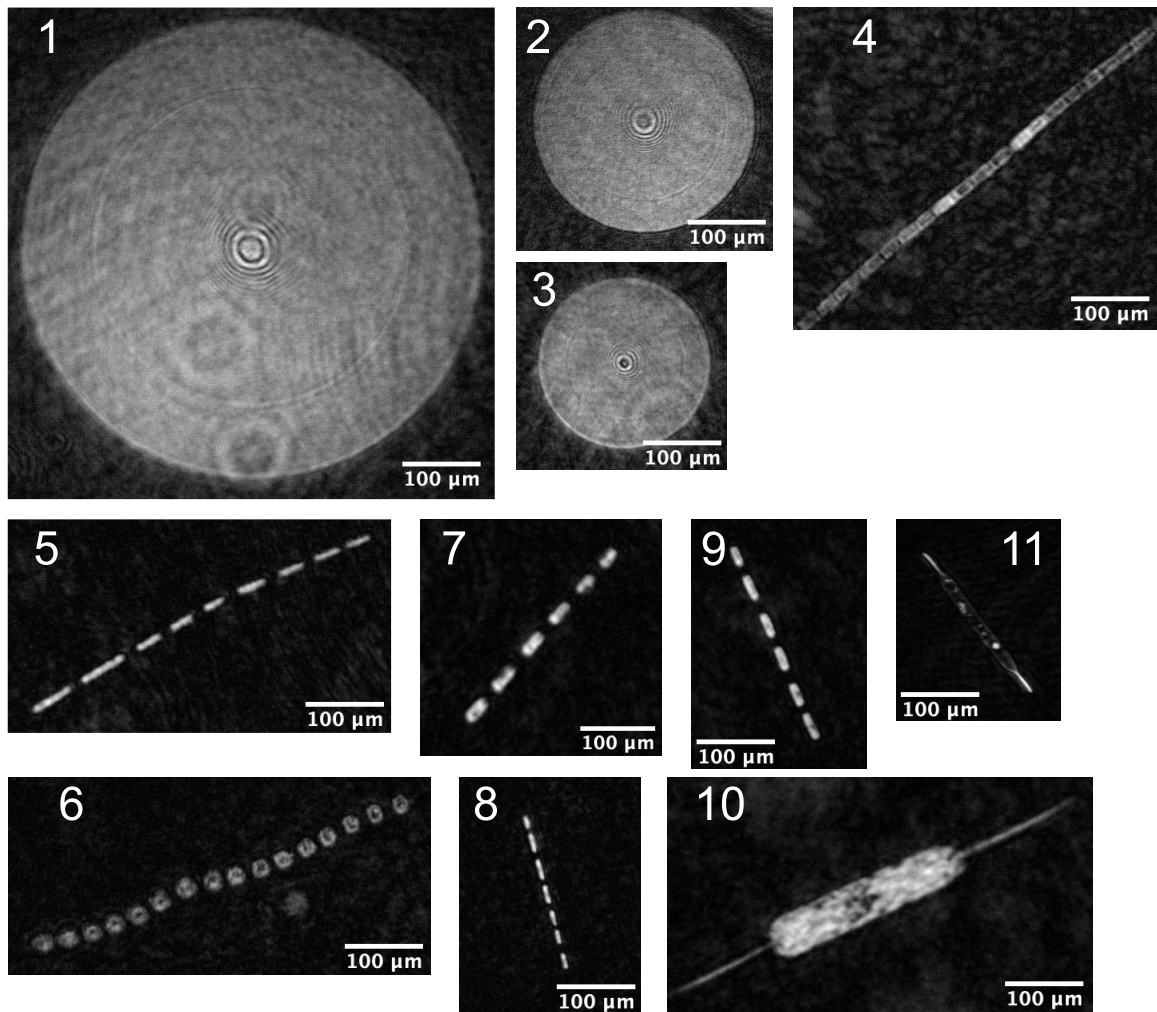
Figure 4.1. A collage of objects detected from March-April 2020 aboard the Queen of Alberni ferry. Scale bars indicate 100 µm. Based on a preliminary survey, the raw holograms from springtime contained so much material, biological and non-biological, that detecting in-focus objects was less common than in other deployments. Taxa included centric (1-3), chain-forming (4-9), and the diatom *Ditylum* (10-11).

Overall, my experience working with the HoloSea microscope and 4Deep software under numerous conditions, including the stationary and mobile sampling efforts described here, and from experimentally tuning the object detection pipeline, suggests the optimal lower size detection limit is > 50 µm, which simultaneously reduces the number of unidentifiable particles and produces the sharpest images. During object

detection, despite several qualitative comparisons between identical samples, it also appears that the local adaptive thresholding advocated in Chapter 2 does not improve, and actually degrades, image sharpness and clarity compared to global adaptive thresholding. This is likely due to the prevalence of non-uniform cells especially chain-forming taxa in many samples, which display intermittent peaks in pixel intensity, and result in incomplete objects detected by local adaptive algorithms. More work is needed to comprehensively compare each thresholding algorithm for greyscale plankton objects.

4.5 Conclusions

There is a collective interest to monitor and forecast environmental changes at large and local scales, such as changes to the Arctic and Subarctic water masses driven by sea ice depletion, or to monitor the impacts of aquaculture and agriculture on biological productivity and water quality. Future ocean observatories will include bio-optical sensors deployed onto stationary and mobile platforms, ideally semi-autonomous and incorporated into an array of interoperable, web-enabled sensors for synoptic observations of the physical and chemical environment. The next generation of observations should move beyond bulk optical properties that are currently available, and towards finer-scale measurements of how biomass, nutrients, and energy are divided across the web of life (Whitt *et al.* 2020). We have shown that digital in-line holographic microscopes are well suited for laboratory and *in-situ* deployments onboard ships-of-opportunity, and that classification can be automated for major plankton groups. Automation techniques will become increasingly important given that most imaging modes are collecting quantities of data in real-time that are unfeasible to analyze and

interpret manually. We have also complemented quantitative imaging with modern high-throughput sequencing and demonstrated a broad survey of micro and mesoplankton across the Newfoundland Shelf. In conclusion, the progress made here indicates that modern imaging and molecular technologies with adequate computational and statistical tools can be complementary, where imaging is currently best suited for abundance and biomass estimates of limited groups and metabarcoding provides deeper estimates of taxonomic richness. These strengths should be considered based on the aims of a study hoping to illuminate a significant fraction of the ocean's microbial community.

BIBLIOGRAPHY

Abadi M, Barham P, Chen J, Chen Z, Davis A, Dean J, *et al.* TensorFlow: A system for large-scale machine learning. 12th Proc USENIX Symp. Oper. Syst. Des. Implement. (OSDI) 2016; 21.

Aitchison J. The Statistical Analysis of Compositional Data. J R Stat Soc Series B Stat Methodol. 1985; 44: 139–160. https://doi.org/10.1111/j.2517-6161.1982.tb01195.x.

Álvarez E, López-Urrutia Á, Nogueira E. Improvement of plankton biovolume estimates derived from image-based automatic sampling devices: Application to FlowCAM. J Plankton Res. 2012; 34: 454–469. https://doi.org/10.1093/plankt/fbs017.

Álvarez E, Moyano M, López-Urrutia Á, Nogueira E, Scharek R. Routine determination of plankton community composition and size structure: A comparison between FlowCAM and light microscopy. J Plankton Res. 2014; 36: 170–184. https://doi.org/10.1093/plankt/fbt069.

Amir A, McDonald D, Navas-Molina JA, Kopylova E, Morton JT, Zech Xu Z, *et al.* Deblur Rapidly Resolves Single-Nucleotide Community Sequence Patterns. MSystems. 2017; 2: mSystems.00191-16, e00191-16. https://doi.org/10.1128/mSystems.00191-16.

Anderson MJ. A new method for non-parametric multivariate analysis of variance. Austral Ecol. 2001; 26: 32–46.

Andrews S. FastQC: a quality-control tool for high-throughput sequence data. Babraham Institute, Cambridge, United Kingdom. 2010: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Axler K, Sponaugle S, Briseño-Avena C, Hernandez F, Warner S, Dzwonkowski B, Dykstra S, Cowen R. Fine-scale larval fish distributions and predator-prey dynamics in a coastal river-dominated ecosystem. Mar Ecol Prog Ser. 2020; 650: 37–61. https://doi.org/10.3354/meps13397.

Behrenfeld MJ, Boss ES. Resurrecting the Ecological Underpinnings of Ocean Plankton Blooms. Annu Rev Mar Sci. 2014; 6: 167–194. https://doi.org/10.1146/annurev-marine-052913-021325.

Behrenfeld MJ, O'Malley RT, Siegel DA, McClain CR, Sarmiento JL, Feldman GC., *et al.* Climate-driven trends in contemporary ocean productivity. Nature. 2006; 444: 752–755. https://doi.org/10.1038/nature05317.

Benfield M, Grosjean P, Culverhouse P, Irigoien X, Sieracki ME, Lopez-Urrutia A., *et al.* RAPID: Research on Automated Plankton Identification. Oceanography. 2007; 20: 172–187.

Bianco V, Memmolo P, Carcagnì P, Merola F, Paturzo M, Distante C, Ferraro P. Microplastic Identification via Holographic Imaging and Machine Learning. Adv Intell Syst. 2020; 2: 1900153.

Biard T, Stemmann L, Picheral M, Mayot N, Vandromme P, Hauss H. *et al*. *In situ* imaging reveals the biomass of giant protists in the global ocean. 2016; Nature, 532: 504–507.

Bochdansky AB, Jericho MH, Herndl GJ. Development and deployment of a point-source digital inline holographic microscope for the study of plankton and particles to a depth of 6000 m: Deep-sea holographic microscopy. Limnol Oceanogr: Methods. 2013; 11: 28–40.

Bochdansky AB, Clouse MA, Herndl GJ. Eukaryotic microbes, principally fungi and labyrinthulomycetes, dominate biomass on bathypelagic marine snow. ISME J. 2017; 11: 362–373. https://doi.org/10.1038/ismej.2016.113.

Bokulich NA, Kaehler BD, Rideout JR, Dillon M, Bolyen E, Knight R, Huttley GA, Gregory Caporaso J. Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. Microbiome. 2018; 6: 90. https://doi.org/10.1186/s40168-018-0470-z.

Bolaños LM, Karp-Boss L, Choi CJ, Worden AZ, Graff JR, Haëntjens N, *et al.* Small phytoplankton dominate western North Atlantic biomass. ISME J. 2020; 14: 1663–1674. https://doi.org/10.1038/s41396-020-0636-0.

Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, *et al.* Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nature Biotech. 2019; 37: 852–857. https://doi.org/10.1038/s41587-019-0209-9.

Bork P, Bowler C, de Vargas C, Gorsky G, Karsenti E, Wincker P. Tara Oceans studies plankton at planetary scale. Science. 2015; 348: 873–873. https://doi.org/10.1126/science.aac5605.

Boyd K, Eng KH, Page CD. Area under the Precision-Recall Curve: Point Estimates and Confidence Intervals. In C. Salinesi, M. C. Norrie, Ó. Pastor (Eds.), Advanced Information Systems Engineering (Vol. 7908, pp. 451–466). 2013; Springer Berlin Heidelberg.

Bradski G. The OpenCV Library. Dr Dobb's Journal of Software Tools, 2000.

Briseño-Avena C, Schmid MS, Swieca K, Sponaugle S, Brodeur RD, Cowen RK. Three-dimensional cross-shelf zooplankton distributions off the Central Oregon Coast during anomalous oceanographic conditions. Prog Oceanogr. 2020; 188: 102436. https://doi.org/10.1016/j.pocean.2020.102436.

Callahan BJ, McMurdie PJ, Holmes SP. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. ISME J. 2017; 11: 2639–2643. https://doi.org/10.1038/ismej.2017.119.

Calle ML. Statistical Analysis of Metagenomics Data. Genomics Inform. 2019; 17: e6. https://doi.org/10.5808/GI.2019.17.1.e6.

Campbell RW, Roberts PL, Jaffe J. The Prince William Sound Plankton Camera: A profiling in situ observatory of plankton and particulates. ICES J Mar Sci. 2020; 77: 1440–1455. https://doi.org/10.1093/icesjms/fsaa029.

Casas-Monroy O, Linley RD, Adams JK, Chan FT, Drake DAR, Bailey SA. Relative Invasion Risk for Plankton across Marine and Freshwater Systems: Examining Efficacy of Proposed International Ballast Water Discharge Standards. PLoS ONE. 2015; 10: e0118267.

Centurioni LR, Turton J, Lumpkin R, Braasch L, Brassington G, Chao Y, *et al.* Global in situ Observations of Essential Climate and Ocean Variables at the Air–Sea Interface. Front Mar Sci. 2019; 6: 419. https://doi.org/10.3389/fmars.2019.00419.

Cermeño P, Teixeira IG, Branco M, Figueiras FG, Marañón E. Sampling the limits of species richness in marine phytoplankton communities. J Plankton Res, 2014; 36: 1135–1139. https://doi.org/10.1093/plankt/fbu033.

Colin S, Coelho LP, Sunagawa S, Bowler C, Karsenti E, Bork P, Pepperkok R, de Vargas C. Quantitative 3D-imaging for cell biology and ecology of environmental microbial eukaryotes. ELife. 2017; 6, e26066. https://doi.org/10.7554/eLife.26066.

Collell G, Prelec D, Patil KR. Reviving Threshold-Moving: A Simple Plug-in Bagging Ensemble for Binary and Multiclass Imbalanced Data. Neurocomputing. 2018; 275: 330-340.

Comeau AM, Li WKW, Tremblay J-É, Carmack EC, Lovejoy C. Arctic Ocean Microbial Community Structure before and after the 2007 Record Sea Ice Minimum. PLoS ONE. 2011; 6: e27492. https://doi.org/10.1371/journal.pone.0027492.

Comeau AM, Douglas GM, Langille M. Microbiome Helper: A custom and streamlined workflow for microbiome research. mSystems, 2017; 2: e00127-16.

Corrêa I, Drews P, Botelho S, de Souza MS, Tavano VM. Deep Learning for Microalgae Classification. In: Machine learning and applications (ICMLA), 2017 16th IEEE international conference on machine learning and applications. 2017; p. 20–5.

Chollet F. Keras. 2015; https://github.com/ fchollet/keras.

Chollet F. Xception: Deep Learning with Depthwise Separable Convolutions. arXiv preprint. 2017; doi: arXiv:1610.02357v3.

Cowen RK, Guigand CM. In situ ichthyoplankton imaging system (ISIIS): System design and preliminary results: In situ ichthyoplankton imaging system. Limnol. Oceanogr.: Methods. 2008; 6: 126–132.

Crameri F, Shephard GE, Heron PJ. The misuse of colour in science communication. Nature Comm. 2020; 11: 5444. https://doi.org/10.1038/s41467-020-19160-7.

Dai J, Wang R, Zheng H, Ji G, Qiao X. ZooplanktoNet: Deep convolutional network for zooplankton classification. Ocean 2016 – Shanghai. 2016; doi: 10.1109/OCEANSAP.2016.7485680.

Davis J, Goadrich M. The relationship between Precision-Recall and ROC curves. Proceedings of the 23rd International Conference on Machine Learning. 2006; 233–240.

Decelle J, Not F. Acantharia. In John Wiley & Sons Ltd (Ed.), ELS (pp. 1–10). 2015; John Wiley & Sons Ltd. https://doi.org/10.1002/9780470015902.a0002102.pub2.

Deiner K, Bik HM, Mächler E, Seymour M, Lacoursière-Roussel A, Altermatt F, *et al.* Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. Molecular Ecol. 2017; 26:, 5872–5895. https://doi.org/10.1111/mec.14350.

del Campo J, Sieracki ME, Molestina R, Keeling P, Massana R, Ruiz-Trillo I. The others: Our biased perspective of eukaryotic genomes. Trends Ecol Evol. 2014; 29: 252–259. https://doi.org/10.1016/j.tree.2014.03.006.

Deng J, Dong W, Socher R, Li LJ, Li K, Fei-Fei L. Imagenet: A large-scale hierarchical image database. In: 2009 IEEE conference on computer vision and pattern recognition, 2009; 248–255.

de Vargas C, Audic S, Henry N, Decelle J, Mahe F, Logares R. *et al.* Eukaryotic plankton diversity in the sunlit ocean. Science. 2015; 348: 1261605–1261605. https://doi.org/10.1126/science.1261605.

Di Bella JM, Bao Y, Gloor GB, Burton JP, Reid G. High throughput sequencing methods and analysis for microbiome research. J Microbiol Methods. 2013; 95: 401–414. https://doi.org/10.1016/j.mimet.2013.08.011.

Dierssen HM. Perspectives on empirical approaches for ocean color remote sensing of chlorophyll in a changing climate. PNAS. 2010; 107: 17073–17078. https://doi.org/10.1073/pnas.0913800107.

DFO. Re-evaluation of the Placentia Bay-Grand Banks Area to Identify Ecologically and Biologically Significant Areas. DFO Can Sci Advis Sec Ecosystem Status Rep. 2019/040.

DFO. Oceanographic Conditions in the Atlantic Zone in 2019. DFO Can Sci Advis Sec Sci. Advis Rep. 2020/028.

Djurhuus A, Pitz K, Sawaya NA, Rojas-Márquez J, Michaud B, Montes E, *et al.* Evaluation of marine zooplankton community structure through environmental DNA metabarcoding: Metabarcoding zooplankton from eDNA. Limnol Oceanogr: Methods. 2018; 16: 209–221. https://doi.org/10.1002/lom3.10237.

Dolan JR, Agatha S, Coats DWC, Montagnes DJS, Stoecker DK. Biology and Ecology of Tintinnid Ciliates: Models for Marine Plankton. Oxford: Wiley/ Blackwell. 2012.

Duarte CM, Regaudie-de-Gioux A, Arrieta JM, Delgado-Huertas A, Agustí S. The Oligotrophic Ocean Is Heterotrophic. Annu Rev Mar Sci. 2013; 5: 551–569. https://doi.org/10.1146/annurev-marine-121211-172337.

Ducklow H, Steinberg D, Buesseler K. Upper Ocean Carbon Export and the Biological Pump. Oceanography. 2001; 14: 50–58. https://doi.org/10.5670/oceanog.2001.06.

Dunker S, Boho D, Wäldchen J, Mäder P. Combining high-throughput imaging flow cytometry and deep learning for efficient species and life-cycle stage identification of phytoplankton. BMC Ecol. 2018; 18: 51. doi: https://doi.org/10.1186/s12898-018-0209-5.

Dutkiewicz S, Cermeno P, Jahn O, Follows MJ, Hickman AE, Taniguchi DAA, Ward BA. Dimensions of marine phytoplankton diversity. Biogeosciences. 2020; 17: 609–634. https://doi.org/10.5194/bg-17-609-2020.

Eickbush TH, Eickbush DG. Finely Orchestrated Movements: Evolution of the Ribosomal RNA Genes. Genetics. 2007; 175: 477–485. https://doi.org/10.1534/genetics.107.071399.

Ester M, Kriegel HP, Xu X. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. Proc Second Know Int Conf Dis Data Min. 1996; 6: 226-231.

Faillettaz R, Picheral M, Luo JY, Guigand C, Cowen RK, Irisson JO. Imperfect automatic image classification successfully describes plankton distribution patterns. Meth Oceanogr. 2016; 15–16: 60–77.

Faure E, Not F, Benoiston A-S, Labadie K, Bittner L, Ayata S-D. Mixotrophic protists display contrasted biogeographies in the global ocean. ISME J. 2019; 13: 1072–1083. https://doi.org/10.1038/s41396-018-0340-5.

Ferri C, Hernández-Orallo J, Modroiu R. An experimental comparison of performance measures for classification. Pattern Recognit Lett. 2009; 30: 27–38.

Fiksel J, Datta A, Amouzou A, Zeger S. Generalized Bayes Quantification Learning under Dataset Shift. J Am Stat Assoc. 2021; 1–19. https://doi.org/10.1080/01621459.2021.1909599.

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. https://doi.org/10.1093/plankt/fbp098.

Fisheries and Oceans Canada. Atlantic Zone Monitoring Program Website. Retrieved 9 May 2021 from Fisheries and Oceans Canada.

Fowler BL, Neubert MG, Hunter-Cevera KR, Olson RJ, Shalapyonok A, Solow AR, Sosik HM. Dynamics and functional diversity of the smallest phytoplankton on the Northeast US Shelf. PNAS. 2020; 117: 12215–12221. https://doi.org/10.1073/pnas.1918439117.

Fratantoni PS, McCartney MS. Freshwater export from the Labrador Current to the North Atlantic Current at the Tail of the Grand Banks of Newfoundland. Deep Sea Res Part I Oceanogr Res Pap. 2010; 57: 258–283. https://doi.org/10.1016/j.dsr.2009.11.006.

Fuller SD, Myers RA. The Southern Grand Bank: A marine protected area for the world. World Wildlife Fund Canada. 2004; Halifax, Nova Scotia. 99p.

Gabor D. A new microscopic principle. Nature. 1948; 161: 777-778.

Garcia-Sucerquia J, Ramırez JH, Prieto DV. Improvement of the signal-to-noise ratio in digital holography. Rev Mex Fis. 2005; 51: 76-81.

Garcia-Sucerquia J, Xu W, Jericho SK, Klages P, Jericho MH, Kreuzer HJ. Digital in-line holographic microscopy. Appl Opt. 2006; 45: 836-850.

Geider R, MacIntyre H, Kana T. Dynamic model of phytoplankton growth and acclimation:responses of the balanced growth rate and the chlorophyll a:carbon ratio to light, nutrient-limitation and temperature. Mar Ecol Prog Ser. 1997; 148: 187–200. https://doi.org/10.3354/meps148187.

Gloor GB, Macklaim JM, Pawlowsky-Glahn V, Egozcue JJ. Microbiome Datasets Are Compositional: And This Is Not Optional. Front Microbiol. 2017; 8: 2224. https://doi.org/10.3389/fmicb.2017.02224.

Gluckman J. Scale variant image pyramids, In: 2006 Computer Vision and Pattern Recognition. 2006; doi: http://dx.doi.org/10.1109/CVPR.2006.265.

Gong W, Marchetti A. Estimation of 18S Gene Copy Number in Marine Eukaryotic Plankton Using a Next-Generation Sequencing Approach. Frontiers in Marine Science. 2019; 6: 219. https://doi.org/10.3389/fmars.2019.00219.

[a]González P, Castaño A, Chawla NV, Coz JJD. A Review on Quantification Learning. ACM Computing Surveys. 2017; 50: 1–40.

[b]González P, Álvarez E, Díez J, López-Urrutia Á, del Coz JJ. Validation methods for plankton image classification systems: Validation methods for plankton image classification systems. Limnol Oceanogr: Methods. 2017; 15: 221–237.

González P, Castaño A, Peacock EE, Díez J, Del Coz JJ, Sosik HM. Automatic plankton quantification using deep features. J Plankton Res. 2019; 41: 449–463.

Göröcs Z, Tamamitsu M, Bianco V, Wolf P, Roy S, Shindo K. *et al.* A deep learning-enabled portable imaging flow cytometer for cost-effective, high-throughput, and label-free analysis of natural water samples. Light: Sci Appl. 2018; 7: doi: 10.1038/s41377-018-0067-0.

Gorsky G, Ohman MD, Picheral M, Gasparini S, Stemmann L, Romagnan JB, Cawood A, Pesant S, Garcia-Comas C, Prejger F. Digital zooplankton image analysis using the zooscan integrated system. J Plankton Res. 2010; 32: 285–303.

Greer AT, Cowen RK, Guigand CM, Hare JA. Fine-scale planktonic habitat partitioning at a shelf-slope front revealed by a high-resolution imaging system. J Mar Syst. 2015; 142: 111–125.

Gregory AC, Zayed AA, Conceição-Neto N, Temperton B, Bolduc B, Alberti A, *et al.* Marine DNA Viral Macro- and Microdiversity from Pole to Pole. Cell. 2019; 177: 1109-1123.e14. https://doi.org/10.1016/j.cell.2019.03.040.

Guillou L, Bachar D, Audic S, Bass D, Berney C, Bittner L, *et al.* The Protist Ribosomal Reference database (PR2): A catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. Nucleic Acids Res. 2012; 41: D597–D604. https://doi.org/10.1093/nar/gks1160.

Guo B, Nyman L, Nayak AR, Milmore D, McFarland M, Twardowski MS, Sullivan JM, Yu J, Hong J. Automated plankton classification from holographic imagery with deep convolutional neural networks. Limnol Oceanogr: Methods. 2021; 19: 21–36. https://doi.org/10.1002/lom3.10402.

Hallegraeff GM. Ocean Climate Change, Phytoplankton Community Responses, and Harmful Algal Blooms: A Formidable Predictive Challenge. J Phycol. 2010; 46: 220–235. https://doi.org/10.1111/j.1529-8817.2010.00815.x.

Han G, Lu Z, Wang Z, Helbig J, Chen N. Seasonal variability of the Labrador Current and shelf circulation off Newfoundland. J. Geophysical Research. 2008; 113: doi:10.1029/2007JC004376.

Hansen LK, Salamon P. Neural network ensembles. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1990; 12: 993–1001.

Harrison WG, Yngve Børsheim K, Li WKW, Maillet GL, Pepin P, Sakshaug E, *et al.* Phytoplankton production and growth regulation in the Subarctic North Atlantic: A comparative study of the Labrador Sea-Labrador/Newfoundland shelves and Barents/Norwegian/Greenland seas and shelves. Prog Oceanography. 2014; 114: 26–45. https://doi.org/10.1016/j.pocean.2013.05.003.

Hasle GR, Syvertsen EE. Marine Diatoms In: Tomas, CR. Identifying Marine Phytoplankton. San Diego: Academic Press; 1997. ISBN 0-12-693018-X-XV. 858pp.

Hays G, Richardson A, Robinson C. Climate change and marine plankton. Trends Ecol Evol. 2005; 20: 337–344.

He K, Zhang X, Ren S, Sun J. Identity mappings in deep residual networks. In: European conference on computer vision; 2016; p. 630–45.

Head EJH, Pepin P. Spatial and inter-decadal variability in plankton abundance and composition in the Northwest Atlantic (1958–2006). J Plankton Res. 2010; 32: 1633–1648. https://doi.org/10.1093/plankt/fbq090.

Henson SA, Dunne JP, Sarmiento JL. Decadal variability in North Atlantic phytoplankton blooms. J Geophys Res. 2009; 114: C04013. https://doi.org/10.1029/2008JC005139.

Hobson PR, Krantz EP, Lampitt RS, Rogerson A, Watson J. A preliminary study of the distribution of plankton using hologrammetry. Opt Laser Technol. 1997; 29: 25–33.

Holte J, Talley LD, Gilson J, Roemmich D. An Argo mixed layer climatology and database, Geophys Res Lett. 2017; 44: 5618–5626. https://doi.org/10.1002/2017GL073426.

Hrycik AR, Shambaugh A, Stockwell JD. Comparison of FlowCAM and microscope biovolume measurements for a diverse freshwater phytoplankton community. J Plankton Res. 2019; 41: 849–864. https://doi.org/10.1093/plankt/fbz056.

Hugerth LW, Andersson AF. Analysing Microbial Community Composition through Amplicon Sequencing: From Sampling to Hypothesis Testing. Front Microbiol. 2017; 8: 1561. https://doi.org/10.3389/fmicb.2017.01561.

Irwin AJ, Finkel ZV, Schofield OME, Falkowski PG. Scaling-up from nutrient physiology to the size-structure of phytoplankton communities. J Plankton Res. 2006; 28: 459–471. https://doi.org/10.1093/plankt/fbi148.

Irwin AJ, Finkel ZV, Müller-Karger FE, Troccoli Ghinaglia L. Phytoplankton adapt to changing ocean environments. PNAS. 2015; 112: 5762–5766. https://doi.org/10.1073/pnas.1414752112.

Jakbsen HH, Carstensen J, Harrison PJ, Zingone A. Estimating time series phytoplankton carbon biomass: Inter-lab comparison of species identification and comparison of volume-to-carbon scaling ratios. Estuar Coast Shelf Sci. 2015; 162: 143–150. https://doi.org/10.1016/j.ecss.2015.05.006.

Janocha K, Czarnecki WM. On Loss Functions for Deep Neural Networks in Classification. arXiv preprint. 2017; doi: arXiv:1702.05659v1.

Jericho MH, Kreuzer HJ. Point Source Digital In-Line Holographic Microscopy, In: Ferraro, P. Wax, A., Zalevsky, Z., editors. Coherent Light Microscopy, Berlin: Springer; 2011. p. 3–30.

Jericho MH, Kreuzer HJ, Kanka M, Riesenberg R. Quantitative phase and refractive index measurements with point-source digital in-line holographic microscopy. Appl Opt. 2012; 51: 1503-1515.

Jericho SK, Klages P, Nadeau J, Dumas EM, Jericho MH, Kreuzer HJ. In-line digital holographic microscopy for terrestrial and exobiological research. Planet Space Sci. 2010; 58: 701–705. https://doi.org/10.1016/j.pss.2009.07.012.

Jian C, Salonen A, Korpela K. Commentary: How to Count Our Microbes? The Effect of Different Quantitative Microbiome Profiling Approaches. Front Cell Infect Microbiol. 2021; 11: 627910. https://doi.org/10.3389/fcimb.2021.627910.

Kanka M, Riesenberg R, Kreuzer HJ. Reconstruction of high-resolution holographic microscopic images. Opt Letters. 2009; 34: 1162-1164.

Kara AB, Rochford PA, Hurlburt HE. Mixed layer depth variability over the global ocean. J Geophys Res. 2003; 108: 3079. https://doi.org/10.1029/2000JC000736.

Karnan C, Jyothibabu R, Kumar TMM, Jagadeesan L, Arunpandi N. On the accuracy of assessing copepod size and biovolume using FlowCAM and traditional microscopy. Indian J Geo Mar Sci. 2017; 46: 1261-1264.

Kemper B, Carl D, Schnekenburger J, Bredebusch I, Schäfer M, Domschke W, von Bally G. Investigation of living pancreas tumor cells by digital holographic microscopy. J Biomed Opt. 2006; 11: doi: 10.1117/1.2204609.

Keeling PJ, Burki F, Wilcox HM, Allam B, Allen EE, Amaral-Zettler LA, *et al.* The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): Illuminating the Functional Diversity of Eukaryotic Life in the Oceans through Transcriptome Sequencing. PLoS Biol. 2014; 12: e1001889. https://doi.org/10.1371/journal.pbio.1001889.

[a]Kelley D. "*The oce package*" in Oceanographic Analysis with R; Springer: New York, NY, USA, 2018; pp. 91-101.

[b]Kelley D. "*Practical Operating Procedures*" in Oceanographic Analysis with R; Springer: New York, NY, USA, 2018; pp. 119-186.

Kingma D, Ba J. Adam: A method for stochastic optimization. arXiv preprint. 2014; doi: arXiv:1412.6980v9.

Kornblith S, Shlens J, Le QV. Do Better ImageNet Models Transfer Better? arXiv preprint. 2018; doi: arXiv:1805.08974v3.

Kreuzer HJ, Nakamura K, Wierzbicki A, Fink HW, Schmid H. Theory of the point source electron microscope. Ultramicroscopy. 1992; 45: 381-403.

Kreuzer HJ, Jericho MHM. Digital In-line Holographic Microscopy. Imag Micro. 2007; 9: 63–65.

Krizhevsky A, Sutskever I, Hinton G. ImageNet classification with deep convolutional neural networks. Adv Neural Info Proc Sys. 2012; Accessed 19 Aug 2019.

Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv preprint. 2015; doi: arXiv:1502.03167v3.

Laws EA. Archie JW. Appropriate use of regression analysis in marine biology. Mar Biol. 1981; 65: 13–16. doi: 10.1007/BF00397062

Leblanc K, Arístegui J, Armand L, Assmy P, Beker B, Bode A, *et al.* A global diatom database – abundance, biovolume and biomass in the world ocean. Earth System Science Data. 2012; 4: 149–165. https://doi.org/10.5194/essd-4-149-2012.

LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015; 521: 436–444.

Lima-Mendez G, Faust K, Henry N, Decelle J, Colin S, Carcillo F, *et al.* Determinants of community structure in the global plankton interactome. Science. 2015; 348: 1262073–1262073. https://doi.org/10.1126/science.1262073.

Ling H, Sridhar K, Gollapudi S, Kumar J, Ohgami RS. Measurement of cell volume using in-line digital holography. Microscopy. 2021; dfaa077. https://doi.org/10.1093/jmicro/dfaa077.

Lombard F, Boss E, Waite AM, Vogt M, Uitz J, Stemmann, *et al.* Globally Consistent Quantitative Observations of Planktonic Ecosystems. Front Mar Sci. 2019; 6: doi: 10.3389/fmars.2019.00196.

Lovell DR, Chua X-Y, McGrath A. Counts: An outstanding challenge for log-ratio analysis of compositional data in the molecular biosciences. NAR Genom Bioinform. 2020; 2: lqaa040. https://doi.org/10.1093/nargab/lqaa040.

Luddington IA, Kaczmarska I, Lovejoy C. Distance and Character-Based Evaluation of the V4 Region of the 18S rRNA Gene for the Identification of Diatoms (Bacillariophyceae). PLoS ONE. 2012; 7: e45664. https://doi.org/10.1371/journal.pone.0045664

Lumini A, Nanni L. Deep learning and transfer learning features for plankton classification. Ecol Inform. 2019; 51: 33–43.

Luo JY, Irisson JO, Graham B, Guigand C, Sarafraz A, Mader C, Cowen RK. Automated plankton image analysis using convolutional neural networks: Automated plankton image analysis using CNNs. Limnol Oceanogr: Methods. 2018; 16: 814–827.

MacLeod N, Benfield M, Culverhouse P. Time to automate identification. Nature. 2010; 467: 154–155. https://doi.org/10.1038/467154a.

MacNeil L. Holographic Transfer Learning. 2020; https://github.com/LiamMacNeil/Holographic-Transfer-Learning.

MacNeil L, Missan S, Luo J, Trappenberg T, LaRoche J. Plankton classification with high-throughput submersible holographic microscopy and transfer learning, BMC Ecol Evol. 123. https://doi.org/10.1186/s12862-021-01839-0.

Malkiel E, Alquaddoomi O, Katz J. Measurements of plankton distribution in the ocean using submersible holography. Meas Sci Technol. 1999; 10: 1142-1152.

Mars Brisbin M, Brunner OD, Grossmann MM, Mitarai S. Paired high-throughput, in situ imaging and high-throughput sequencing illuminate acantharian abundance and vertical distribution. Limnol Oceanogr. 2020; 65: 2953–2965. https://doi.org/10.1002/lno.11567.

Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. J. 2011; 17:10_12 DOI 10.14806/ej.17.1.200.

McMurdie PJ, Holmes S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLoS ONE. 2013; 8: e61217. https://doi.org/10.1371/journal.pone.0061217.

McMurdie PJ, Holmes S. Waste Not, Want Not: Why Rarefying Microbiome Data Is Inadmissible. PLoS Comput Biol. 2014; 10: e1003531. https://doi.org/10.1371/journal.pcbi.1003531.

Mcnair H, Hammond CN, Menden-Deuer S. Phytoplankton carbon and nitrogen biomass estimates are robust to volume measurement method and growth environment. J Plankton Res. 2021; 43: 10.

Menden-Deuer S, Lessard EJ. Carbon to volume relationships for dinoflagellates, diatoms, and other protist plankton. Limnol Oceanogr. 2000; 45: 569–579. https://doi.org/10.4319/lo.2000.45.3.0569.

Moberg EA, Sosik HM. Distance maps to estimate cell volume from two-dimensional plankton images: Distance map cell volume algorithm. Limnol Oceanogr: Meth. 2012; 10: 278–288. https://doi.org/10.4319/lom.2012.10.278.

Mohsen A, Park J, Chen Y-A, Kawashima H, Mizuguchi K. Impact of quality trimming on the efficiency of reads joining and diversity analysis of Illumina paired-end reads in the context of QIIME1 and QIIME2 microbiome analysis frameworks. BMC Bioinform. 2019; 20: 581. https://doi.org/10.1186/s12859-019-3187-5.

Montagnes DJS, Franklin M. Effect of temperature on diatom volume, growth rate, and carbon and nitrogen content: Reconsidering some paradigms. Limnol Oceanogr. 2001; 46: 2008–2018. https://doi.org/10.4319/lo.2001.46.8.2008.

Morabito C, Bournaud C, Maës C, Schuler M, Aiese Cigliano R, Dellero Y, et al. The lipid metabolism in thraustochytrids. Prog Lipid Res. 2019; 76: 101007. https://doi.org/10.1016/j.plipres.2019.101007.

Nayak AR, Malkiel E, McFarland MN, Twardowski MS, Sullivan JM. A Review of Holography in the Aquatic Sciences: In situ Characterization of Particles, Plankton, and Small Scale Biophysical Interactions. Front Mar Sci. 2021; 7: 572147. https://doi.org/10.3389/fmars.2020.572147.

Nickolls J, Buck I, Garland M, Skadron K. Scalable Parallel Programming with CUDA, Queue, GPU Computing. 2008; 6: 40-53.

Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, et al. vegan: Community Ecology Package. 2013.

Olson RJ, Sosik HM. A submersible imaging-in-flow instrument to analyze nano-and microplankton: Imaging FlowCytobot: In situ imaging of nano- and microplankton. Limnol Oceanogr: Methods. 2007; 5: 195–203.

Otsu NA. threshold selection method from gray-level histogram. IEEE Trans Syst Man Cybern. 1979; 9: 62–66.

Orenstein EC, Beijbom O, Peacock EE, Sosik HM. WHOI-plankton-a large scale fine grained visual recognition benchmark dataset for plankton classification. arXiv preprint. 2015; doi: arXiv:1510.00745v1.

Orenstein EC, Beijbom O. Transfer Learning and Deep Feature Extraction for Planktonic Image Data Sets. Proc IEEE Int Conf Comput Vis. 2017; 1082–1088, doi: 10.1109/WACV.2017.125.

[a]Orenstein EC, Ratelle D, Briseño-Avena C, Carter ML, Franks PJS, Jaffe JS, Roberts PLD. The Scripps Plankton Camera system: A framework and platform for in situ microscopy. Limnol Oceanogr: Meth. 2020; 18: 681–695. https://doi.org/10.1002/lom3.10394.

[b]Orenstein EC, Kenitz KM, Roberts PLD, Franks PJS, Jaffe JS, Barton AD. Semi- and fully supervised quantification techniques to improve population estimates from machine classifiers. Limnol Oceanogr: Meth. 2020; 18: 739–753. https://doi.org/10.1002/lom3.10399.

Pan J, del Campo J, Keeling PJ. Reference Tree and Environmental Sequence Diversity of Labyrinthulomycetes. J Eukaryotic Microbiol. 2017; 64: 88–96. https://doi.org/10.1111/jeu.12342.

Pastore VP, Zimmerman TG, Biswas SK, Bianco S. Annotation-free learning of plankton for classification and anomaly detection. Sci Rep. 2020; 10: https://doi.org/10.1038/s41598-020-68662-3.

Pearson K. Mathematical contributions to the theory of evolution — on a form of spurious correlation which may arise when indices are used in the measurement of organs. Proc R Soc Lond. 1897; 60: 489–498. doi:10.1098/rspl.1896.0076.

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, *et al*. Scikit-learn: Machine Learning in Python. Machine Learning in Python. 2011; 12: 2825-2830.

Pepin P, Colbourne E, Maillet G. Seasonal patterns in zooplankton community structure on the Newfoundland and Labrador Shelf. Prog Oceanogr. 2011; 91: 273–285. https://doi.org/10.1016/j.pocean.2011.01.003.

Picheral M, Guidi L, Stemmann L, Karl DM, Iddaoud G, Gorsky G. The Underwater Vision Profiler 5: An advanced instrument for high spatial resolution studies of particle size spectra and zooplankton: Underwater vision profiler. Limnol Oceanogr: Methods. 2010; 8: 462–473. https://doi.org/10.4319/lom.2010.8.462.

Provost F. Machine learning from imbalanced data sets 101. Proceedings of the AAAI-2000 Workshop on Imbalanced Data Sets. 2000.

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, *et al*. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. Nucleic Acids Res. 2012; 41: D590–D596. https://doi.org/10.1093/nar/gks1219.

Quinn TP, Erb I, Gloor G, Notredame C, Richardson MF, Crowley TM. A field guide for the compositional analysis of any-omics data. GigaScience. 2019; 8: giz107. https://doi.org/10.1093/gigascience/giz107.

Raes EJ, Bodrossy L, van de Kamp J, Bissett A, Ostrowski M, Brown MV, *et al*. Oceanographic boundaries constrain microbial diversity gradients in the South Pacific Ocean. PNAS. 2018; 115: E8266–E8275. https://doi.org/10.1073/pnas.1719335115.

Raghukumar S. Ecology of the marine protists, the Labyrinthulomycetes (Thraustochytrids and Labyrinthulids). European Journal of Protistology. 2002; 38: 127–145. https://doi.org/10.1078/0932-4739-00832.

R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2019: URL https://www.R-project.org/.

Reichstein M, Camps-Valls G, Stevens B, Jung M, Denzler J, Carvalhais N, Prabhat. Deep learning and process understanding for data-driven Earth system science. Nature. 2019; 566: 195–204. https://doi.org/10.1038/s41586-019-0912-1.

Richardson PL. Florida Current, Gulf Stream, and Labrador Current. John H, in Encyclopedia of Ocean Sciences. 2001; edited by J.H. Steele, pp. 1054-1064, Academic Press, Oxford, doi: http://dx.doi.org/10.1006/rwos.2001.0357.

Rivenson Y, Wu Y, Ozcan A. Deep learning in holography and coherent imaging. Light: Sci Appli. 2019; 8: https://doi.org/10.1038/s41377-019-0196-0.

Rodríguez-Ramos T, Dornelas M, Marañón E, Cermeño P. Conventional sampling methods severely underestimate phytoplankton species richness. J Plankton Res. 2014; 36: 334–343. https://doi.org/10.1093/plankt/fbt115.

Rognes T, Flouri T, Nichols B, Quince C, Mahe F. VSEARCH: a versatile open source tool for metagenomics PeerJ. 2016; 4:e2584 doi: 10.7717/peerj.2584 https://doi.org/10.7717/peerj.2584.

Roselli L, Paparella F, Stanca E, Basset A. New data-driven method from 3D confocal microscopy for calculating phytoplankton cell biovolume: Calculating Phytoplankton Cell Biovolume. J Microscopy. 2015; 258: 200–211. https://doi.org/10.1111/jmi.12233.

Rotermund LM, Samson J, Kreuzer HJ. A Submersible Holographic Microscope for 4-D In-Situ Studies of Micro-Organisms in the Ocean with Intensity and Quantitative Phase Imaging. J Mar Sci: Res Develop. 2015; 6: doi: 10.4172/2155-9910.1000181.

Ryabov A, Kerimoglu O, Litchman E, Olenina I, Roselli L, Basset A, Stanca E, Blasius B. Shape matters: The relationship between cell geometry and diversity in phytoplankton. Ecol Lett. 2021; 24: 847–861. https://doi.org/10.1111/ele.13680.

Saccà A. Methods for the estimation of the biovolume of microorganisms: A critical review: Methods biovolume of microorganisms. Limnol Oceanogr: Meth. 2017; 15: 337–348. https://doi.org/10.1002/lom3.10162.

Saito T, Rehmsmeier M. The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets. PLoS ONE. 2015; 10(3): e0118432.

Sanders HL. Marine Benthic Diversity: A Comparative Study. Am Nat. 1968: 102: 243–282. https://doi.org/10.1086/282541.

Sargent EC, Hitchcock A, Johansson SA, Langlois R, Moore CM, LaRoche J, Poulton AJ, Bibby T. S. Evidence for polyploidy in the globally important diazotroph *Trichodesmium*. FEMS Microbiol Lett. 2016; 363: fnw244. https://doi.org/10.1093/femsle/fnw244.

Sauzède R, Lavigne H, Claustre H, Uitz J, Schmechtig C, D'Ortenzio F, Guinet C, Pesant S. Vertical distribution of chlorophyll a concentration and phytoplankton community composition from in situ fluorescence profiles: A first database for the global ocean. Earth Syst. Sci. Data. 2015; 7: 261–273,

Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint. 2014; doi: arXiv:1409.1556v6.

Sharif Razavian A, Azizpour H, Sullivan J, Carlsson S. CNN features off-the-shelf: an astounding baseline for recognition. arXiv preprint. 2014; doi: arXiv:1403.6382v3.

Sheng J, Malkiel E, Katz J. Digital holographic microscope for measuring three-dimensional particle distributions and motions. Appl Opt. 2006; 45: 3893-3901.

Schmidhuber J. Deep Learning in Neural Networks: An Overview. Neural Netw. 2015; 61: 85–117.

Schnars U, Jüptner WP. Digital recording and numerical reconstruction of holograms. Meas Sci Technol. 2002; 13: R85. doi: 10.1088/0957- 0233/13/9/201.

Sieracki C, Sieracki M, Yentsch C. An imaging-in-flow system for automated analysis of marine microplankton. Marine Ecology Progress Series. 1998; 168: 285–296. https://doi.org/10.3354/meps168285.

Silverman JD, Roche K, Mukherjee S, David LA. Naught all zeros in sequence count data are the same. Comput Struct Biotech J. 2020; 18: 2789–2798. https://doi.org/10.1016/j.csbj.2020.09.014.

Sherr E, Sherr B. Role of microbes in pelagic food webs: A revised concept. Limnol Oceanogr. 1988; 33(5): 1225–1227. https://doi.org/10.4319/lo.1988.33.5.1225.

Sogin ML, Morrison HG, Huber JA, Welch MD, Huse SM, Neal PR, et al. Microbial diversity in the deep sea and the underexplored "rare biosphere." PNAS. 2006; 103: 12115–12120. doi: 10.1073/pnas.0605127103.

Sun H, Benzie PW, Burns N, Hendry DC, Player MA, Watson J. Underwater digital holography for studies of marine plankton. Philos Trans R Soc A: Math, Phys Eng Sci. 2008; 366: 1789–1806.

Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. J Mach Learn Res. 2014; 15: 1929-1958.

Strigl D, Kofler K, Podlipnig S. Performance and Scalability of GPU-Based Convolutional Neural Networks. 2010 18th Euromicro Conference on Parallel, Distributed and Network-Based Processing. 2010; 317–324. https://doi.org/10.1109/PDP.2010.43.

Stoeck T, Bass D, Nebel M, Christen R, Jones MDM, Breiner HW, Richards TA. Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. Mol Ecol. 2010; 19: 21–31. https://doi.org/10.1111/j.1365-294X.2009.04480.x.

Sverdrup HU. Oceanography for Meteorologists. Prentice Hall. 1942; 246 pp.

Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the Inception Architecture for Computer Vision. arXiv preprint. 2015; doi: arXiv:1512.00567v3.

Tapics T, Gregory-Eaves I, Huot Y. The private life of *Cystodinium*: *In situ* observation of its attachments and population dynamics. J Plankton Res. 2021; fbab025. https://doi.org/10.1093/plankt/fbab025.

Tharwat A. Classification assessment methods. Appl Comput Informa. 2018; 1-3. https://doi.org/10.1016/j.aci.2018.08.003.

Therriault J-C, Petrie B, Pepin P, Gagnon J, Gregory D, Helbig J, *et al.* Proposal for a northwest Atlantic zonal monitoring program. Can Tech Rep Hydrogr Ocean Sci. 1998; 194: vii+57p.

Tomas CR. Identifying Marine Phytoplankton. San Diego: Academic Press. 1997; ISBN 0-12-693018-X-XV. 858pp.

Torti A, Lever MA, Jorgensen BB. Origin, dynamics, and implications of extracellular DNA pools in marine sediments. Mar Genomics. 2015; 24: 185–196.

Tréguer P, Bowler C, Moriceau B, Dutkiewicz S, Gehlen M, Aumont O, *et al.* Influence of diatom diversity on the ocean biological carbon pump. Nature Geo. 2018; 11: 27–37. https://doi.org/10.1038/s41561-017-0028-x.

Trudnowska E, Lacour, L, Ardyna M, Rogge A, Irisson JO, Waite AM, Babin M, Stemmann L. Marine snow morphology illuminates the evolution of phytoplankton blooms and determines their subsequent vertical export. Nature Communications. 2021; 12: 2816. https://doi.org/10.1038/s41467-021-22994-4.

van Noord N, Postma E. Learning scale-variant and scale-invariant features for deep image classification. Pattern Recognit. 2017; 61: 583–592.

Vogel, S. Life in moving fluids: The physical biology of flow. Princeton University Press. 1981; 320 pp.

Vogt M, O'Brien C, Peloquin J, Schoemann V, Breton E, Estrada M, *et al.* Global marine plankton functional type biomass distributions: Phaeocystis spp. [Preprint]. Oceanography – Biological. 2012; https://doi.org/10.5194/essdd-5-405-2012.

Vollath D. Automatic focusing by correlative methods. J Microscopy. 1987; 147: 279-288.

Walcutt NL, Knörlein B, Cetinić I, Ljubesic Z, Bosak S, Sgouros T, *et al.* Assessment of holographic microscopy for quantifying marine particle size and concentration. Limnol Oceanogr: Methods. 2020; doi: 10.1002/lom3.10379.

Weiss K, Khoshgoftaar TM, Wang D. A survey of transfer learning. J Big Data. 2016; 3: https://doi.org/10.1186/s40537-016-0043-6.

Weiss S, Xu ZZ, Peddada S, Amir A, Bittinger K, Gonzalez, A, *et al*. Normalization and microbial differential abundance strategies depend upon data characteristics. Microbiome. 2017; 5: 27. https://doi.org/10.1186/s40168-017-0237-y.

[a]Willis A, Bunge J. Estimating diversity via frequency ratios: Estimating Diversity via Ratios. Biometrics. 2015; 71: 1042–1049. https://doi.org/10.1111/biom.12332.

[b]Willis A, Bunge J, Whitman T. Inference for changes in biodiversity. ArXiv:1506.05710 [q-Bio, Stat]. 2015; http://arxiv.org/abs/1506.05710.

Willis A, Bunge J, Whitman T. Improved detection of changes in species richness in high diversity microbial communities. J R Stat Soc: C-Appl Stat. 2017; 66: 963–977. https://doi.org/10.1111/rssc.12206.

Willis AD. Rarefaction, Alpha Diversity, and Statistics. Front Microbiol. 2019; 10: 2407. https://doi.org/10.3389/fmicb.2019.02407.

Whitt C, Pearlman J, Polagye B, Caimi F, Muller-Karger F, Copping A, *et al.* Future Vision for Autonomous Ocean Observations. Front Mar Sci. 2020; 7: 697. https://doi.org/10.3389/fmars.2020.00697.

Woese, CR, Fox GE. Phylogenetic structure of the prokaryotic domain: The primary kingdoms. PNAS. 1977; 74: 5088–5090. https://doi.org/10.1073/pnas.74.11.5088.

Wright SW, Jeffrey SW. Pigment Markers for Phytoplankton Production. In J. K. Volkman (Ed.), Marine Organic Matter: Biomarkers, Isotopes and DNA (Vol. 2N, pp. 71–104). Springer-Verlag. 2006; https://doi.org/10.1007/698_2_003.

Xu W, Jericho MH, Meinertzhagen IA, Kreuzer HJ. Digital in-line holography for biological applications. Proc Natl Acad Sci USA. 2001; 98: 11301–11305.

Yadav S, Shukla S. Analysis of k-Fold Cross-Validation over Hold-Out Validation on Colossal Datasets for Quality Classification. 2016 IEEE 6th International Conference on Advanced Computing (IACC). 2016; 78–83.

Yosinski J, Clune J, Bengio Y, Lipson H. How transferable are features in deep neural networks? arXiv preprint. 2014; doi: arXiv:1411.1792v1.

Zetsche E, Mallahi A, Dubois F, Yourassowsky C, Kromkamp J, Meysman FJR. Imaging-in-Flow: Digital holographic microscopy as a novel tool to detect and classify nanoplanktonic organisms. Limnol Oceanogr: Methods. 2014; 12: 757–775.

Zetsche EM, Baussant T, Meysman FJR, van Oevelen D. Direct Visualization of Mucus Production by the Cold-Water Coral *Lophelia pertusa* with Digital Holographic Microscopy. PLoS ONE. 2016; 11: e0146766. https://doi.org/10.1371/journal.pone.0146766.

Zhu F, Massana R, Not F, Marie D, Vaulot D. Mapping of picoeucaryotes in marine ecosystems with quantitative PCR of the 18S rRNA gene. FEMS Microbiol Ecol. 2005; 52: 79–92. https://doi.org/10.1016/j.femsec.2004.10.006.
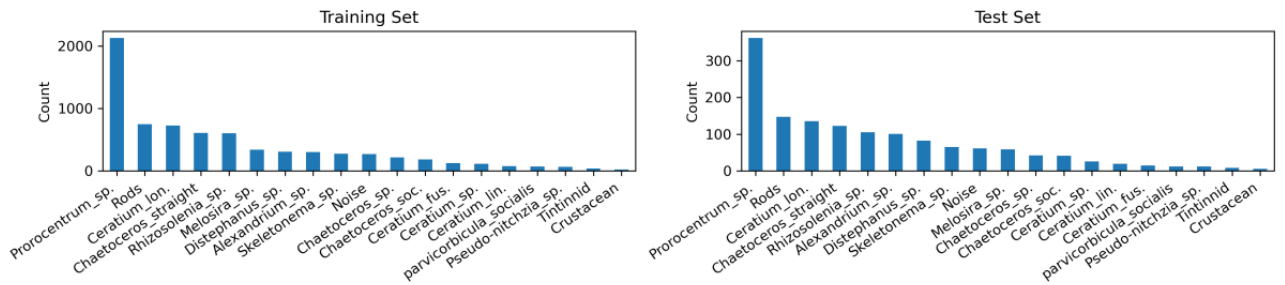
Figure A.1. Left to right, distribution of taxa abundance for training set— where the distribution ratios are maintained during stratified cross validation— and the test set.
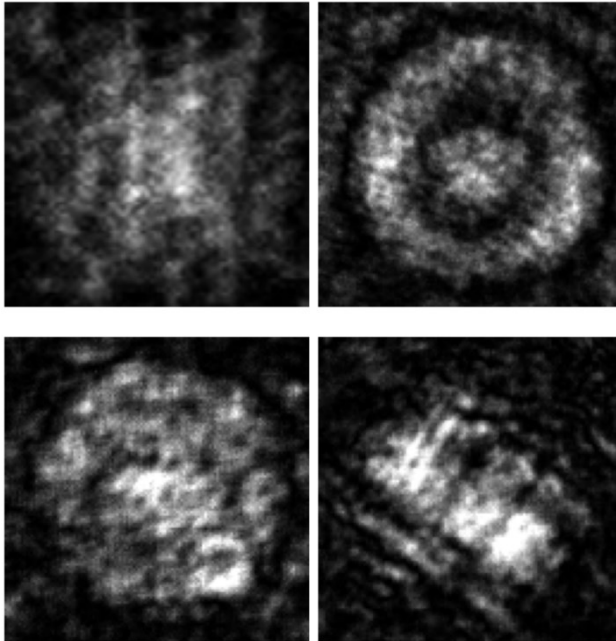
Figure A.2. Four classified noise objects with no resolvable features. Image artefacts are a challenge for any imaging system operating in real time, where both the imaging mode and environment can create non-biological objects.

Table A.1. The reference paper of four CNNs, their convolutional layers, the weighted layers that are changed during backpropagation, and broad overview of their key features.

Model	Reference	Convolutions	Trainable Parameters	Description
VGG16	Simonyan and Zisserman (2014)	13	138,357,544	Stacked sequence of 3×3 convolution filters and max pooling layers. Three fully connected layers compile features. Runner up in ILSVRC 2014.
InceptionV3	Szegedy *et al.* (2016)	48	23,851,784	Parallel convolution filters and pooling layers to extract features with different filter sizes (multiple sized receptive fields) known as depth wise separable convolutions. Parallel filters recognize features of multiple sizes within a layer. The Inception module (GoogLeNet) winner in ILSVRC 2014.
ResNet50V2	He *et al.* (2016)	50	25,613,800	ResNets composed of sequential layers in blocks, with identity connections between blocks that sum the output of each block with the previous block. These shortcut connections are most important during learning— the backpropagation algorithm— where error gradients can be conveyed to the earliest layers. ResNets overcame vanishing gradients in learning deeper, more complex models, winning ILSVRC 2015. ResNetV2 only applies non-linearity (ReLu) before convolutions within a block.
Xception	Chollet (2017)	71	22,910,480	Extreme Inception modules perform 1×1 convolution before depth wise separable convolutions with shortcut connections between convolution blocks analogous to ResNets. The top accuracy from the ILSVRC outperforms VGG16, InceptionV3, and ResNet152
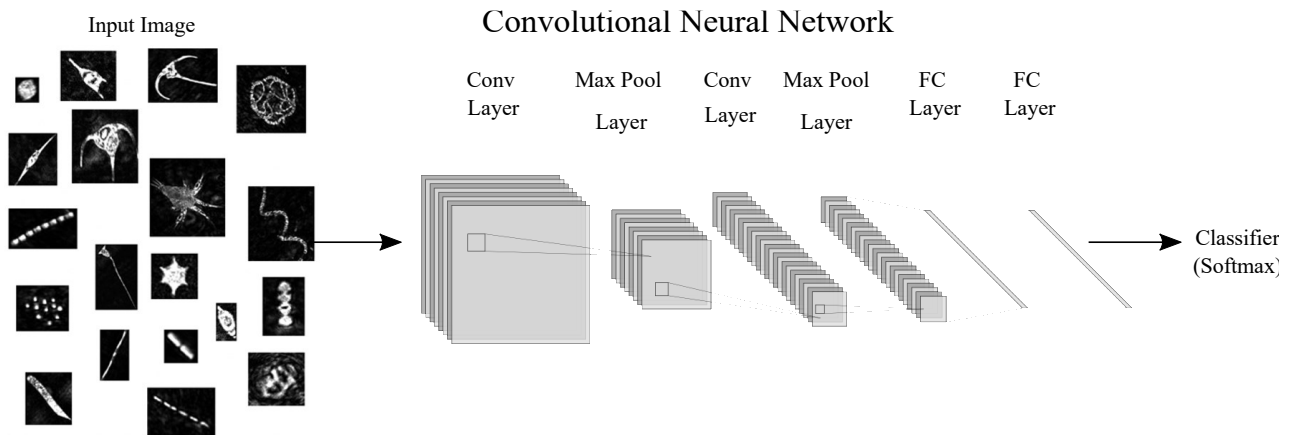
Figure A.3. Network architecture for basic CNN. The inputs are the resized (128×128) images containing in-focus objects. CNN schematic created in http://alexlenail.me/NN-SVG/LeNet.html.

Table A.2. Total time and memory expended for training and evaluating each model averaged for feature extraction and fine tuning.

Model	Time (minutes)	Memory Consumption (GB)
VGG16	83	3.61
InceptionV3	79	0.83
ResNetV2	74	1.28
Xception	71	1.24

Table A.3. Average performance of each model for each threshold metric on the test set for each fold.

|		Threshold Metrics (%)			
	Model	Accuracy	Precision	Recall	F1-Score
Feature Extraction	VGG16	88.2 ± 0.6	88.4 ± 0.5	88.1 ± 0.6	87.8 ± 0.9
	InceptionV3	79.8 ± 1.2	77.2 ± 2.0	82.6 ± 0.7	78.9 ± 1.4
	ResNetV2	88.2 ± 0.8	88.6 ± 0.7	88.1 ± 07	87.9 ± 0.9
	Xception	90.1 ± 0.6	89.8 ± 0.9	90.7 ± 0.4	89.8 ± 0.7
Retraining Deeper Layers	VGG16	88.7 ± 0.3	88.5 ± 0.3	88.8 ± 0.6	88.3 ± 0.4
	InceptionV3	79.7 ± 0.09	77.8 ± 1.9	81.9 ± 0.03	79.0 ± 1.2
	ResNetV2	87.9 ± 1.0	87.6 ± 1.3	88.6 ± 0.7	87.5 ± 1.1
	Xception	90.2 ± 1.0	89.9 ± 2.0	90.8 ± 1.0	90.0 ± 1.8
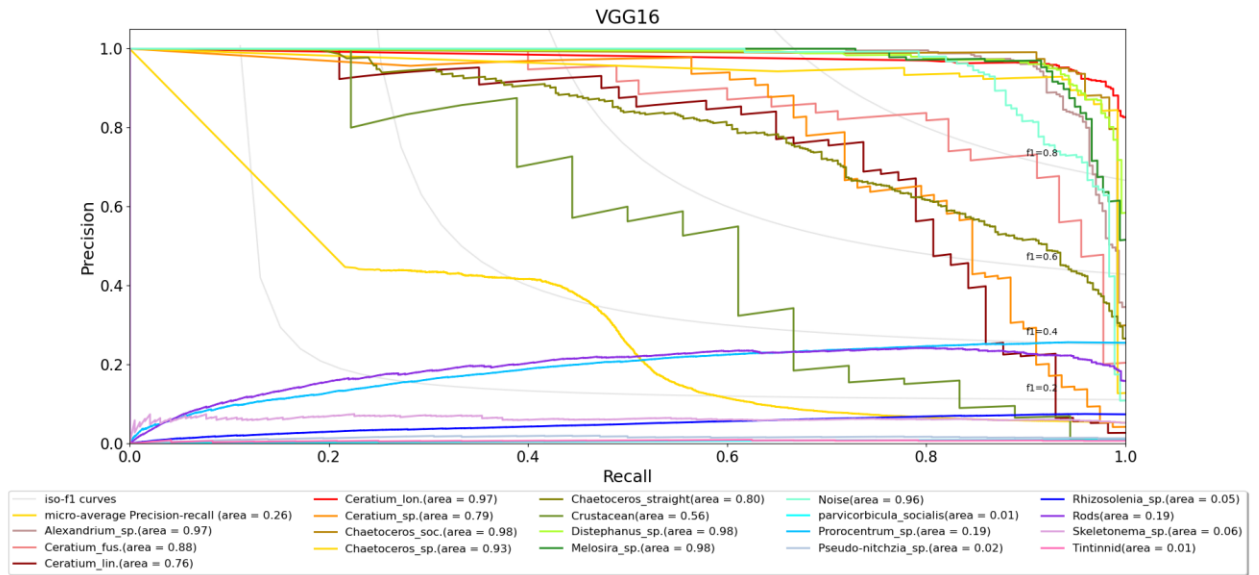
Figure A.4. Precision-recall curves of the InceptionV3, with iso-curves for their harmonic mean F1-score, and the area under the curve (AUC-PR).
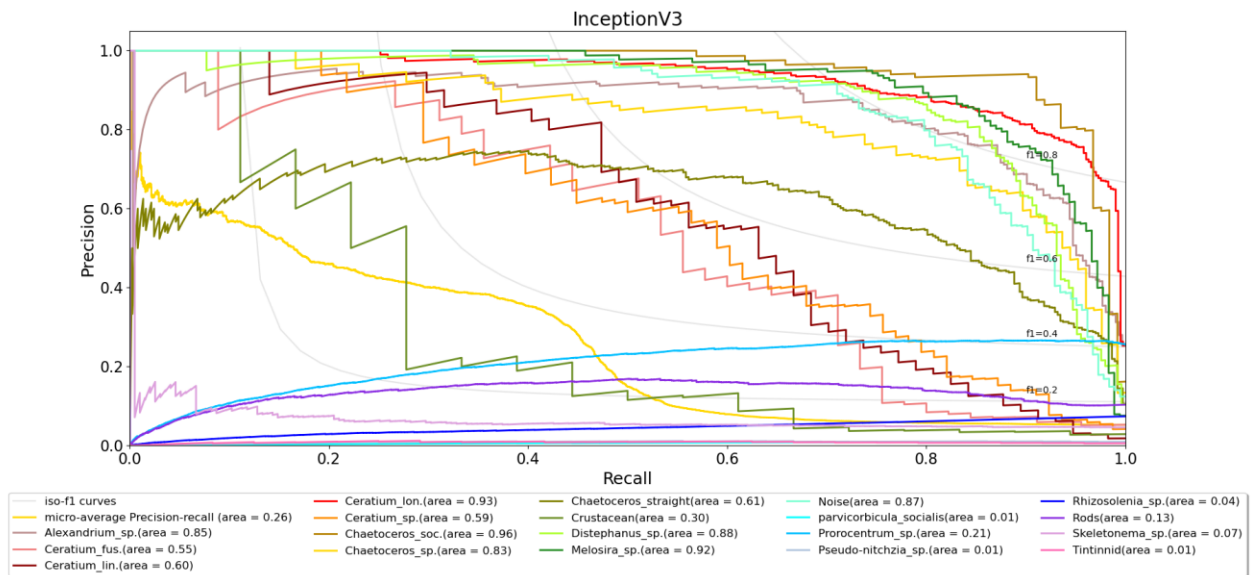
Figure A.5. Precision-recall curves of the InceptionV3, with iso-curves for their harmonic mean F1-score, and the area under the curve (AUC-PR).
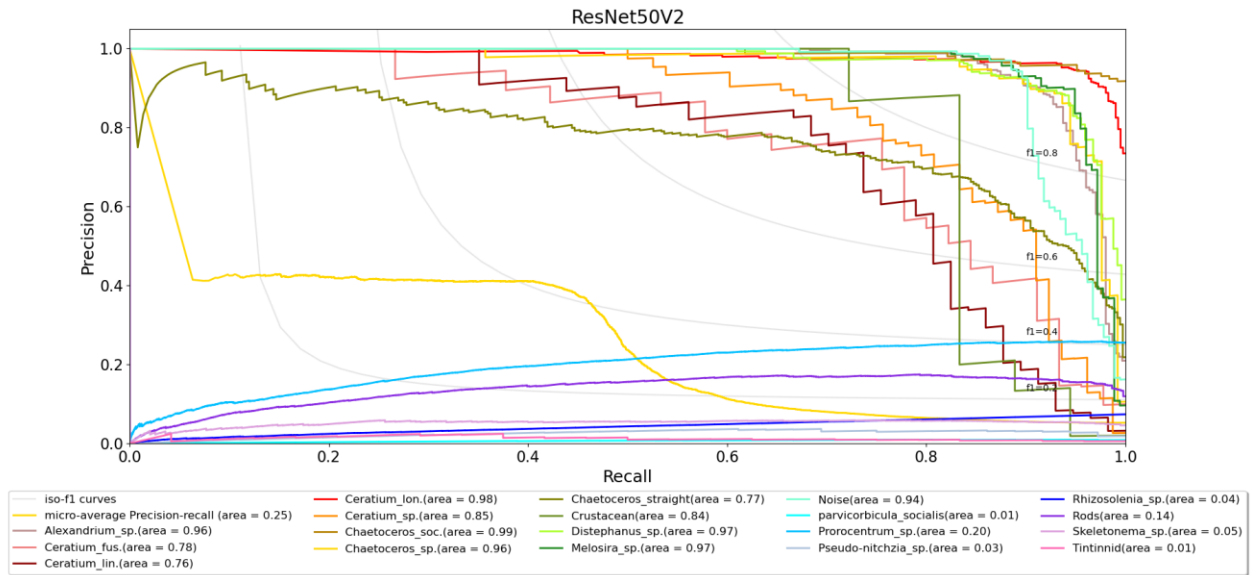
Figure A.6. Precision-recall curves of the InceptionV3, with iso-curves for their harmonic mean F1-score, and the area under the curve (AUC-PR).
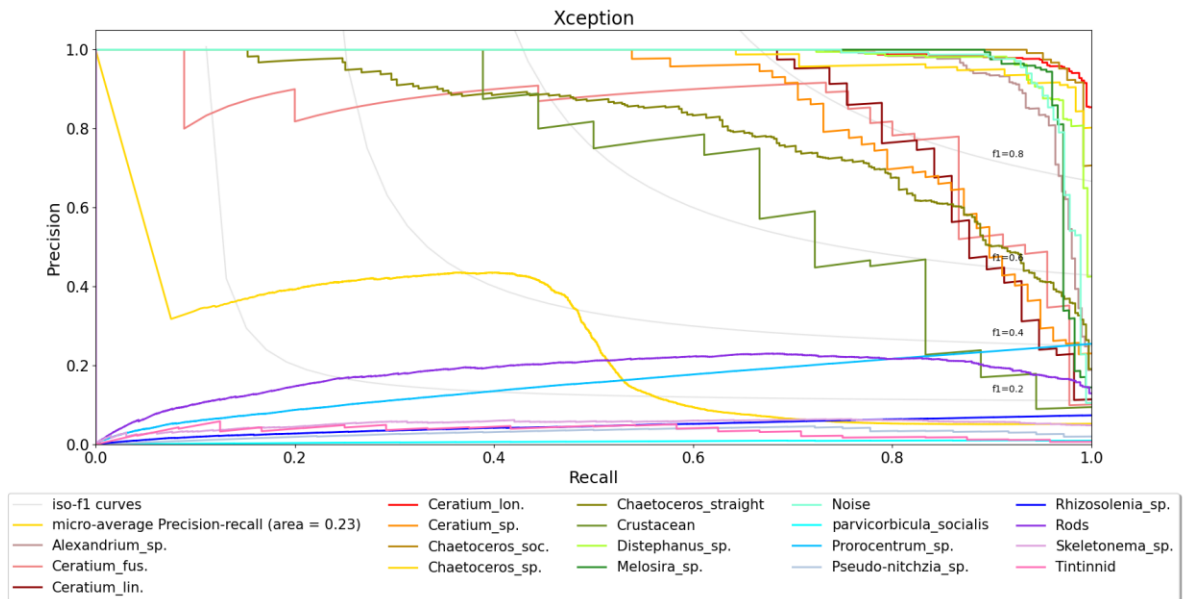

Figure A.7. Precision-recall curves of the Xception model for each class, with iso-curves for their harmonic mean F1-score, and the area under the curve (AUC-PR).

**APPENDEIX B — COPYRIGHT AGREEMENT FOR CHAPTER 2**

Chapter 2 is a pre-copyedited, author-produced document of an article accepted for publication in BMC Ecology and Evolution following peer review. The version of record "MacNeil et al. (2021) Plankton classification with high-throughput submersible holographic microscopy and transfer learning is now available online: https://doi.org/10.1186/s12862-021-01839-0.
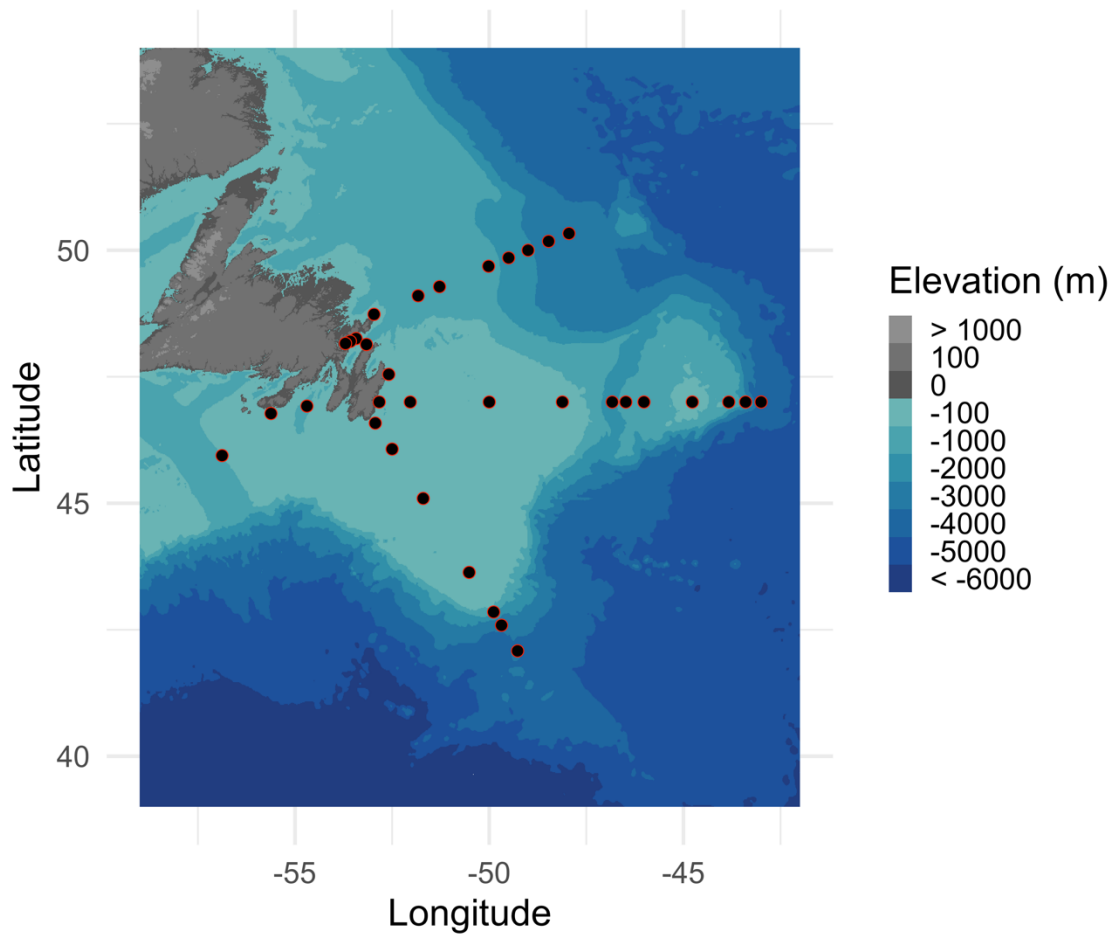
Figure C.1. The sample locations for the full imaging and metabarcoding dataset, totalling 51 samples from 34 stations.
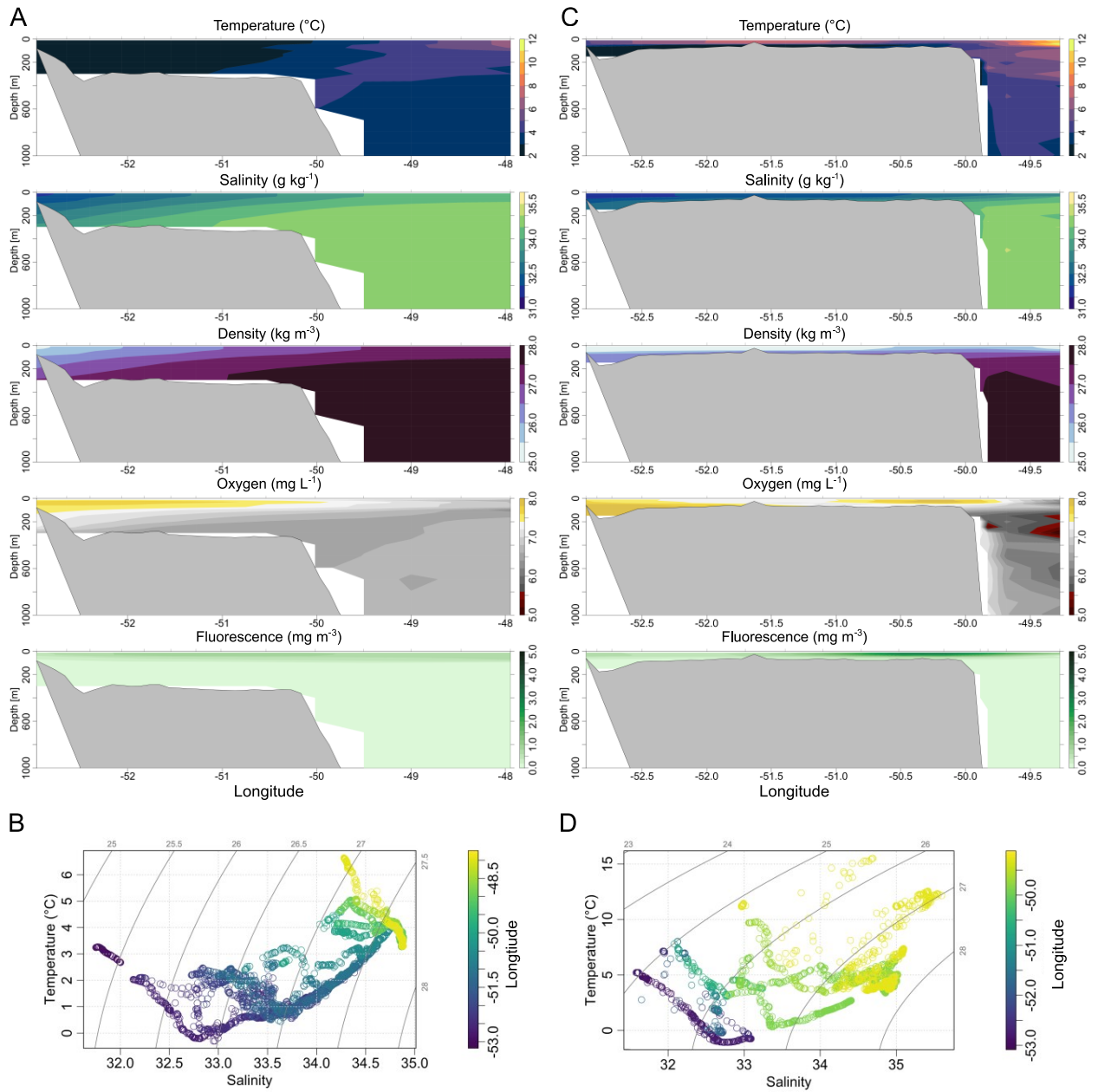
Figure C.2. The A-B) Bonavista Banks and C-D) SE Grand Banks section plots down to 1000 m depth with corresponding temperature-salinity diagrams across the longitudinal gradient of the shelf. The location of each profile cast is indicated as tick marks on the top x-axis of each section plot. The isopycnals on the temperature-salinity diagrams indicate constant density.
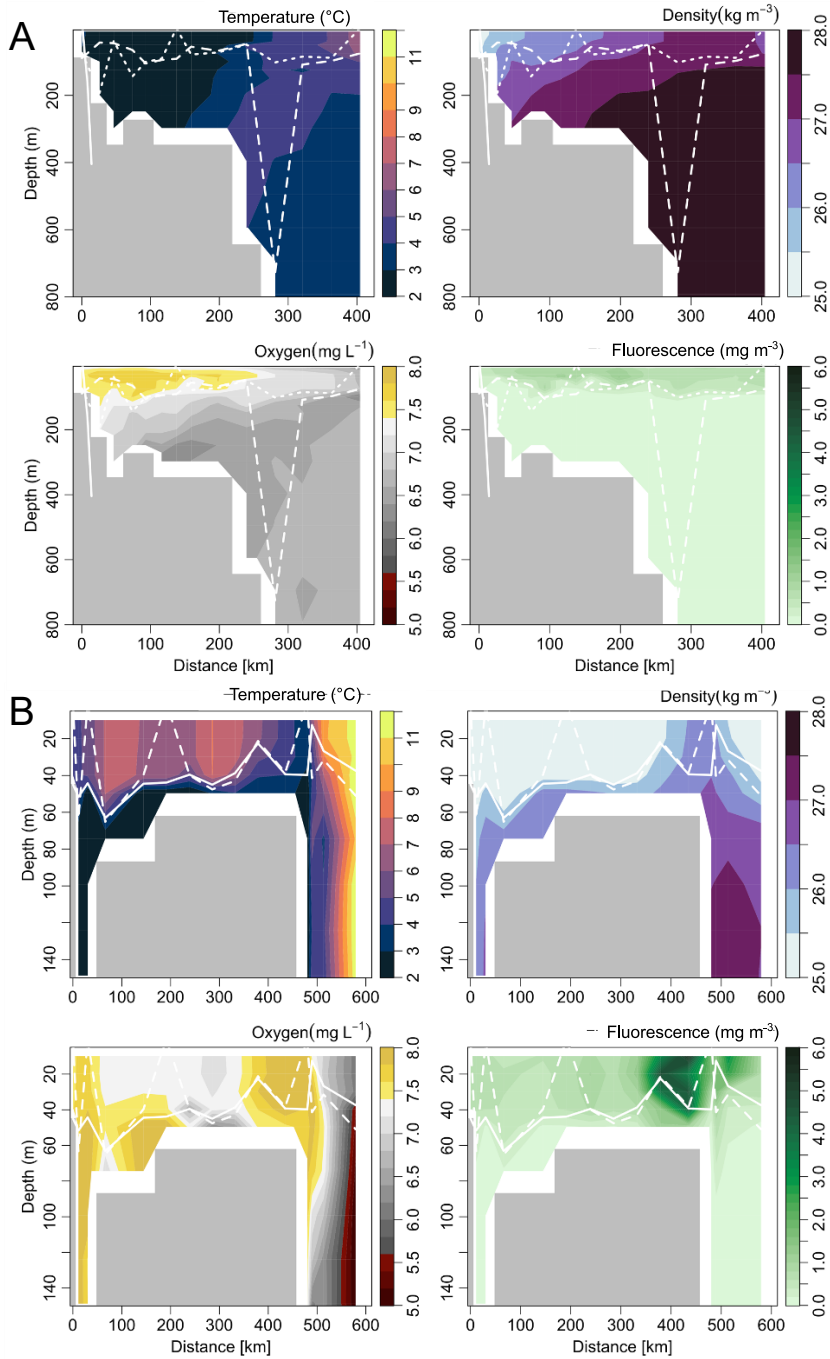
Figure C.3. Mixed layer depth estimates for both criterion (short dashed lines) and derivative (long dashed lines) methods on the A) Bonavista Banks and B) SE Grand Banks. The lines are plotted over each section plot, however high-resolution bathymetry could not be visualized when mixed layer depth estimates are plotted. The sections also contain subtle differences to Figure 3.3 and Figure C.1 because it is unsmoothed, but the general water column structure remains the same.