



HAL
open science

Machine learning techniques to characterize functional traits of plankton from image data

Eric C Orenstein, Sakina-Dorothee Ayata, Frédéric Maps, Tristan Biard, Érica C Becker, Fabio Benedetti, Thibault de Garidel-Thoron, Jeffrey S Ellen, Filippo Ferrario, Sarah L C Giering, et al.

► To cite this version:

Eric C Orenstein, Sakina-Dorothee Ayata, Frédéric Maps, Tristan Biard, Érica C Becker, et al.. Machine learning techniques to characterize functional traits of plankton from image data. 2021. hal-03482282v1

HAL Id: hal-03482282

<https://hal.univ-lorraine.fr/hal-03482282v1>

Preprint submitted on 15 Dec 2021 (v1), last revised 30 Jun 2022 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Machine learning techniques to characterize functional traits of plankton from image data

Eric C. Orenstein¹, Sakina-Dorothee Ayata¹, Frédéric Maps², Tristan Biard³, Érica C. Becker⁴, Fabio Benedetti⁵, Thibault de Garidel-Thoron⁶, Jeffrey S. Ellen⁷, Filippo Ferrario², Sarah L.C. Giering⁸, Tamar Guy-Haim⁹, Laura Hoebeke¹⁰, Morten Hvitfeldt Iversen¹¹, Thomas Kiørboe¹², Jean-François Lalonde¹³, Arancha Lana¹⁴, Martin Laviale¹⁵, Fabien Lombard¹, Tom Lorimer¹⁶, Séverine Martini¹⁷, Albin Meyer¹⁵, Klas Ove Möller¹⁸, Barbara Niehof¹, Mark D. Ohman⁷, Cédric Pradalier¹⁹, Jean-Baptiste Romagnan²⁰, Simon-Martin Schröder²¹, Virginie Sonnet²², Heidi M. Sosik²³, Lars S. Stemmann¹, Michiel Stock¹⁰, Tuba Terbiyik-Kurt²⁴, Nerea Valcárcel-Pérez²⁵, Laure Vilgrain¹, Guillaume Wacquet²⁶, Anya M. Waite²⁷, Jean-Olivier Irisson¹

Keywords: Functional trait-based ecology, plankton, imaging, machine learning

¹ Sorbonne Université, CNRS, Laboratoire d'Océanographie de Villefranche, Villefranche-sur-Mer, France

² Département de Biologie, Université Laval, Québec, Canada

³ Laboratoire d'Océanologie et de Géosciences, Univ. Littoral Côte d'Opale, Univ. Lille, CNRS, UMR 8187, Wimereux, France

⁴ Universidade Federal de Santa Catarina, UFSC, Florianópolis, Santa Catarina, Brazil

⁵ ETH Zürich, Institute of Biogeochemistry and Pollutant Dynamics, Zürich, Switzerland

⁶ Aix-Marseille Université, CNRS, IRD, Coll. de France, INRAE, CEREGE, Aix en Provence, France

⁷ Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA, USA

⁸ Ocean Biogeosciences, National Oceanography Centre, Southampton, United Kingdom

⁹ National Institute of Oceanography, Israel Oceanographic and Limnological Research, Haifa, Israel

¹⁰ KERMIT, Department of Data Analysis and Mathematical Modelling, Ghent University, Ghent, Belgium

¹¹ Wegener Institute Helmholtz Center for Polar and Marine Research, Bremerhaven, Germany

¹² Centre for Ocean Life, DTU-Aqua, Technical University of Denmark, Denmark

¹³ Laboratoire de vision et systèmes numériques, Université Laval, Québec, Canada

¹⁴ Institut Mediterrani d'Estudis Avançats (IMEDEA, UIB-CSIC), Balearic Islands, Spain.

¹⁵ Université de Lorraine, CNRS, LIEC, Metz, France

¹⁶ Eawag, Überlandstrasse 133, CH-8600 Dübendorf, Switzerland

¹⁷ Aix Marseille Univ., Université de Toulon, CNRS, IRD, MIO UM, Marseille, France

¹⁸ Helmholtz-Zentrum Hereon, Institute of Coastal Ocean Dynamics, Geesthacht, Germany

¹⁹ GeorgiaTech Lorraine, CNRS IRL GT-CNRS, Metz, France

²⁰ IFREMER, Centre Atlantique, unité Ecologie et Modèles pour l'Halieutique (EMH), Rue de l'île d'Yeu, BP, Nantes Cedex, France

²¹ Kiel University, Kiel, Germany

²² Graduate School of Oceanography, University of Rhode Island, Narragansett, RI, USA

²³ Woods Hole Oceanographic Institution, Woods Hole, MA USA

²⁴ Cukurova University, Faculty of Fisheries, Department of Basic Sciences, Adana, Turkey

²⁵ Centro Oceanográfico de Málaga, IEO, CSIC. 29640 Fuengirola, Spain

²⁶ IFREMER, Laboratoire Environnement Ressources, Boulogne-sur-Mer, France

²⁷ Ocean Frontier Institute, Dalhousie University, Halifax, Nova Scotia, Canada

Abstract

Plankton imaging systems supported by automated classification and analysis have improved ecologists' ability to observe aquatic ecosystems. Today, we are on the cusp of reliably tracking plankton populations with a suite of lab-based and in situ tools, collecting imaging data at unprecedentedly fine spatial and temporal scales. But these data streams have potential well beyond examining the abundances of different taxa; the individual images themselves contain a wealth of information on functional traits. Here we outline traits that could be measured from image data, suggest computer vision approaches to extract functional trait information from the images, and discuss promising avenues for novel studies. The approaches we discuss are data agnostic and are broadly applicable to other aquatic or terrestrial organisms.

Introduction

Over the past 40 years, in situ and laboratory-based plankton imaging systems have developed rapidly and proven capable of describing the distribution of organisms on unprecedented spatio-temporal scales, from microns to ocean basins and from seconds to decades (Ortner et al. 1979, Davis et al. 1992, Olson & Sosik 2007, Picheral et al. 2010, Sieracki et al. 2010, Gorsky et al. 2010, Ohman et al. 2018). There is a huge variety of plankton imaging devices with a range of technical and ecological applications: living or preserved samples; in the laboratory or in the field; snapshots or continuous video; color or grayscale; coherent or incoherent illumination (Benfield et al. 2007, Lombard et al. 2019). Each approach is limited by engineering constraints to a specific portion of the plankton size spectrum and taxonomic tree. Combining such modalities would allow aquatic ecologists to study the full size spectrum of plankton communities and provide a holistic view of the planktonic ecosystem (Stemmann & Boss 2012, Romagnan et al. 2015, Lombard et al. 2019). These imaging systems enable the simultaneous estimation of the taxonomic structure of a community and measurement of morphological characteristics of individual organisms. The amount of recorded data has increased immensely, both in quantity and in quality, due to improvements in instrumentation – higher pixel density sensors, faster bus speed to move bits from the sensors to the processors, and more powerful embedded computers – and developments in image post-processing. Moreover, using imaging techniques on historic plankton samples has given scientists the opportunity to revisit archived samples (García-Comas et al. 2011, Peacock et al. 2014).

The principal objective of plankton imaging has been to extract quantitative, population-level information through taxonomic recognition of individual images. Manual annotation of the data remains a non-trivial task that requires expert knowledge and lots of time, even with semi-automated recognition (Culverhouse et al. 2003, 2014, Grosjean et al. 2004). Labelling has remained the major bottleneck limiting scientists' ability to analyze image data or use it for monitoring purposes (MacLeod et al. 2010). Scientists have started to look towards supervised machine learning meth-

ods – algorithms that learn to annotate new data based on a set of human-generated training examples – to expedite classification efforts (Irisson et al. In press). Until the 2010s, such methods typically involved extracting hand-engineered numerical features from a curated set of training images followed by tuning an ensemble or margin-based classifier (Simpson et al. 1991, Blaschko et al. 2005, Sosik & Olson 2007, Gorsky et al. 2010). Computers were programmed to measure shape attributes – such as area, eccentricity, or aspect ratio – and texture metrics – like variance in grey levels – from the images. These routines were generally written for data collected by a specific system and required manual parameter tuning to ensure accurate measurements. Recently, aquatic ecologists have started exploring popular deep learning techniques which instead seek to learn representations of the organisms directly from the images (LeCun et al. 2015, González et al. 2017, Orenstein & Beijbom 2017, Luo et al. 2018, Ellen et al. 2019). While avoiding the need for defining clear shape and texture attributes, deep methods require an enormous amount of annotated training data to appropriately tune (Bengio et al. 2013, Sun et al. 2017).

Functional trait-based approaches (FTBAs) describing plankton ecosystems have been developed in parallel with advances in imaging systems (Litchman & Klausmeier 2008, Litchman et al. 2013, Kiørboe et al. 2018, Martini et al. 2021). Rather than considering the taxonomic identity of an organism, FTBAs instead characterize them by their specific combination of functional traits – characteristics of individual organisms that impact their fitness via resource acquisition, growth, reproduction, and survival (Violle et al. 2007). Traits information transcends taxa and may have several advantages over classical taxonomic information: (i) organisms are distributed in the environment mainly according to their traits, not their taxonomic classification; (ii) ecosystem functions depend on the traits of constituent organisms, not their taxonomic composition; and (iii) while aquatic ecosystems are inhabited by a myriad of species (e.g. hundreds of thousands of eukaryotic taxa in the upper ocean; de Vargas et al. 2015), key taxa-transcending traits are few (Martini et al. 2021). Trait-based descriptions therefore have the potential to encapsulate the complexity of aquatic ecosystems

in a concise number of metrics (Kjørboe et al. 2018).

The functional trait-based perspective suggests that obtaining information on the characteristics of individual organisms directly from images is an efficient alternative to sorting the data taxonomically (Martini et al. 2021). Indeed, efforts have already been made to reanalyze hand-engineered image features originally drawn for classification purposes. This approach has revealed cryptic patterns in plankton communities, such as the morphological characteristics of particulate matter in the global ocean and the feeding habits of marine copepods along an Arctic sea-ice gradient (Trudnowska et al. 2021, Vilgrain et al. 2021). Such studies are just beginning to combine high-throughput imaging systems and automated analysis techniques for trait-based studies. We believe there is much more to be done and learned by leveraging a suite of powerful new machine learning tools in the context of functional trait-based ecology.

This paper aims to highlight how imaging and machine learning might be combined to gain direct

access to plankton functional traits from individual images and, eventually, enrich our understanding of pelagic ecosystems. First, we detail traits that lend themselves to observation images, regardless of how the image has been acquired (in situ or in the lab, high throughput or piecewise acquisition, etc.). We distinguish between traits that have a clear visual signature and can be *measured* from the image itself and traits that are *inferred* from the image, taking into account additional assumptions or contextual variables. The second section is dedicated to machine learning methods that could be used to estimate functional traits from images. We encourage ecologists to adopt an “evaluation first” design paradigm to inform data annotation and algorithm selection. We conclude with several forward-looking suggestions regarding potential statistical analyses of the traits collected by machine learning algorithms, future instrument design to collect functional traits data, and extensions of these methods to other organisms.

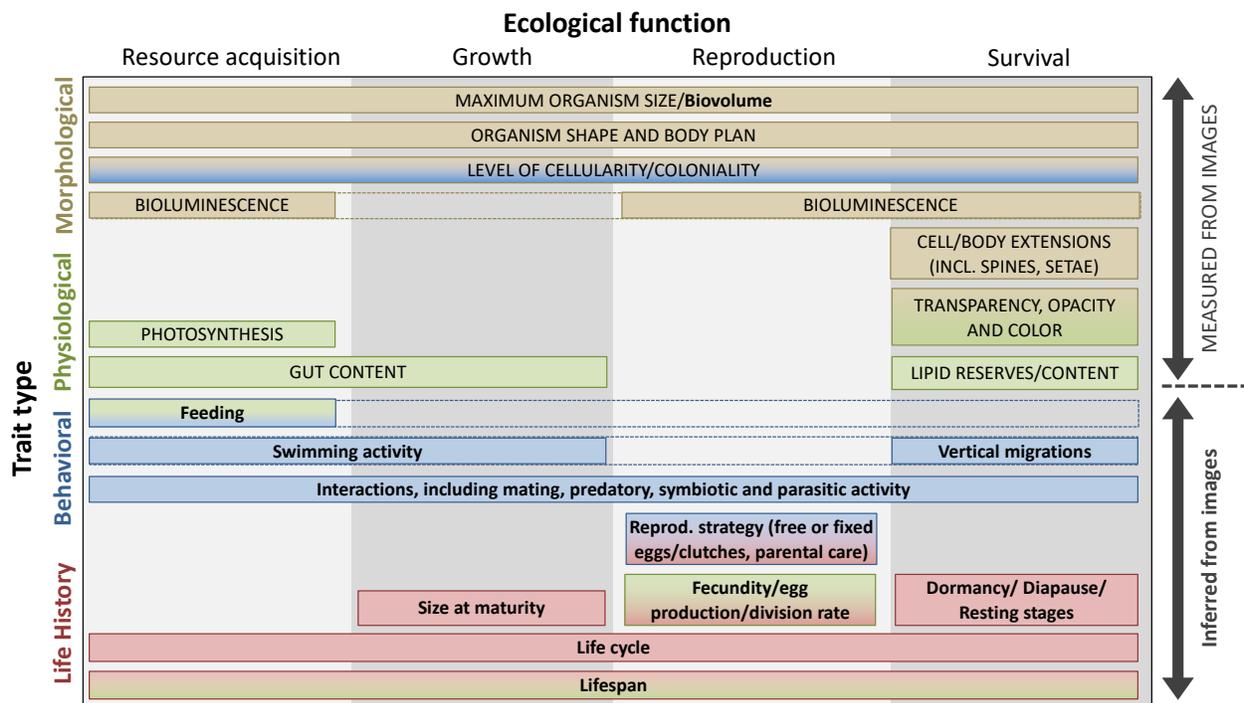


Figure 1: Plankton functional traits that can be estimated from images, following the unified typology of Martini et al. (2021). Trait types along the y-axis follow the order of Section 2 below. Measured traits, ones that can be quantified solely from images, are in capital letters. Inferred traits, which require additional information beyond raw pixels, are written in bold text.

Plankton traits from images

Imaging of planktonic organisms enables studies that track whole communities while collecting individual-level metrics such as functional traits (Fig. 1). In situ systems, in particular, can capture images of undisturbed creatures that reveal their pose, position, behavioural responses, and local interactions between organisms (Ohman 2019, Vilgrain et al. 2021).

For the rest of this paper, we primarily discuss images that are already segmented into Regions of Interest (ROIs). For most plankton imaging systems, it is relatively easy to remove targets from full frame images since they are sparsely distributed against uniform background intensities (Gorsky et al. 2010, Orenstein et al. 2020). Plankton imaging systems usually collect ROIs by distinguishing foreground pixels from the background without regard to the actual nature of the object. Therefore, the instrument does not discriminate between living creatures and particles such as marine snow (faecal pellets and other organic detritus). At any time and depth in pelagic systems, abundance and biomass of marine snow dominate and thus form the bulk of collected image data (Alldredge & Silver 1988, Stemmann & Boss 2012, Ohman et al. 2018, Trudnowska et al. 2021). ROIs of marine snow are an important signal measured by imaging systems, but are not the focus of this work. In this section, we will focus on individual organisms' traits, following the typology presented in Fig. 1, and discuss how these traits could be measured or inferred from images.

Measured traits

Size: the “master trait”

Cell or body size is a functional trait that can be easily measured from a calibrated image (Fig 2; e.g. a-c). Size transcends the scales of organization of biological systems, from the individual to the whole ecosystem through allometric relationships involving metabolism, development, feeding and mobility (Fig. 1; Platt & Denman 1977, Hansen et al. 1994, Litchman et al. 2013, Blanchard et al. 2017). In aquatic ecology, the size structure of the

whole community of organisms has been recognized as a property explaining trophic network organization and functioning since the early 1970s (Sheldon et al. 1972, Sieburth et al. 1978). It is hence often referred to as a “master trait”.

Size should be understood in all its dimensions, not only in terms of length (L^1) but also surface (L^2) and volume (L^3) or, equivalently, individual biomass (see section 2.2). The major axis length is a first order estimate of size readily accessible in images. Surface can be estimated from the cross-sectional area an object occupies in the image plane. Volume can likewise be derived from the particle's 2D projection under a set of assumptions (Moberg & Sosik 2012). These metrics, computed either through length measurements or equivalent spherical diameter (ESD), are the standard approach to estimating size spectra from sets of individual plankton image data. The measurements, however, must be calibrated between systems and deployments since they can vary due to foreshortening as a function of the object's orientation relative to the image plane, the imaging system itself, or the image processing pipeline (Giering et al. 2020a).

Beyond size, other morphological traits have not been studied as extensively since plankton were historically sampled with nets or bottles. While effective for taxonomic identification, these extractive techniques can damage or destroy fragile aquatic organisms, hampering efforts to collect information about the relationship between the individual morphology¹ and its environment (Remsen et al. 2004, Whitmore et al. 2019). In situ imaging devices can, however, measure these intrinsic and contextual elements without altering an individual's expression of its morphological traits (Ellen et al. 2019).

Organism shape and body plan

The shape of an individual organism is intimately linked to its biological function and the way it interacts with the physical environment (Kiørboe 2011b, Hirst 2012, Ryabov et al. 2021). The particular hydrodynamic environment an organism lives in imposes strict constraints on body plans. A single organism can experience several orders of

¹ Note that in this paper “morphology” should be understood in its biological sense, i.e. the visually identifiable properties of an object, rather than in its computer vision sense, i.e. the numerical characteristics derived from the binary mask of the object (its “imprint” in the image).

magnitude of Reynolds numbers while growing from an egg to adulthood, from a “sticky” laminar flow to the turbulent flow humans experience while swimming. As a result, it usually develops different locomotion strategies throughout its ontogeny. Copepods, for example, develop from spherical nauplius larvae that generate propulsion with appendages surrounding their body, to torpedo-shaped copepodite stages equipped with appendages allowing them to hover or jump by hundreds of body lengths per second to escape threats detected via flow disturbances (Fig. 2i,j; Kiørboe et al. 2010).

Unicellular organisms, such as diatoms or rhizarians, can likewise alter the shape of the biological unit they are forming, e.g. by adopting colonial strategies and forming colonies or chains (Fig. 2a-c,z,δ; Biard & Ohman 2020, Kenitz et al. 2020). This strategy can be related to the need to (i) counter predation by inflating size, (ii) modify the surface-to-volume ratio to increase resource acquisition, and (iii) modify the drag coefficient to control settling velocities, or any combination thereof (Nielsen 2006, Du Clos et al. 2021). Such strategies can impact global-scale biogeochemical cycles by altering the size-spectrum of the planktonic community and the balance between the productive surface layer and the interior of the ocean.

Cell/Body extensions

Many planktonic organisms bear spines, stout setae, and other defensive structures that inhibit consumption by predators. Such structures (elongated spines on diatoms – Fig. 2c, stout caudal setae of some harpacticoid copepods, etc.) are optically resolvable and likely to confer a reduction in predation rate.

In other organisms, body extensions are involved in feeding. In situ imaging has revealed that the volume of water filtered by these aquatic organisms may be markedly different than expected from net-collected individuals (Ohman 2019). For example, recent results from mesopelagic foraminifera have shown that the complete resolution of calcite spines and rhizopods can increase the surface area of the foraging apparatus by 100 to 1,000 times relative to what was expected from the hard test alone (Fig. 2α; Gaskell et al. 2019). In situ images of several types of tentacle-bearing predators, including cnidaria and ctenophores, have also

shown dramatically larger surface areas when the full extent of natural swimming and hunting postures is taken into account (Whitmore et al. 2019). Many pteropods and appendicularians use large and elaborate mucus extensions that filter vast quantities of water (Fig. 2β; Alldredge 1981, Burridge et al. 2017). These fragile gelatinous structures have broad implications for the biological carbon pump by increasing particle flux to the deep ocean (Lombard & Kiørboe 2010, Katija et al. 2017). The new information revealed by imagery has forced a re-evaluation of the classical approach to aquatic trophic network dynamics and caused an update of some allometric relationships between body size and filtering rates and even, in some instances, prey size (Conley et al. 2018).

Finally, many organisms (especially crustaceans) have a variety of mechanosensory and chemosensory setae that are crucial for prey and threat detection and which can be resolved by imaging systems (Fig. 2o,p). For example, prey-predator interactions have been observed for the marine copepod *Oithona plumifera* by analysing frames from experimental video recordings (Jiang & Paffenhöfer 2008). This advancement could have important consequences for trophic network analyses and our understanding of matter and energy flows, since the sensory abilities of zooplanktonic organisms are tightly linked to their feeding strategies (Kiørboe 2011a).

Bioluminescence

Bioluminescence is the emission of light by organisms through a chemical reaction. It is a ubiquitous trait, known to exist in about three quarters of planktonic organisms down to 4000 m (Martini & Haddock 2017). Light emitted by organisms can differ in terms of pattern and intensity: luminous spots on a body-part, repeated pulses of bright light by gelatinous organisms, diffuse sources of luminous excreted material, or chains of luminescent structures (Priede et al. 2006). Regardless of its manifestation, bioluminescence is strongly related to intraspecific communication – such as mate finding – or predator-prey interactions (Haddock et al. 2010). Until recently, most of the biological functions attributed to this trait have been inferred from the organism ecology, morphology and the characteristics of the bioluminescent emission (wavelength, intensity, patterns, chemistry).

Image-based plankton traits from ML

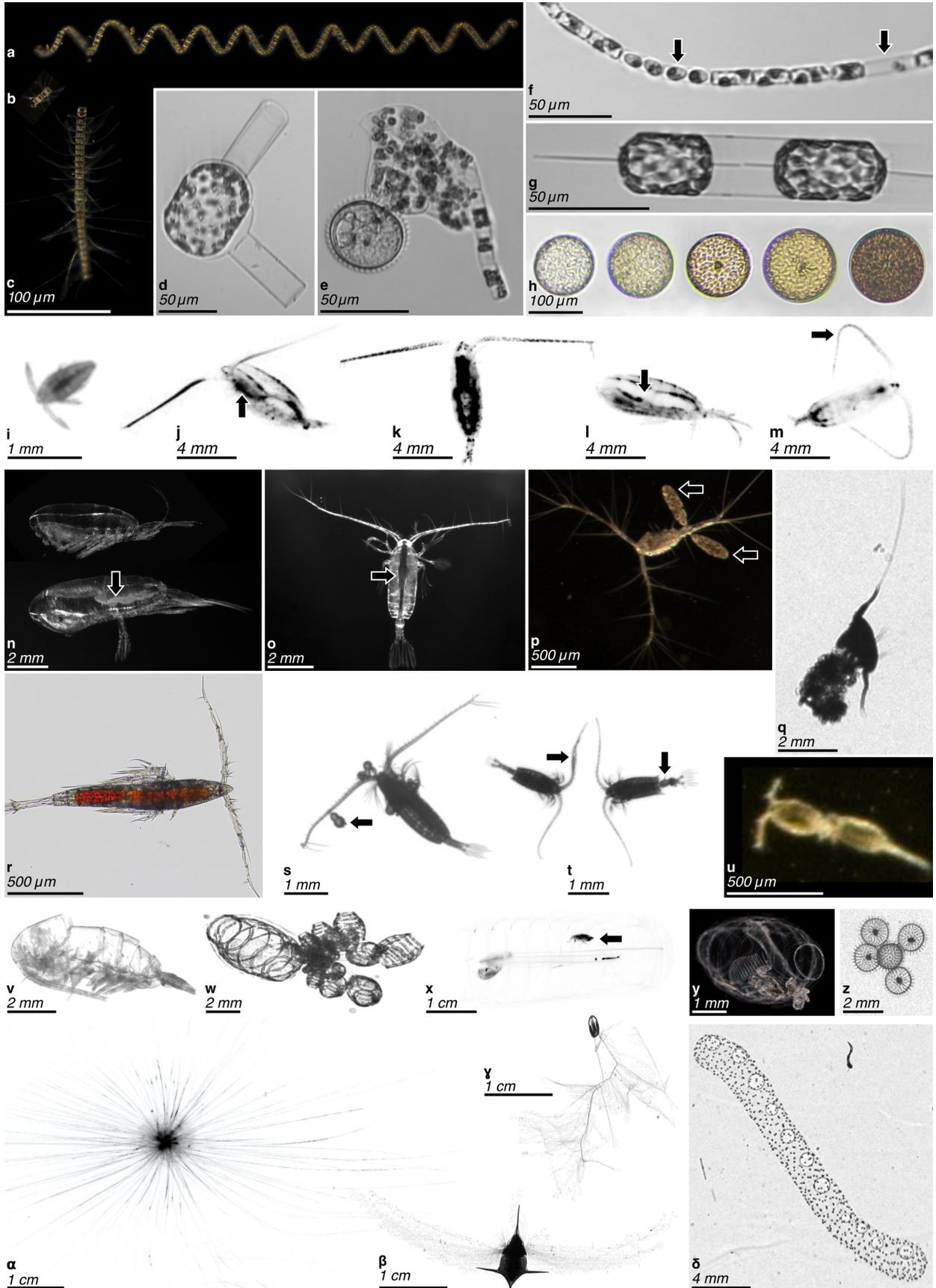


Figure 2: Example of plankton images on which traits can be identified. a-h: diatoms, i-v: copepods, w-δ: other taxa. a-c: chains of *Chaetoceros spp.* of different sizes [Scripps Pier Cam, SPC]; note the long spines on c. d: Sexual stage of *Guinardia flaccida* [Imaging FlowCytobot, IFCB]. e: dinoflagellate consuming a diatom chain (*Guinardia delicatula*) by external digestion in a feeding veil (pallium) [IFCB]. f: *Guinardia delicatula* infected with parasite (first arrow) or as an empty frustule (second arrow) [IFCB]. f: *Ditylum brightwellii* cell dividing [IFCB]. h: Coscinodiscophycidae (centric diatoms) containing various amounts of pigments [Planktoscope]. i: nauplius stage of a crustacean [ZooScan], j-m: calanoid copepods [Underwater Vision Profiler 5], note the full gut (arrow) and active posture with antennae deployed on j, the pigmented (dark) body parts on j to l, the lipid sac (arrow) and resting posture, with antennae along the body on l, and the curved antennae (arrow) associated with a jump of the copepod on m. n: immature (top) and mature (bottom, with visible oocytes – arrow) female of *Calanus hyperboreus* [Lightframe On-sight Key species Investigation, LOKI]. o: *Gaetanus brevispinus* displaying many sensory setae on its antennae and a well visible gut (arrow) [LOKI]. p: another copepod with well visible setae and two egg sacs (arrows) [SPC]. q: copepod associated with (possibly feeding on) a marine snow particle [ZooGlider]. r: *Microsetella sp.* displaying many spines and intense coloration, likely from its gut content [Planktoscope]. s: Calanoid copepod with parasite dinoflagellates (arrow) [ZooCAM]. t: male (with geniculate antennae – arrow, left) and female (with bulging genital segment – arrow, right) of *Centropages sp.* [ZooCAM]. u: *Oncaea* mating [SPC]. v: empty copepod carcass or molt [ZooScan]. w: Doliolid budding [ISIS]. x: Salp with an amphipod inside (arrow) [UVP]. y: transparent Doliolid [SPC]. z: a few solitary Rhizaria, family Aulospheridae [ZooGlider], to be contrasted with δ. α: Foraminifera with long cell extensions [UVP]. β: Pteropod (dark) with part of its mucus net deployed (grey). γ: Ctenophore, family Mertensiidae, with very long fishing tentacles deployed [ISIS]. δ: a colonial Rhizaria, order Collodaria [ZooGlider].

Bioluminescent signals are easy to spot, either by in situ observations or in the laboratory after collecting organisms, and can be spontaneous or mechanically stimulated. Hence, it remains difficult to assess whether measurements reflect natural behaviour or the organism's response to stimuli associated with the imaging method. Recent deployments of high sensitivity cameras on remotely

operated vehicles have enabled the acquisition of images of bioluminescence in situ. In the near future, we expect to see such tools used for consistent measurements of this trait, especially if we can accurately assess the level of disturbance created by the sampling devices (Ohman et al. 2018). In the meantime, signal analysis of bioluminescence intensity over time is a promising approach to better understand individual signals as well as community changes over time (Cronin et al. 2016, Messié et al. 2019).

Transparency, opacity and color

Tissue transparency is an important characteristic of the body plan of metazoans, including cnidarians, ctenophores, chaetognaths, and pelagic tunicates (contrast Fig. 2k,t with Fig. 2w,x,y). The combination of nearly transparent tissue with high water content allows organisms to be large, often with reduced metabolic requirements and reduced susceptibility to visual predators (Acuña et al. 2011). Opacity can likewise offer crucial information about many zooplanktonic organisms since body structures, such as gonads (Fig. 2k,n) and guts (Fig. 2j,o), can have variable content and appearances in the image (Vilgrain et al. 2021). Gut fluorescence, for example, has been used for several decades to estimate grazing rates on phytoplankton (Pasternak 1994), and can reveal whether feeding was recent or not (Sourisseau et al. 2008). Other pigments may be linked to physiological functions, such as carotenoids which offer oxidative stress protection to marine and freshwater copepods (Hylander et al. 2015, Schneider et al. 2016).

For unicellular organisms, pigmentation has been measured for decades from individual cells as a monochromatic or multispectral fluorescence signal (Olson et al. 1991, 2003) and can now be combined with imaging (Olson & Sosik 2007, Sieracki et al. 2010). Individual pictures can yield the identification and quantification of intracellular structures, such as chloroplasts (Fig. 2h) of photosynthetic symbionts, epibionts, or parasites (Fig. 2f). Pigments can also be measured either in their cell of origin (Fig. 2h), or inside the vacuoles or the gut of a consumer (Fig. 2r). Thus a combination of image measurements may be required to assess whether the pigmentation actually originates from the individual observed, especially in the complex

case of mixotrophic organisms (Flynn et al. 2013, Lauffer et al. 2015).

Photosynthesis

Among the main light-acquisition traits (Litchman & Klausmeier 2008, Kiørboe et al. 2018, Martini et al. 2021), the most common ones can be estimated at the individual level from images using microscope-based saturating-flash fluorescence measurements (e.g. quantum yield of photochemistry, photosynthesis-irradiance curves related parameters, non-photochemical quenching, functional absorption cross section). This has been performed in natural phytoplankton populations (Olson et al. 1996, Villareal 2004, Dijkman & Kromkamp 2006). Recent developments in hyperspectral imaging combined with optical microscopy will allow estimation of pigment content (Fig. 2h) and spectral absorption properties (Mélédér et al. 2013) at the individual scale (Xu et al. 2020).

Lipid reserves

Lipid allocation is a key component of energy reserves and buoyancy for a variety of organisms. In organisms that have adapted to conditions of intermittent resources, reserves can be found in dedicated internal structures that could be measured directly from images: vacuoles in unicellular plankton, lipid droplets in both unicellular and metazoan organisms, and lipid sacs in zooplankton thriving in extreme seasonal conditions (e.g. upwelling along eastern boundary currents, monsoon or polar oceans - Fig. 2l). Lipids can represent up to 80% of body volume in polar copepods, for example, and contribute actively to parts of the global carbon cycle through the *lipid pump* (Jónasdóttir et al. 2015). Imaging can be used to visually estimate these lipid reserves in copepods using geometric approaches akin to measuring whole body volume and biomass (Schmid et al. 2018). Lipid droplets in individual phytoplankton cells can be quantified using holotomographic imaging devices (Jung et al. 2018), though this technique is not feasible yet for in situ measurement.

Inferred traits

The functional traits discussed so far can be measured directly from images. Many others cannot but could rather be estimated by combining visible features and environmental context metrics ob-

tained from the sampling device metadata (temperature, light level, etc.).

Biomass and biovolume

Biovolume and biomass can be estimated from an organism or cell measured area using empirical relationships specifically derived for a particular imaging instrument and target organism (Hillebrand et al. 1999, Menden-Deuer & Lessard 2000, Moberg & Sosik 2012). Biomass estimates from biovolume can involve simple conversion factors and ratios that may vary between organisms. For example, a jellyfish and a crustacean of the same ESD have distinct biomass conversion factors (Lehette & Hernández-León 2009, McConville et al. 2017, Giering et al. 2019). These relationships can be generic for whole groups or calibrated for individual species or development stages (Ikeda et al. 2007).

Feeding and metabolic rates

In situ images of relatively transparent aquatic organisms can be used to estimate the volume swept while feeding (Fig. 2a-y), the digestive tract fullness (Fig. 2j,o), and the frequency of actual feeding interactions (Fig. 2e,q). Ingestion rates, gut transit time and maximum feeding rates could subsequently be estimated with concurrent measurements of environmental variables such as temperature and prey concentrations (Wirtz 2013). Metabolic activities, such as respiration and excretion, scale with organism size (Banse 1976, Ikeda et al. 2001), allowing the estimation of, for example, zooplankton community respiration or excretion as the sum of all individual organisms' activities (Ikeda 2014).

Swimming and activity

Aquatic organisms have evolved diverse locomotive structures with clear morphological signatures like gas vacuoles, flagella, and cilia. An organism's specific motility strategy affects the trade-off between energy intake through feeding and death by predation (Harvey & Menden-Deuer 2012). It also indicates if the organism can actively influence its position in the water column and overcome turbulent mixing (Gallager et al. 2004) (Fig. 2m, of a copepod "jumping"). Activity patterns can potentially be inferred from still images based on the position of body and appendages. For example, Vilgrain et al. (2021) suggested that the "body contour complexity" defined by their multivariate

analysis of thousands of individual images of Arctic copepods revealed that copepods were hovering and foraging in the marginal ice zone in spring when the phytoplankton bloom occurred (Fig. 2j,k). Under continuous ice cover, individual copepods displayed a posture typical of dormancy (Fig. 2l).

Diel Vertical Migration (DVM) is a peculiar case of swimming activity that influences the fitness of a range of highly mobile aquatic organisms. Zooplankton evolved this behaviour as a trade-off between maximizing feeding opportunities on the primary producers thriving in the well-lit surface and minimizing mortality from visual predators (Aksnes & Giske 1990, Möller et al. 2015). Recent results suggest that both the propensity to migrate vertically and the amplitude of DVM varies with the body size of planktonic copepods (Ohman & Romagnan 2016). The accurate characterization of metazoan body size could then be combined with visual indices of activity level, to predict the migration behaviour of zooplanktonic organisms.

Interactions

Interactions among organisms or between organisms and particles can be split in two categories: the first involves interactions between distinct entities spread within the water column (e.g. predator and prey – Fig. 2e), while the second occurs at the scale of an individual (e.g. parasitism – Fig. 2f,s,x). In situ imaging systems that record the undisturbed relative position of individuals could be used to reveal traits related to spatial interactions, such as the feeding (Fig. 2e,q), mating (Fig. 2u), parasite infection (Fig. 2f), or predation. When coupled with the appropriate information extraction technique (see section 3), such data will allow scientists to better understand inter- and intra-specific relationships, from intracellular symbiosis and parasitism to planktonic interactions within microhabitats like marine snow or other detrital matter (Peacock et al. 2014, Nishibe et al. 2015).

Reproduction strategy, fecundity, and division rate

Reproduction strategies form a continuum bounded by r (many offspring, no brooding) and K (few offspring, brooding) strategies (Jaspers et al. 2018). Taxa with r -strategies tend to have a smaller size at maturity and to produce large clutches – many eggs are released during a single spawning event – while K -strategists tend to produce larger adults

and release a few eggs at a time. K -strategists' eggs are also often much larger, provided with substantial lipid reserves, and benefit from parental care (e.g. brooding pouches from amphipods or egg sacs from copepods).

In situ imaging systems, with sufficient pixel resolution, could be used to count free eggs in the water column, individuals within egg sacs (Fig. 2p; Möller et al. 2015) or to estimate the occurrence of mature oocytes in females' gonads (Fig. 2k,n; Niehoff 2007). Daily individual egg production rates could be estimated from these measures – coupled with simultaneous in situ temperature data, the estimated proportion of mature females, and female size at maturity – via known allometric relationships and temperature-dependent, species-specific hatching time (Niehoff 2003). For asexually reproducing organisms, new budding individuals could be directly counted on the image (Fig. 2w), and frequency of cell division (Fig. 2g) could be used to infer protist population level division (Campbell et al. 2010, Brosnahan et al. 2015). Monitoring the evolution of taxonomically-resolved planktonic cell size in both time and space could provide information on productivity levels of a system (Hofmann et al. 2019).

Life cycles and lifespan

The majority of plankton are unicellular organisms whose cell morphology is tightly linked to the phase of an individual's development cycle, including sexual stages (Fig. 2d) and dormant stages (Brosnahan et al. 2015). Most multicellular zooplankton species also have morphologically distinct life stages (e.g., Fig. 2i,j). Usually, the individual develops from an egg to adulthood through several moulting or metamorphosis events. This process requires a significant energetic investment over time, but it ensures that the morphology of growing individuals is well adapted to their changing environment (Sainmont 2014). The appearance of early larval life stages of crustaceans, pteropods, and polychaetes in the water column is indicative of recent reproduction and successful hatching. Images of sufficiently high resolution to identify these larval stages can thus yield the reproductive phase of a local population and contribute to estimates of the ecosystem productivity. Images could also provide further information on individual ages. Zooplankton age is crucial for

studying population and community dynamics but is notoriously difficult to determine (Kjørboe et al. 2015). We are not aware of any attempt to accurately estimate a numerical value of the age of individual zooplankton directly from in situ images. It is unlikely that this could ever be achieved, except for very peculiar cases where aging would somehow “scar” an individual in a typical and predictable way.

Dormancy and resting stages

Resting stages, such as diatom dormant spores, can be inferred from images and used to better estimate carbon fluxes as they play a significant role in carbon export (Salter et al. 2012, Rembauville et al. 2018). Similarly, zooplankton that accumulate lipid reserves can undergo dormancy (both quiescence and true diapause – Fig. 2l). This trait is associated with a shift in metabolic rate, activity level, habitat, swimming behavior, and survival (Baumgartner & Tarrant 2017). Thus, it has a significant impact on individual fitness, local population dynamics, and the biological carbon pump (Jónasdóttir et al. 2019). Imaging has been used to characterize dormant and active copepods in the Arctic by collecting individual-level measurements of stage development, lipid content, and associated depth distribution (Schmid et al. 2018). Such individual measurements can be combined with in situ temperature data to estimate an individual’s respiration rates at the depth of dormancy and, eventually, infer the duration of its dormancy phase (Maps et al. 2014).

Mortality

Living versus dead organisms can be distinguished both manually and automatically in image data (Fig. 2f,v; Picheral et al. 2017, Reimann et al. 2020). The use of high-throughput imaging devices deployed in situ offer new opportunities to estimate the spatio-temporal variation in the proportion of dead individuals within natural phyto- and zooplankton communities. More generally, the accumulation over time and space of finely resolved trait distributions (e.g. development stages, sizes, proportion of dead individuals) could significantly increase the usefulness and precision of statistical approaches used to estimate mortality rates and life span based on cohort detection and monitoring (Shaw et al. 2021).

Computer vision and machine learning approaches to estimate traits from images

Computer vision and machine learning techniques offer scalable solutions to automatically quantify functional traits from plankton images. Successful deployment of these methods will require careful consideration of the target trait, development of specialized annotated data, selection of the most appropriate algorithm, and application of a consistent evaluation metric.

To this end, we recommend adopting an evaluation-first design paradigm – where the specific task and quantitative evaluation metric are defined before annotation begins – to motivate the construction of these resources and the necessary workflow (Gupta et al. 2019). Following this principle will encourage scientists to think carefully about the algorithm they wish to use, the criteria for success, and therefore what sort of labelled data they need. This output-centric approach will smooth the transition from labelling to inference.

In order to measure functional traits from images, one must design a pipeline from human annotation to deployment on unlabelled data (Fig. 3). Here, we discuss how to construct a training dataset; potential computer vision and machine learning approaches to estimate traits from images; and how to evaluate the trained systems.

Datasets and labelling

The critical step in any supervised machine learning experiment is the construction of an appropriately labelled training set. For classification tasks, this entails deciding on the appropriate set of classes to consider and annotating the data until there are sufficient labelled samples – typically 100s to 1,000s per class for deep neural networks – to train the desired model. Training systems to detect and measure traits requires human effort beyond sorting images to generate appropriate labelled data. The exact approach depends on the target trait and corresponding computational method.

Annotation formats

Machine learning frameworks expect data in specific standard formats. To train a classifier, for example, training images are often expected to be sorted into unique folders or accompanied by a

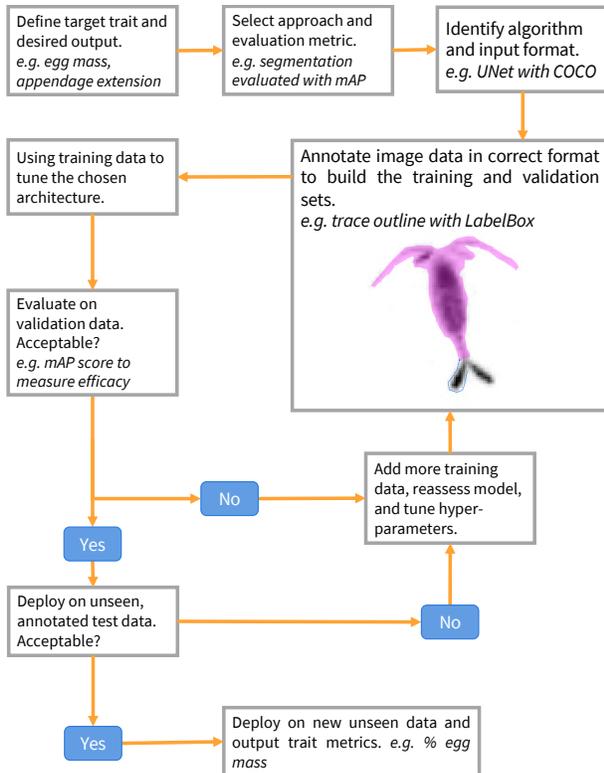


Figure 3: Workflow diagram for a computer vision approach targeting a specific functional trait extracted from plankton image data. In this example, a group decides to target egg-bearing copepods with a UNet segmentation model. A human annotator selects ovigerous tissue from a copepod image and outputs the mask in the COCO data format. The model is then trained and evaluated using the mean Average Precision (Box 2; SI 1). Note that this workflow is not specific to plankton and could also be used for other types of organisms.

numeric list indicating the class names. Object detection, segmentation, and keypoint detection frameworks all likewise require annotations in specific formats for training.

There are two common natural image datasets used for benchmarking new approaches to such tasks: PASCAL Visual Object Classes (VOC)² and Common Objects in COntext (COCO)³ (Everingham et al. 2010, Lin et al. 2014). These two databases contain many thousands of images with bounding boxes and pixel level segmentation labels. Publicly available domain specific datasets often ad-

here to the VOC or COCO formats to facilitate experiments and interoperability.

Scientists seeking to fine tune existing architectures trained on these datasets might consider adopting their exact annotation formats. Alternatively, operators could save annotations in whatever format is most convenient and write their own data loader in the selected toolkit. PyTorch, for example, has a generic Dataset object that can be modified to accommodate any input and output.

Annotation tools

There is a broad universe of tools for data annotation, ranging from freeware interfaces to industry-scale software packages (e.g. LabelBox, RectLabel, and MATLAB Image Labeler). Many of these are open source and can be adapted to include contextual metadata or other parameters. Several tools have been written specifically for *in situ* marine imagery that facilitate integration of oceanographic metadata (Gomes-Pereira et al. 2016, Picheral et al. 2017). The choice of annotation software should be dictated by ease of use and the ability to output in standard data formats. Practitioners should invest time researching contemporary options and select a package that is actively being supported. Annotation requirements and format standards will undoubtedly evolve as the technology changes. At the time of writing, researchers should look for services that (1) output data in COCO or VOC format; (2) have an interface for polygon object detection labelling; (3) support superpixel segmentation labelling assistance; and (4) allow hierarchical annotations. We note that annotation interfaces are typically independent of databases for storing and working with the resulting data.

Data augmentation

Annotation efforts of any kind are expensive in terms of both human labor and monetary cost, limiting the availability of training data. Data augmentation is a common procedure to subtly alter the appearance of labelled ROIs – via geometric transformations or filtering operations – to avoid overfitting the model. Done appropriately, the approach can improve the representational power of the automated system. There is also potential for producing synthetic training data via 3D models or

² <http://host.robots.ox.ac.uk/pascal/VOC/voc2007/html/doc/index.html>

³ <https://cocodataset.org/#format-data>

Box 1: Definitions of a few computational terms, highlighting important, but subtle, differences.

Computer vision (CV) A broad subfield of computer science dedicated to using a computer to interpret images and video sequences.

Machine learning (ML) A set of statistical approaches that attempt to discern patterns in data, either automatically or based on explicit human instructions.

Supervised ML ML techniques that teach a computer to recognize patterns using a set of expert-curated examples, such as annotated images.

Unsupervised ML ML methods that attempt to group data together without human intervention. Clustering algorithms are a common example. Their performance is often difficult to evaluate.

Training set A collection of data annotated by human experts for teaching a computer how to interpret information. Building the labelled dataset is the most time-consuming and critical part of an ML workflow.

Validation set A separate human labelled dataset used to evaluate a trained system. These data are entirely independent of the training set and should represent conditions the system might encounter in the field. Also referred to as test data.

Feature-based learning ML algorithms that operate on a reduced, hand-engineered feature space. Each data point is cast as a vector of measurements and used to tune a set of parameters that dictate how the model works.

Deep neural networks (DNNs) A type of representational algorithm that learns directly from raw data. DNNs layer many mathematical abstractions on top of each other to connect input information to a desired output. Through iterative training, the system learns the most salient features of the input. Modern DNNs often have numerous layers and billions of weights.

Transfer learning A shortcut for training DNNs by repurposing a network originally trained for a different task.

image merging before training (Mahmood et al. 2020). All these operations should be undertaken conservatively and in consideration of the target trait to avoid overfitting one's model to the augmented data.

[Datasets for transfer learning and feature extraction](#)

Transfer learning is an effective way to reduce the initial annotation effort when preparing data to train a deep neural network architecture for a new

task. The parameters of a model originally trained on a generic dataset like ImageNet are fine-tuned to the new target dataset (Yosinski et al. 2014). The method has been used for classification tasks for plankton data (Orenstein & Beijbom 2017, Cheng et al. 2019). There are currently three large, publicly available plankton image datasets sorted for classification tasks that can be leveraged for such designs (Sosik et al. 2014, Cowen et al. 2015, Elineau et al. 2018).

Transfer learning is also an effective approach for repurposing trained networks for other computer vision tasks such as object detection or semantic segmentation (Girshick et al. 2014, Long et al. 2015, Ouyang et al. 2016). Doing so for plankton images will require generating sufficiently large labelled datasets akin to those produced for classification tasks. Codifying the annotation format across imaging systems – and making the resulting datasets public – will speed development of effective systems. Moreover, if these datasets are made public, consistent data formats will ease interoperability and speed development.

Feature extraction from deep nets – using internal representations of a model – is another viable option for rapidly generating features and classifiers from previously trained generic classifiers (Donahue et al. 2014). These weights can be used for unsupervised clustering or to train ensemble or margin classifiers (González et al. 2019, Schröder et al. 2020).

[Algorithms, approaches, and evaluation](#)

There are a myriad of algorithms and approaches that have been developed for computer vision tasks that could be adapted for trait-based projects. As a motivating example, consider two hypothetical studies: one examining ovigerous copepods (Fig. 4) and another looking at parasitized diatoms (Fig. 5; Peacock et al. 2014). Here, we detail several candidate techniques that might be used to develop a pipeline to generate the required data: hand-crafted features, semantic classification, object detection, segmentation, deep regression, and key-point estimation. We note there are many specific algorithms within each category that are worth exploring. Here, we seek to give a broad overview of each area without making specific recommendations as to the exact approach to implement.

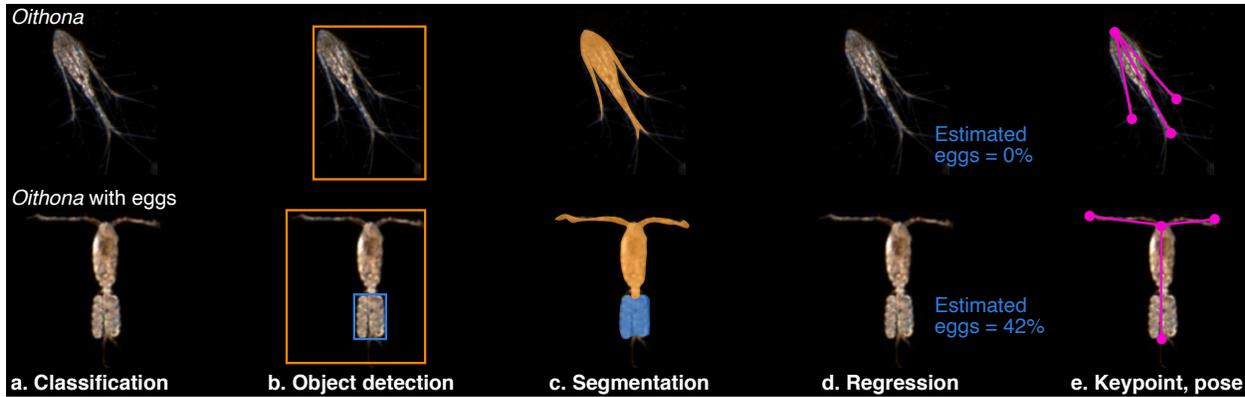


Figure 4: Examples of several techniques for trait extraction from zooplankton images. The hypothetical use case is examining ovigerous copepods imaged by the Scripps Plankton Camera system. The top panel is a non-egg bearing copepod. The bottom panel is an individual carrying an egg-sac. (a) Automated classifiers could be trained to add a semantic descriptor to the taxonomic class. (b) Object detection finds the organism and desired trait. (c) Segmentation algorithms classify the pixels as belonging to the organism or the trait. (d) Regression estimates the percentage of pixels that represent the trait. (e) Keypoint/pose estimation finds body nodes (red dots) and connects them (yellow lines) to estimate orientation or appendage extension.

Handcrafted features and measurable traits

Morphological traits measurements can be made by extracting hand-engineered features from ROIs. Such measurements are made directly from the pixels and might include major axis length, equivalent spherical diameter (ESD), or opacity. Techniques for doing so are well-established and standard in many in situ imaging pipelines (Blaschko et al. 2005, Sosik & Olson 2007, Gorsky et al. 2010). If the instrument produces color images, the color channels could be used to infer information regarding an organism's pigment content or transparency. Calibrating these approaches can be done

concisely by comparing automated measures with a ground truth set made by a human expert.

Morphological traits measurements (as described in section 2.1) are the most directly interpretable, but require careful calibration of the pixel size and sample volume of a particular instrument. Likewise, interoperability can be challenging since the feature extraction technique will need to be rewritten for a new instrument – there is no general use feature extractor for different systems. However, once measurements are made and converted into realistic units, the metrics are easily comparable across devices.

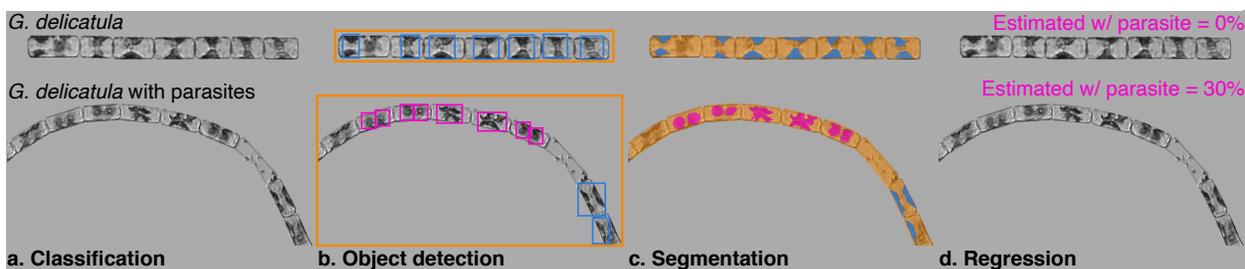


Figure 5: Examples of several techniques for trait extraction from phytoplankton. The hypothetical use case is examining parasitized diatom chains imaged by the Imaging FlowCytobot. The top panel is a healthy *Guinardia delicatula*. The bottom panel is a parasitized chain. (a) Automated classifiers could be trained to add a semantic descriptor to the taxonomic class. (b) Object detection finds the entire chain, chloroplasts, and parasites. (c) Segmentation algorithms classify individual pixels as belonging to the organism, chloroplasts, or parasites. (d) Regression estimates the amount of an image that corresponds to the organite/parasite biovolume.

Classification with trait-specific classes

Automated classifiers can be retrained to target specific traits based on an appropriately labelled set of images. If one wishes to separate ovigerous copepods from other copepods, they could add a training category specifying the semantic class “with eggs” (Fig. 4a). Similar semantic categories could be constructed to address traits such as body extension, presence of parasites, or behavioral signatures (e.g. extended appendages, body position, etc). Any classification method could be used in this manner: ensemble, margin, deep nets or otherwise. This has already been done to some extent in the publicly available IFCB and ISIIS datasets (Sosik et al. 2014, Cowen et al. 2015). Both datasets currently contain a number of classes distinguished with semantic descriptors related to traits: “G_delicatula_parasite” or “G_delicatula_external_parasite” from the IFCB (Fig. 5a); and “trichodesmium_puff” or “trichodesmium_tuft” from the ISIIS. Several classification efforts on these and other datasets already include such trait information (Luo et al. 2018, Ellen et al. 2019, González et al. 2019).

These efforts can be expanded in a targeted way by labeling (or re-labeling) images with trait information and tuning a new algorithm. The resulting classifier could then be run on new unlabelled data or as a second stage analyser on the output of a more general computer annotator. Evaluating the output could be done with raw accuracy on a per class basis or as the mean average precision (Box 2). As with any automated classification pipeline, care will need to be taken to assess issues related to dataset shift (Moreno-Torres et al. 2012, González et al. 2017).

Object detection for trait identification

Object detection pipelines are built to identify, select, and propose labels for subregions in an image (Zhao et al. 2019). Such techniques could be used to select egg sacs from the rest of a copepod’s body or parasites inside a diatom chain (Figs. 4b; 5b). Other target traits might be bioluminescent structures, lipid reserves, or specific body parts. The models could also be used to locate organisms in a full frame image (Katija et al. 2021). Training object detection models for such tasks will require new training sets that include bounding box coordinates and labels targeting the traits. The resulting

bounding boxes could be used for count based studies like estimating the number of ovigerous individuals or the percentage of a population that has been feeding.

Object detection was historically accomplished by proposing regions based on hand-engineered features from an input image and a localizer searching the feature space (Lienhart & Maydt 2002, Viola & Jones 2004, Felzenszwalb et al. 2010). Recently, deep neural networks – such as Faster Regions with CNN Features (RCNN), Single Shot Multibox Detector (SSD), and You Only Look Once (YOLO) – have been specifically designed to identify and select subregions of images (Girshick et al. 2014, Liu et al. 2016a, Redmon et al. 2016). These techniques are generally used to draw bounding boxes around objects in natural image datasets like COCO and VOC. Object detectors are typically evaluated with the mean Average Precision score (Box 2).

Semantic and instance segmentation

Instead of selecting bounding boxes around subregions, segmentation algorithms attempt to delineate sections of the image at the pixel level. Rather than returning boxes surrounding interesting areas of an image, segmenters produce masks where each pixel is tagged as belonging to a particular region. Segmentation approaches are appropriate for trait studies that require estimating mass. A computer could be trained, for example, to select pixels that represent pigmented regions of an organism, gut content, the egg tissue, or parasites (Figs. 4c; 5c). The resulting masks could then be used to estimate an organism’s state (healthy/starved, ovigerous/not) or the proportion of tissue dedicated to such functions.

Preparing data for training segmentation algorithms will require very detailed human delineated masks. The annotation task thus requires more time and attention from the human expert than object detection or classification labelling. Super-pixel-based assistance workflows can be used to accelerate the annotation task (King et al. 2018). Segmentation tasks are also typically more computationally difficult than detection. There are again many flavors of segmentation algorithms. At the time of writing, UNet, Segnet, and Mask RCNN are among the most widely applied models for generic object segmentation tasks (Ronneberger et

al. 2015, Badrinarayanan et al. 2017, He et al. 2017). The performance of segmentation algorithms is typically stated as the Average Precision at different levels of Intersection over Union (Box 2).

The computer vision community makes a distinction between “semantic” segmentation – labelling all pixels with a class – and “instance” segmentation – separating each occurrence of a given class. For example, consider an image of two touching copepods (Fig. 2u): semantic segmentation would label all “copepod” pixels as a single region while instance segmentation would label them as “copepod 1” and “copepod 2”.

Deep regression

The ultimate goal of any of the techniques we have mentioned is to output an estimate of some relevant trait – angle relative to horizontal, extension of limbs, presence of eggs, etc. Rather than learning an intermediate task like segmentation or object detection, deep regression algorithms seek to train a model to directly output these estimates. Regression approaches are thus appropriate for calculating numerical properties of an object. Returning to the copepod egg example, rather than training an object detector or segmentation algorithm, an operator could instead build a network that outputs an estimate of the percentage of the ROI occupied by eggs (Fig. 4d). Likewise, a deep regressor could be trained to estimate the percentage of a chain ROI that is parasitized (Fig. 5d).

Deep regression approaches have been successfully implemented on photos of humans to directly estimate head pose and a subject's age (Liu et al. 2016b, Rothe et al. 2018). Using deep regression in this manner will discard other information in the image, but might be a more efficient means of getting the value of a trait that can be expressed as a continuous number (Lathuilière et al. 2020). The efficacy of these techniques can be expressed as the mean squared error of the predicted estimate.

Pose or keypoint estimation

Pose estimation algorithms attempt to teach computers to recognize how a human body is positioned in an image (Sapp and Taskar 2013; Andriluka et al. 2014). Many algorithms search for and connect keypoints that together describe a very rough skeleton. Keypoint analysis is ideal for studying organism orientation and limb extension. By identifying head, tail, and extremities, the com-

puter can estimate what direction the organism is facing relative to the camera and the angle of its body parts (Fig. 4e). Such an approach could also be used to investigate interactions between organisms from full frame images.

Currently, pose estimators are almost exclusively used to find people and determine their 3D orientation. Several recent papers have suggested adapting these methods to animal images both in the wild and the lab (Cao et al. 2019; Li et al. 2020). Keypoint and pose tasks are relatively new as compared to object detection and segmentation. A few contemporary approaches are PersonLab, OpenPose, and DensePose (Cao et al. 2017; Papan-dreou et al. 2017; Alp Güler et al. 2018). Keypoints are typically evaluated by specifying a threshold distance from the true point location. Point proposals can then be evaluated with accuracy as a function of the threshold – akin to a receiver operator curve.

Evaluation metrics

There are many evaluation metrics used to assess the efficacy of machine learning systems, and several that are particularly important to the methods we have discussed (Box 2; SI 1). Selection of the appropriate metric can better guide training and deployment of ML systems by providing informative feedback on the system performance. Classifier accuracy (*acc*), the average percent of correctly returned labels over all classes, is the most commonly reported metric for planktonic imaging systems. *acc* is a useful but flawed metric, often obscuring important errors (Tharwat 2020). Practitioners should rather rely on more informative metrics like precision, recall, and the F1-score (Box 2; SI 1).

Object detectors and segmentation algorithms are often evaluated with the *mAP* score, as defined in the VOC and COCO challenges. The number encapsulates the trade-off between precision at different levels of recall or Intersection over Union (SI 1). Versions of *mAP* are found in many areas of computer vision and will prove useful in ecological applications. Other approaches for trait measurements will likely have different evaluation metrics. Time and energy should be put into selecting and understanding these numbers, as they will guide model development and deployment.

Box 2: Common evaluation metrics for automated classifiers.

Accuracy (<i>acc</i>)	$\frac{t_p + t_n}{t_p + f_p + f_n + t_n}$	The ratio of the number of correct labels returned by an automated classifier over the total number of samples.
Precision (<i>p</i>)	$\frac{t_p}{t_p + f_p}$	The ratio of the number of correct labels over the total number of samples assigned to that class.
Recall (<i>r</i>)	$\frac{t_p}{t_p + f_n}$	The ratio of the number of correct labels over the actual number of samples in the class.
F1-Score (<i>F₁</i>)	$\frac{t_p}{t_p + \frac{1}{2}(f_p + f_n)}$	The harmonic mean of precision and recall that summarizes model performance in a single metric scaling between 0 and 1.
Average Precision (<i>AP</i>)	$\frac{1}{n+1} \sum_{r \in \{0, 1/n, 2/n, \dots, 1\}} p(r)$	The average of the precision for different levels of the recall.
Mean Average Precision (<i>mAP</i>)	$\frac{\sum_{c=1}^Q AP_c}{Q}$	The mean <i>AP</i> over all <i>Q</i> classes. It is a summary statistic that describes how a model does over all classes.
Intersection over Union (<i>IoU</i>)	$\frac{A \cap B}{A \cup B}$	<i>A</i> ∩ <i>B</i> represents the area in pixels of the overlap between the region proposed by the computer (<i>A</i>) and the ground truth (<i>B</i>). <i>A</i> ∪ <i>B</i> is the total number of pixels contained in both regions. IoU is a metric mostly used for detection and segmentation tasks.

Perspectives and conclusion

Using ML for trait estimation: a promising tool

Digital imaging systems, coupled with automated analysis techniques, are increasingly used by aquatic ecologists. Much of the community's effort has been focused on developing classification approaches, largely delineated along taxonomic lines (Irison et al. In press). We believe that the extraction of functional traits is the next evolution in machine learning-enabled analysis of ecological images. With traits values derived from in situ images, aquatic ecologists could automatically estimate functional diversity, better link diversity to ecosystem functioning, and parameterize and test trait-based models to assess the health of aquatic ecosystems (Martini et al. 2021).

As with classification efforts, developing trait extraction systems will require a substantial investment of expert human analysis. Images will need to be closely scrutinized and appropriately

labelled, sometimes at the pixel level, to perform experiments and validate new approaches. To this end, we strongly advocate researchers employ the evaluation first design paradigm, where the target trait and criteria for success are chosen before any annotation begins. Doing so will ensure that labels are stored in the appropriate format and enough data are annotated for the experiments, thus facilitating interoperability.

All of the computational approaches presented in Section 3 have the potential to yield significant new scientific results. We believe that deep regression and pose estimation are especially promising for plankton trait studies. Pose estimation could be used to examine the orientation of a group of organisms relative to horizontal or the angle of their appendages relative to their body. Deep regression would be effective for estimating proportion in an image: e.g. how extended are the appendages or mucus net, how full is the gut, how big is the lipid reserve, how curved or deformed is the shell, etc. This approach will require segmentation style an-

notations but is ultimately relatively simple to train and deploy. We stress that the specific approaches and algorithms discussed in this paper are only representative; there are a multitude of viable combinations of traits, machine learning techniques, and statistical analyses. Moreover, new computer vision approaches are being released every year; many of the specific models mentioned in this paper will very likely be obsolete within a few years.

We anticipate that functional trait-based approaches via imaging will be maximally effective when undertaken with a suite of instruments. Facilitating studies at such a broad scale requires conscientious calibration efforts for all instruments capturing images of plankton. Indeed, calibration is arguably more important for trait extraction than for taxonomic classification. At a minimum, the pixel size has to be well-established and reported to ensure that recovered measurements are appropriately scaled. While the estimation of pixel size is possible for microscopic imaging systems that are imaging precisely in a single plane, pixel size calibration cannot be achieved with single cameras; stereo vision or active illumination is required to recover the scale. Ideally all sources of variability, such as intensity and blur characteristics, will be quantified and made widely available (Vandromme et al. 2012; Giering et al. 2020). While sharing these parameters will aid cross-dataset comparisons, true interoperability will be achieved only through extensive intercalibration studies.

From individual traits to functional diversity and ecosystem functioning

When appropriately applied and validated, the computer vision and machine learning techniques we discussed will yield previously inaccessible data: consistent, large-scale quantitative measures of trait expression at the individual level. With these measurements, single trait distributions can be estimated in space and time (Barton et al. 2013; Brun et al. 2016; Carmona et al. 2016) and intraspecific variability of traits could be documented (Martini et al. 2021). Statistical models can subsequently be implemented to better understand the interplay between trait expression and environmental factors by deriving empirical relationships between them. These empirical relationships could

further be used to structure and parameterize mechanistic models of planktonic communities based on equations describing functional types and traits (Follows et al. 2007; Stemmann and Boss 2012; Ward et al. 2012; Le Quéré et al. 2016; Serra-Pompei et al. 2020). Marrying large-scale observational capacity and modelling in this manner would allow scientists to examine evolutionary constraints on trait expression at unprecedented spatiotemporal scales.

The sheer volume of trait data these approaches could produce has important implications for how aquatic ecologists measure diversity (Fontana et al. 2018). Several descriptors of functional diversity have been defined – such as functional richness, divergence, or evenness (Villéger et al. 2008) – that could be studied with image-based functional trait data. Mathematically, these diversity metrics are properties of a hypervolume in a high-dimensional space defined by the functional traits of interest. Many metrics have been proposed and several user guides have been written about their definition, properties, and utility for studying functional diversity (Schleuter et al. 2010; Mouillot et al. 2013; Carmona et al. 2016; Schmera et al. 2017; Legras et al. 2018). It is important to bear in mind that these tools are sensitive to the species richness within an observed community, as well as the number and type of traits considered (Lefcheck et al. 2015; Legras et al. 2020). These caveats suggest that substantial effort should be devoted to understanding how image-based measurements and intra-specific variability affect the trait hyperspace being studied. It is possible to circumvent such metrics by clustering individuals or taxa in a trait space defined by measurements of individual morphologies (Trudnowska et al. 2021; Kruk et al. 2021).

The estimation of functional traits at the individual level will also yield further knowledge on ecosystem functioning by giving access to bulk properties at the population or community levels, including community-scale insight into the ecology of patchiness, individual interactions, and associated advantages for resource acquisition such as light and nutrients (Litchman and Klausmeier 2008). For instance, sex ratio could also be estimated from images, since male and female zooplankton are often distinguished by traits that affect physiology, behaviour and eventual fitness

(Heuschele et al. 2013). Females typically invest more resources in gonad development and offspring production, often incurring higher predation risks as a trade-off. For example, during DVM, females of the krill species *Meganyctiphanes norvegica* migrate significantly closer to the surface to forage at night than males (Tarling 2003). This behaviour leads to 40% higher energy intake, yet increases the risk of visual predation, resulting in the male-to-female sex ratio shifting from 1:1 at the beginning of the summer to 3:1 by the onset of winter. The ability to record finely resolved vertical profiles of sex-associated trait frequencies could allow direct quantification of mating seasonality and potential mating success and output. Such observations could also allow detecting potential ecological perturbations as deviations from a baseline distribution of traits. At the community scale, image-based functional traits such as size, photosynthetic traits, metabolic rates or vertical migration, provide further access to ecosystem functioning through the estimation of primary production (Litchman et al. 2015) or carbon export (Stamieszkin et al. 2015; Archibald et al. 2019; Kiko et al. 2020).

Desirable qualities of imaging instruments

Future technology development should consider trait analysis in the hardware and software design phase. Four particular capabilities would be informative for trait-based studies: (i) the ability to record color images, (ii) the capacity to record bursts of video, (iii) the storage to retain full-frame images (or the necessary metadata to reconstruct them from ROIs), and (iv) 3-D imaging capabilities. Color information, either as standard 3-channel color or hyperspectral, will facilitate studies seeking to quantify pigmentation, gut content, or bioluminescence. Recording video will make it easier to assess interactions between individuals. Retaining full frame images would allow revisiting them and detecting objects that were not targets during the initial processing. Increased storage capacity and speed will allow instruments to hold more data, such as videos and full frames, while maintaining deployment duration. 3-D imaging via stereo imaging, structure from motion, or Z-stack image acquisition would provide better descriptions of organism volume and other structures.

All of these new modalities are rapidly becoming realistic. Disk space is continually becoming less expensive, bus rates are getting faster, and embedded computers are getting more power efficient. While the computer components are available, actual construction and deployment will necessitate significant development time and effort. In particular, given enough power, on-board processing techniques could be explored to automatically detect possible interactions or other potentially interesting events to trigger video recording.

Beyond plankton

Plankton are far from the only organisms sampled by imaging systems. Indeed, ecologists collect untold exabytes of image and video data, including other marine organisms such as fish, benthic organisms (Beyan and Browman 2020), mammals (O'Connell et al. 2010; Karnowski et al. 2016) and freshwater benthic organisms (Milošević et al. 2020). Research groups that produce such data have already begun to leverage ML techniques to analyze their data, largely relying on object detection and taxonomic classification algorithms (Alken et al. 2019; Mahmood et al. 2020; Kloster et al. 2020). There remains much to be learned by studying these data streams with an eye toward trait-based analyses. So far, only a few studies have quantified traits automatically and have generally focused on size estimates (Álvarez-Ellacuría et al. 2020).

The computer vision methods we have identified could be adapted to estimate the number of parasites per fish in salmon farms (via object detection), the bleaching of coral reefs (via segmentation or deep regression (Nielsen et al. 2018)), or the cell deformations of benthic diatoms in response to pollutants (with key point estimation or deep regression). Conversely, plankton studies could benefit from approaches developed to automatically analyze images of other aquatic organisms (e.g., Beyan & Bromann 2020 and references therein). Scientists studying benthic macro-organisms, fish, or marine mammals with images will have unique sets of technological challenges associated with automating analysis (e.g. complex backgrounds for in situ benthic images or overlapping targets when recording pelagic fish). Despite such difficulties, functional traits could still be estimated from im-

ages and feed multi-compartment studies (Martini et al. 2021). The eventual combination of image data sets to cover the wide range of organism size and habitat will likely require both careful intercalibration and new mathematical methods.

Conclusion

As a final note, we would like to advocate for two goals that we should pursue as a community: (1) more open and efficient sharing of trait-annotated datasets, and (2) development of educational programs at the interface of computer science and ecology. The first goal is obvious and one that has been discussed widely. The benefits of open-access annotations are manifest: combining large volumes of publicly available labelled data with fine-tuning procedures could speed testing and deployment of new techniques. For instance, the Kaggle challenge on plankton image classification (<https://kaggle.com/c/datasciencebowl/>) has promoted the use of Convolutional Neural Networks (CNNs). This stressed how beneficial data sharing can be, since the problems that ecologists are facing with their images (e.g. taking size into account for ML or dealing with unbalanced datasets) can be challenging for data scientists as well. Therefore, aquatic ecologists should be encouraged to share or continue sharing their image data and models. The second goal will require more time and effort, but will likewise benefit the field as a whole. Most universities do not offer interdisciplinary tracks for students to pursue machine learning for ecological data. The establishment of dedicated programs that combine computer vision, machine learning, and aquatic ecology training should be promoted. Offering degrees and certificates at this interface would better prepare future scientists and establish a talent pool of creative practitioners.

Acknowledgements

This article emerged from the ideas discussed during the ARTIFACTZ workshop (Artificial intelligence application to the identification of functional traits of zooplankton from high-resolution images) funded by Sentinelle Nord (Université Laval, Québec), MODELIFE (UCA), and LOV (SU/

CNRS). The ARTIFACTZ workshop was organized by SDA, JOI and FM in Villefranche-sur-Mer, France, 24-26 April 2019. The book of abstracts of the workshop is available here: <https://hal.sorbonne-universite.fr/hal-02862387>. SDA acknowledges funding from CNRS for her sabbatical in 2018-2020. Additional support was provided by the Institut des Sciences du Calcul et des Données (ISCD) of Sorbonne Université (SU) through the support of the sponsored junior team FORMAL (From ObseRving to Modeling oceAn Life), especially through the post-doctoral contract of EO. French co-authors also wish to thank public taxpayers who fund their salaries. The authors have no conflicts of interest to declare.

This work is a contribution to the scientific program of Québec Océan and the Takuvik Joint International Laboratory (UMI3376; CNRS - Université Laval).

FM was supported by an NSERC Discovery Grant (RGPIN-2014-05433).

MS is supported by the Research Foundation - Flanders (FWO17/PDO/067).

FB received support from the European Union's Horizon 2020 research and innovation programme under grant agreement n°SEP-210591007.

MDO is supported by the Gordon and Betty Moore Foundation and the U.S. National Science Foundation.

ECB is supported by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) under the grant agreement n° 88882.438735/2019-01.

TB is supported by the French National Research Agency (ANR-19-CE01-0006).

NVP is supported by the Spanish State Research Agency, Ministry of Science and Innovation (PTA2016-12822-I)

FL is supported by the Institut Universitaire de France (IUF)

HMS was supported by the Simons Foundation (561126) and the U.S. National Science Foundation (CCF-1539256, OCE-1655686). Emily Peacock is gratefully acknowledged for expert annotation of IFCB images.

LS was supported by the Chair VISION from CNRS/Sorbonne Université.

References

- Acuña, J. L., Á. López-Urrutia, and S. Colin. 2011. Faking giants: the evolution of high prey clearance rates in jellyfishes. *science* **333**: 1627–1629.
- Aksnes, D. L., and J. Giske. 1990. Habitat profitability in pelagic environments. *Mar. Ecol. Prog. Ser.* Oldendorf **64**: 209–215.
- Allredge, A. L. 1981. The impact of appendicularian grazing on natural food concentrations in situ 1. *Limnol. Oceanogr.* **26**: 247–257.
- Allredge, A. L., and M. W. Silver. 1988. Characteristics, dynamics and significance of marine snow. *Prog. Oceanogr.* **20**: 41–82.
- Allken, V., N. O. Handegard, S. Rosen, T. Schreyeck, T. Mahiout, and K. Malde. 2019. Fish species identification using a convolutional neural network trained on synthetic data. *ICES J. Mar. Sci.* **76**: 342–349.
- Alp Güler, R., N. Neverova, and I. Kokkinos. 2018. Densepose: Dense human pose estimation in the wild. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 7297–7306.
- Álvarez-Ellacuría, A., M. Palmer, I. A. Catalán, and J.-L. Lisani. 2020. Image-based, unsupervised estimation of fish size from commercial landings using deep learning. *ICES J. Mar. Sci.* **77**: 1330–1339. doi:10.1093/icesjms/fsz216
- Andriluka, M., L. Pishchulin, P. Gehler, and B. Schiele. 2014. 2d human pose estimation: New benchmark and state of the art analysis. *Proceedings of the IEEE Conference on computer Vision and Pattern Recognition*. 3686–3693.
- Archibald, K. M., D. A. Siegel, and S. C. Doney. 2019. Modeling the Impact of Zooplankton Diel Vertical Migration on the Carbon Export Flux of the Biological Pump. *Glob. Biogeochem. Cycles* **33**: 181–199. doi:https://doi.org/10.1029/2018GB005983
- Badrinarayanan, V., A. Kendall, and R. Cipolla. 2017. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **39**: 2481–2495.
- Banase, K. 1976. Rates of Growth, Respiration and Photosynthesis of Unicellular Algae as Related to Cell Size—a Review^{1,2}. *J. Phycol.* **12**: 135–140. doi:10.1111/j.1529-8817.1976.tb00490.x
- Barton, A. D., A. J. Pershing, E. Litchman, N. R. Record, K. F. Edwards, Z. V. Finkel, T. Kiørboe, and B. A. Ward. 2013. The biogeography of marine plankton traits. *Ecol. Lett.* **16**: 522–534.
- Baumgartner, M. F., and A. M. Tarrant. 2017. The Physiology and Ecology of Diapause in Marine Copepods. *Annu. Rev. Mar. Sci.* **9**: 387–411. doi:10.1146/annurev-marine-010816-060505
- Benfield, M. C., P. Grosjean, P. F. Culverhouse, and others. 2007. RAPID: research on automated plankton identification. *Oceanography* **20**: 172–187.
- Bengio, Y., A. Courville, and P. Vincent. 2013. Representation learning: A review and new perspectives. *IEEE Trans. Pattern Anal. Mach. Intell.* **35**: 1798–1828.
- Beyan, C., and H. I. Browman. 2020. Setting the stage for the machine intelligence era in marine science. *ICES J. Mar. Sci.* **77**: 1267–1273. doi:10.1093/icesjms/fsaa084
- Biard, T., and M. D. Ohman. 2020. Vertical niche definition of test-bearing protists (Rhizaria) into the twilight zone revealed by in situ imaging. *Limnol. Oceanogr.* **65**: 2583–2602.
- Blanchard, J. L., R. F. Heneghan, J. D. Everett, R. Trebilco, and A. J. Richardson. 2017. From bacteria to whales: using functional size spectra to model marine ecosystems. *Trends Ecol. Evol.* **32**: 174–186.
- Blaschko, M. B., G. Holness, M. A. Mattar, D. Lisin, P. E. Utgoff, A. R. Hanson, H. Schultz, and E. M. Riseman. 2005. Automatic in situ identification of plankton. *Application of Computer Vision, 2005. WACV/MOTIONS'05 Volume 1. Seventh IEEE Workshops on. IEEE*. 79–86.
- Brosnahan, M. L., L. Velo-Suárez, D. K. Ralston, and others. 2015. Rapid growth and concerted sexual transitions by a bloom of the harmful dinoflagellate *Alexandrium fundyense* (Dinophyceae). *Limnol. Oceanogr.* **60**: 2059–2078. doi:10.1002/lno.10155
- Brun, P., M. R. Payne, and T. Kiørboe. 2016. Trait biogeography of marine copepods—an analysis across scales. *Ecol. Lett.* **19**: 1403–1413.
- Burridge, A. K., E. Goetze, D. Wall-Palmer, S. L. Le Double, J. Huisman, and K. T. Peijnenburg. 2017. Diversity and abundance of pteropods and heteropods along a latitudinal gradient

- across the Atlantic Ocean. *Prog. Oceanogr.* **158**: 213–223.
- Campbell, L., R. J. Olson, H. M. Sosik, A. Abraham, D. W. Henrichs, C. J. Hyatt, and E. J. Buskey. 2010. First Harmful Dinophysis (dinophyceae, Dinophysiales) Bloom in the U.S. Is Revealed by Automated Imaging Flow Cytometry1. *J. Phycol.* **46**: 66–75. doi:10.1111/j.1529-8817.2009.00791.x
- Cao, J., H. Tang, H.-S. Fang, X. Shen, C. Lu, and Y.-W. Tai. 2019. Cross-Domain Adaptation for Animal Pose Estimation. *Proc. IEEE CVF Int. Conf. Comput. Vis.* 9498–9507.
- Cao, Z., T. Simon, S.-E. Wei, and Y. Sheikh. 2017. Realtime multi-person 2d pose estimation using part affinity fields. *Proceedings of the IEEE conference on computer vision and pattern recognition.* 7291–7299.
- Carmona, C. P., F. De Bello, N. W. Mason, and J. Lepš. 2016. Traits without borders: integrating functional diversity across scales. *Trends Ecol. Evol.* **31**: 382–394.
- Cheng, K., X. Cheng, Y. Wang, H. Bi, and M. C. Benfield. 2019. Enhanced convolutional neural network for plankton identification and enumeration. *PLOS ONE* **14**: e0219570. doi:10.1371/journal.pone.0219570
- Conley, K. R., F. Lombard, and K. R. Sutherland. 2018. Mammoth grazers on the ocean’s minuteness: a review of selective feeding using mucous meshes. *Proc. R. Soc. B Biol. Sci.* **285**: 20180056. doi:10.1098/rspb.2018.0056
- Cowen, R. K., K. L. Robinson, and J. Luo. 2015. PlanktonSet 1.0: Plankton imagery data collected from F.G. Walton Smith in Straits of Florida from 2014-06-03 to 2014-06-06 and used in the 2015 National Data Science Bowl (NCEI Accession 0127422).doi:10.7289/v5d21vjd
- Cronin, H. A., J. H. Cohen, J. Berge, G. Johnsen, and M. A. Moline. 2016. Bioluminescence as an ecological factor during high Arctic polar night. *Sci. Rep.* **6**: 36374. doi:10.1038/srep36374
- Culverhouse, P. F., N. Macleod, R. Williams, M. C. Benfield, R. M. Lopes, and M. Picheral. 2014. An empirical assessment of the consistency of taxonomic identifications. *Mar. Biol. Res.* **10**: 73–84. doi:10.1080/17451000.2013.810762
- Culverhouse, P. F., R. Williams, B. Reguera, V. Herby, and S. González-Gil. 2003. Do experts make mistakes? A comparison of human and machine identification of dinoflagellates. *Mar. Ecol. Prog. Ser.* **247**: 5.
- Davis, C. S., S. M. Gallager, and A. R. Solow. 1992. Microaggregations of Oceanic Plankton Observed by Towed Video Microscopy. *Science* **257**: 230–232. doi:10.1126/science.257.5067.230
- Dijkman, N. A., and J. C. Kromkamp. 2006. Photosynthetic characteristics of the phytoplankton in the Scheldt estuary: community and single-cell fluorescence measurements. *Eur. J. Phycol.* **41**: 425–434. doi:10.1080/09670260600937791
- Donahue, J., Y. Jia, O. Vinyals, J. Hoffman, N. Zhang, E. Tzeng, and T. Darrell. 2014. Decaf: A deep convolutional activation feature for generic visual recognition. *Int. Conf. Mach. Learn.* 647–655.
- Du Clos, K. T., L. Karp-Boss, and B. J. Gemmill. 2021. Diatoms rapidly alter sinking behavior in response to changing nutrient concentrations. *Limnol. Oceanogr.* **66**: 892–900. doi:10.1002/lno.11649
- Elineau, A., C. Desnos, L. Jalabert, and others. 2018. ZooScanNet: plankton images captured with the ZooScan. doi:10.17882/55741
- Ellen, J. S., C. A. Graff, and M. D. Ohman. 2019. Improving plankton image classification using context metadata. *Limnol. Oceanogr. Methods* **17**: 439–461. doi:10.1002/lom3.10324
- Everingham, M., L. Van Gool, C. K. Williams, J. Winn, and A. Zisserman. 2010. The pascal visual object classes (voc) challenge. *Int. J. Comput. Vis.* **88**: 303–338.
- Felzenszwalb, P. F., R. B. Girshick, D. McAllester, and D. Ramanan. 2010. Object Detection with Discriminatively Trained Part-Based Models. *IEEE Trans. Pattern Anal. Mach. Intell.* **32**: 1627–1645. doi:10.1109/TPAMI.2009.167
- Flynn, K. J., D. K. Stoecker, A. Mitra, J. A. Raven, P. M. Glibert, P. J. Hansen, E. Granéli, and J. M. Burkholder. 2013. Misuse of the phytoplankton-zooplankton dichotomy: the need to assign organisms as mixotrophs within plankton functional types. *J. Plankton Res.* **35**: 3–11. doi:10.1093/plankt/fbs062
- Follows, M. J., S. Dutkiewicz, S. Grant, and S. W. Chisholm. 2007. Emergent biogeography of microbial communities in a model ocean. *science* **315**: 1843–1846.

- Fontana, S., M. K. Thomas, M. Moldoveanu, P. Spaak, and F. Pomati. 2018. Individual-level trait diversity predicts phytoplankton community properties better than species richness or evenness. *ISME J.* **12**: 356–366. doi:10.1038/ismej.2017.160
- Gallager, S. M., H. Yamazaki, and C. S. Davis. 2004. Contribution of fine-scale vertical structure and swimming behavior to formation of plankton layers on Georges Bank. *Mar. Ecol. Prog. Ser.* **267**: 27–43. doi:10.3354/meps267027
- García-Comas, C., L. Stemann, F. Ibanez, L. Berline, M. G. Mazzocchi, S. Gasparini, M. Picheral, and G. Gorsky. 2011. Zooplankton long-term changes in the NW Mediterranean Sea: Decadal periodicity forced by winter hydrographic conditions related to large-scale atmospheric changes? *J. Mar. Syst.* **87**: 216–226.
- Gaskell, D. E., M. D. Ohman, and P. M. Hull. 2019. Zooglider-based measurements of planktonic foraminifera in the California Current System. *J. Foraminifer. Res.* **49**: 390–404.
- Giering, S. L. C., S. R. Wells, K. M. J. Mayers, and others. 2019. Seasonal variation of zooplankton community structure and trophic position in the Celtic Sea: A stable isotope and biovolume spectrum approach. *Prog. Oceanogr.* **177**: 101943. doi:10.1016/j.pocean.2018.03.012
- Giering, S. L., B. Hosking, N. Briggs, and M. H. Iversen. 2020. The interpretation of particle size, shape, and carbon flux of marine particle images is strongly affected by the choice of particle detection algorithm. *Front. Mar. Sci.* **7**.
- Girshick, R., J. Donahue, T. Darrell, and J. Malik. 2014. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. *Proceedings of the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.* 580–587.
- Gomes-Pereira, J. N., V. Auger, K. Beisiegel, and others. 2016. Current and future trends in marine image annotation software. *Prog. Oceanogr.* **149**: 106–120. doi:10.1016/j.pocean.2016.07.005
- González, P., A. Castaño, N. V. Chawla, and J. J. D. Coz. 2017. A review on quantification learning. *ACM Comput. Surv. CSUR* **50**: 74.
- González, P., A. Castaño, E. E. Peacock, J. Díez, J. J. Del Coz, and H. M. Sosik. 2019. Automatic plankton quantification using deep features. *J. Plankton Res.* **41**: 449–463. doi:10.1093/plankt/fbz023
- Gorsky, G., M. D. Ohman, M. Picheral, and others. 2010. Digital zooplankton image analysis using the ZooScan integrated system. *J. Plankton Res.* **32**: 285–303. doi:10.1093/plankt/fbp124
- Grosjean, P., M. Picheral, C. Warembourg, and G. Gorsky. 2004. Enumeration, measurement, and identification of net zooplankton samples using the ZOOSCAN digital imaging system. *ICES J. Mar. Sci. J. Cons.* **61**: 518–525.
- Gupta, A., P. Dollar, and R. Girshick. 2019. Lvis: A dataset for large vocabulary instance segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.* 5356–5364.
- Haddock, S. H., M. A. Moline, and J. F. Case. 2010. Bioluminescence in the sea. *Annu. Rev. Mar. Sci.* **2**: 443–493.
- Hansen, B., P. K. Bjornsen, and P. J. Hansen. 1994. The size ratio between planktonic predators and their prey. *Limnol. Oceanogr.* **39**: 395–403.
- Harvey, E. L., and S. Menden-Deuer. 2012. Predator-induced fleeing behaviors in phytoplankton: a new mechanism for harmful algal bloom formation? *PLoS One* **7**: e46438.
- He, K., G. Gkioxari, P. Dollar, and R. Girshick. 2017. Mask R-CNN. *Proceedings of the Proceedings of the IEEE International Conference on Computer Vision.* 2961–2969.
- Heuschele, J., S. Eliassen, and T. Kiørboe. 2013. Optimal mate choice patterns in pelagic copepods. *Oecologia* **172**: 399–408. doi:10.1007/s00442-012-2516-4
- Hillebrand, H., C.-D. Dürselen, D. Kirschtel, U. Pollinger, and T. Zohary. 1999. Biovolume Calculation for Pelagic and Benthic Microalgae. *J. Phycol.* **35**: 403–424. doi:10.1046/j.1529-8817.1999.3520403.x
- Hirst, A. G. 2012. Intraspecific scaling of mass to length in pelagic animals: Ontogenetic shape change and its implications. *Limnol. Oceanogr.* **57**: 1579–1590.
- Hofmann, P., A. Chatzinotas, W. S. Harpole, and S. Dunker. 2019. Temperature and stoichiometric dependence of phytoplankton traits. *Ecology* **100**: e02875. doi:https://doi.org/10.1002/ecy.2875

- Hoiem, D., Y. Chodpathumwan, and Q. Dai. 2012. Diagnosing error in object detectors. *European conference on computer vision*. Springer. 340–353.
- Hylander, S., T. Kiørboe, P. Snoeijs, R. Sommaruga, and T. G. Nielsen. 2015. Concentrations of sunscreens and antioxidant pigments in Arctic *Calanus* spp. in relation to ice cover, ultraviolet radiation, and the phytoplankton spring bloom. *Limnol. Oceanogr.* **60**: 2197–2206.
- Ikedo, T. 2014. Respiration and ammonia excretion by marine metazooplankton taxa: synthesis toward a global-bathymetric model. *Mar. Biol.* **161**: 2753–2766.
- Ikedo, T., Y. Kanno, K. Ozaki, and A. Shinada. 2001. Metabolic rates of epipelagic marine copepods as a function of body mass and temperature. *Mar. Biol.* **139**: 587–596. doi:10.1007/s002270100608
- Ikedo, T., F. Sano, and A. Yamaguchi. 2007. Respiration in marine pelagic copepods: a global-bathymetric model. *Mar. Ecol. Prog. Ser.* **339**: 215–219.
- Irisson, J.-O., S.-D. Ayata, D. J. Lindsay, L. Karp-Boss, and L. Stemann. In press. Machine Learning for the Study of Plankton and Marine Snow from Images. *Annu. Rev. Mar. Sci.* **14**. doi:10.1146/annurev-marine-041921-013023
- Jaspers, C., L. Marty, and T. Kiørboe. 2018. Selection for life-history traits to maximize population growth in an invasive marine species. *Glob. Change Biol.* **24**: 1164–1174. doi:https://doi.org/10.1111/gcb.13955
- Jiang, H., and G.-A. Paffenhöfer. 2008. Hydrodynamic signal perception by the copepod *Oithona plumifera*. *Mar. Ecol. Prog. Ser.* **373**: 37–52.
- Jónasdóttir, S. H., A. W. Visser, K. Richardson, and M. R. Heath. 2015. Seasonal copepod lipid pump promotes carbon sequestration in the deep North Atlantic. *Proc. Natl. Acad. Sci.* **112**: 12122–12126.
- Jónasdóttir, S. H., R. J. Wilson, A. Gislason, and M. R. Heath. 2019. Lipid content in overwintering *Calanus finmarchicus* across the Subpolar Eastern North Atlantic Ocean. *Limnol. Oceanogr.* **64**: 2029–2043. doi:https://doi.org/10.1002/lno.11167
- Jung, J., S.-J. Hong, H.-B. Kim, and others. 2018. Label-free non-invasive quantitative measurement of lipid contents in individual microalgal cells using refractive index tomography. *Sci. Rep.* **8**: 6524. doi:10.1038/s41598-018-24393-0
- Karnowski, J., C. Johnson, and E. Hutchins. 2016. Automated video surveillance for the study of marine mammal behavior and cognition. *Anim. Behav. Cogn.* **3**: 255–264.
- Katija, K., P. L. D. Roberts, J. Daniels, and others. 2021. Visual Tracking of Deepwater Animals Using Machine Learning-Controlled Robotic Underwater Vehicles. Proceedings of the Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. 860–869.
- Katija, K., R. E. Sherlock, A. D. Sherman, and B. H. Robison. 2017. New technology reveals the role of giant larvaceans in oceanic carbon cycling. *Sci. Adv.* **3**: e1602374. doi:10.1126/sciadv.1602374
- Kenitz, K. M., E. C. Orenstein, P. L. Roberts, P. J. Franks, J. S. Jaffe, M. L. Carter, and A. D. Barton. 2020. Environmental drivers of population variability in colony-forming marine diatoms. *Limnol. Oceanogr.* **65**: 2515–2528. doi:10.1002/lno.11468
- Kiko, R., D. Bianchi, C. Grenz, H. Hauss, M. Iversen, S. Kumar, A. Maas, and C. Robinson. 2020. Editorial: Zooplankton and Nekton: Gatekeepers of the Biological Pump. *Front. Mar. Sci.* **7**. doi:10.3389/fmars.2020.00545
- King, A., S. M. Bhandarkar, and B. M. Hopkinson. 2018. A Comparison of Deep Learning Methods for Semantic Segmentation of Coral Reef Survey Images. Proceedings of the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. 1394–1402.
- Kiørboe, T. 2011a. How zooplankton feed: mechanisms, traits and trade-offs. *Biol. Rev.* **86**: 311–339.
- Kiørboe, T. 2011b. What makes pelagic copepods so successful? *J. Plankton Res.* **33**: 677–685. doi:10.1093/plankt/fbq159
- Kiørboe, T., A. Andersen, V. J. Langlois, and H. H. Jakobsen. 2010. Unsteady motion: escape jumps in planktonic copepods, their kinematics and energetics. *J. R. Soc. Interface* **7**: 1591–1602.
- Kiørboe, T., S. Ceballos, and U. H. Thygesen. 2015. Interrelations between senescence, life-history traits, and behavior in planktonic copepods.

- Ecology **96**: 2225–2235. doi:<https://doi.org/10.1890/14-2205.1>
- Kjørboe, T., A. Visser, and K. H. Andersen. 2018. A trait-based approach to ocean ecology. *ICES J. Mar. Sci.* **75**: 1849–1863.
- Kloster, M., D. Langenkämper, M. Zurowietz, B. Beszteri, and T. W. Nattkemper. 2020. Deep learning-based diatom taxonomy on virtual slides. *Sci. Rep.* **10**: 14416. doi:[10.1038/s41598-020-71165-w](https://doi.org/10.1038/s41598-020-71165-w)
- Kruk, C., C. Piccini, M. Devercelli, L. Nogueira, V. Accattatis, L. Sampognaro, and A. M. Segura. 2021. A trait-based approach predicting community assembly and dominance of microbial invasive species. *Oikos*. doi:<https://doi.org/10.1111/oik.07694>
- Lathuilière, S., P. Mesejo, X. Alameda-Pineda, and R. Horaud. 2020. A Comprehensive Analysis of Deep Regression. *IEEE Trans. Pattern Anal. Mach. Intell.* **42**: 2065–2081. doi:[10.1109/TPAMI.2019.2910523](https://doi.org/10.1109/TPAMI.2019.2910523)
- Lauffer, M., F. Genty, J. L. Collette, and S. Margueron. 2015. Phytoplankton identification by combined methods of morphological processing and fluorescence imaging. *2015 IEEE Workshop on Environmental, Energy, and Structural Monitoring Systems (EESMS) Proceedings*. Proceedings of the 2015 IEEE Workshop on Environmental, Energy, and Structural Monitoring Systems (EESMS) Proceedings. 131–135.
- Le Quéré, C., E. T. Buitenhuis, R. Moriarty, and others. 2016. Role of zooplankton dynamics for Southern Ocean phytoplankton biomass and global biogeochemical cycles. *Biogeosciences* **13**: 4111–4133. doi:<https://doi.org/10.5194/bg-13-4111-2016>
- LeCun, Y., Y. Bengio, and G. Hinton. 2015. Deep learning. *Nature* **521**: 436–444. doi:[10.1038/nature14539](https://doi.org/10.1038/nature14539)
- Lefcheck, J. S., V. A. Bastazini, and J. N. Griffin. 2015. Choosing and using multiple traits in functional diversity research. *Environ. Conserv.* **42**: 104–107.
- Legras, G., N. Loiseau, and J.-C. Gaertner. 2018. Functional richness: Overview of indices and underlying concepts. *Acta Oecologica* **87**: 34–44. doi:[10.1016/j.actao.2018.02.007](https://doi.org/10.1016/j.actao.2018.02.007)
- Legras, G., N. Loiseau, J.-C. Gaertner, J.-C. Pogiale, and N. Gaertner-Mazouni. 2020. Assessing functional diversity: the influence of the number of the functional traits. *Theor. Ecol.* **13**: 117–126.
- Lehette, P., and S. Hernández-León. 2009. Zooplankton biomass estimation from digitized images: a comparison between subtropical and Antarctic organisms. *Limnol. Oceanogr. Methods* **7**: 304–308. doi:<https://doi.org/10.4319/lom.2009.7.304>
- Li, S., S. Gunel, M. Ostrek, P. Ramdya, P. Fua, and H. Rhodin. 2020. Deformation-Aware Unpaired Image Translation for Pose Estimation on Laboratory Animals. *Proc. IEEE CVF Conf. Comput. Vis. Pattern Recognit.* 13158–13168.
- Lienhart, R., and J. Maydt. 2002. An extended set of Haar-like features for rapid object detection. *Proceedings. International Conference on Image Processing*. Proceedings of the Proceedings. International Conference on Image Processing. I–I.
- Lin, T.-Y., M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick. 2014. Microsoft coco: Common objects in context. *European conference on computer vision*. Springer. 740–755.
- Litchman, E., and C. A. Klausmeier. 2008. Trait-based community ecology of phytoplankton. *Annu. Rev. Ecol. Evol. Syst.* **39**: 615–639.
- Litchman, E., M. D. Ohman, and T. Kjørboe. 2013. Trait-based approaches to zooplankton communities. *J. Plankton Res.* **35**: 473–484. doi:[10.1093/plankt/fbt019](https://doi.org/10.1093/plankt/fbt019)
- Litchman, E., P. de T. Pinto, K. F. Edwards, C. A. Klausmeier, C. T. Kremer, and M. K. Thomas. 2015. Global biogeochemical impacts of phytoplankton: a trait-based perspective. *J. Ecol.* **103**: 1384–1396. doi:<https://doi.org/10.1111/1365-2745.12438>
- Liu, W., D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg. 2016a. SSD: Single Shot MultiBox Detector. *Computer Vision – ECCV 2016*. Springer International Publishing. 21–37.
- Liu, X., W. Liang, Y. Wang, S. Li, and M. Pei. 2016b. 3D head pose estimation with convolutional neural network trained on synthetic images. *2016 IEEE International Conference on Image Processing (ICIP)*. Proceedings of the 2016 IEEE International Conference on Image Processing (ICIP).

- ternational Conference on Image Processing (ICIP). 1289–1293.
- Lombard, F., E. Boss, A. M. Waite, and others. 2019. Globally Consistent Quantitative Observations of Planktonic Ecosystems. *Front. Mar. Sci.* **6**: 196. doi:10.3389/fmars.2019.00196
- Lombard, F., and T. Kiørboe. 2010. Marine snow originating from appendicularian houses: Age-dependent settling characteristics. *Deep Sea Res. Part Oceanogr. Res. Pap.* **57**: 1304–1313.
- Long, J., E. Shelhamer, and T. Darrell. 2015. Fully Convolutional Networks for Semantic Segmentation. Proceedings of the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 3431–3440.
- Luo, J. Y., J.-O. Irisson, B. Graham, C. Guigand, A. Sarafraz, C. Mader, and R. K. Cowen. 2018. Automated plankton image analysis using convolutional neural networks. *Limnol. Oceanogr. Methods* **16**: 814–827. doi:10.1002/lom3.10285
- MacLeod, N., M. Benfield, and P. Culverhouse. 2010. Time to automate identification. *Nature* **467**: 154–155. doi:10.1038/467154a
- Mahmood, A., M. Bennamoun, S. An, F. Sohel, F. Boussaid, R. Hovey, and G. Kendrick. 2020. Automatic detection of Western rock lobster using synthetic data. *ICES J. Mar. Sci.* **77**: 1308–1317. doi:10.1093/icesjms/fsz223
- Maps, F., N. R. Record, and A. J. Pershing. 2014. A metabolic approach to dormancy in pelagic copepods helps explaining inter- and intra-specific variability in life-history strategies. *J. Plankton Res.* **36**: 18–30. doi:10.1093/plankt/fbt100
- Martini, S., and S. H. Haddock. 2017. Quantification of bioluminescence from the surface to the deep sea demonstrates its predominance as an ecological trait. *Sci. Rep.* **7**: 1–11.
- Martini, S., F. Larras, A. Boyé, and others. 2021. Functional trait-based approaches as a common framework for aquatic ecologists. *Limnol. Oceanogr.* **66**: 965–994. doi:https://doi.org/10.1002/lno.11655
- McConville, K., A. Atkinson, E. S. Fileman, J. I. Spicer, and A. G. Hirst. 2017. Disentangling the counteracting effects of water content and carbon mass on zooplankton growth. *J. Plankton Res.* **39**: 246–256.
- Méléder, V., M. Laviale, B. Jesus, J. L. Mouget, J. Lavaud, F. Kazemipour, P. Launeau, and L. Barillé. 2013. In vivo estimation of pigment composition and optical absorption cross-section by spectroradiometry in four aquatic photosynthetic micro-organisms. *J. Photochem. Photobiol. B* **129**: 115–124. doi:10.1016/j.jphotobiol.2013.10.005
- Menden-Deuer, S., and E. J. Lessard. 2000. Carbon to volume relationships for dinoflagellates, diatoms, and other protist plankton. *Limnol. Oceanogr.* **45**: 569–579. doi:10.4319/lo.2000.45.3.0569
- Messié, M., I. Shulman, S. Martini, and S. H. Haddock. 2019. Using fluorescence and bioluminescence sensors to characterize auto- and heterotrophic plankton communities. *Prog. Oceanogr.* **171**: 76–92.
- Milošević, D., A. Milosavljević, B. Predić, and others. 2020. Application of deep learning in aquatic bioassessment: Towards automated identification of non-biting midges. *Sci. Total Environ.* **711**: 135160. doi:10.1016/j.scitotenv.2019.135160
- Moberg, E. A., and H. M. Sosik. 2012. Distance maps to estimate cell volume from two-dimensional plankton images. *Limnol. Oceanogr. Methods* **10**: 278–288.
- Möller, K. O., J. O. Schmidt, M. St. John, and others. 2015. Effects of climate-induced habitat changes on a key zooplankton species. *J. Plankton Res.* **37**: 530–541.
- Moreno-Torres, J. G., T. Raeder, R. Alaiz-Rodríguez, N. V. Chawla, and F. Herrera. 2012. A unifying view on dataset shift in classification. *Pattern Recognit.* **45**: 521–530.
- Mouillot, D., N. A. Graham, S. Villéger, N. W. Mason, and D. R. Bellwood. 2013. A functional approach reveals community responses to disturbances. *Trends Ecol. Evol.* **28**: 167–177.
- Niehoff, B. 2003. Gonad morphology and oocyte development in *Pseudocalanus* spp. in relation to spawning activity. *Mar. Biol.* **143**: 759–768.
- Niehoff, B. 2007. Life history strategies in zooplankton communities: The significance of female gonad morphology and maturation types for the reproductive biology of marine calanoid copepods. *Prog. Oceanogr.* **74**: 1–47. doi:10.1016/j.pocean.2006.05.005

- Nielsen, D. A., K. Petrou, and R. D. Gates. 2018. Coral bleaching from a single cell perspective. *ISME J.* **12**: 1558–1567. doi:10.1038/s41396-018-0080-6
- Nielsen, S. L. 2006. Size-dependent growth rates in eukaryotic and prokaryotic algae exemplified by green algae and cyanobacteria: comparisons between unicells and colonial growth forms. *J. Plankton Res.* **28**: 489–498.
- Nishibe, Y., K. Takahashi, T. Ichikawa, K. Hidaka, H. Kurogi, K. Segawa, and H. Saito. 2015. Degradation of discarded appendicularian houses by oncaeid copepods. *Limnol. Oceanogr.* **60**: 967–976. doi:10.1002/lno.10061
- O’Connell, A. F., J. D. Nichols, and K. U. Karanth. 2010. *Camera traps in animal ecology: methods and analyses*, Springer Science & Business Media.
- Ohman, M. D. 2019. A sea of tentacles: optically discernible traits resolved from planktonic organisms in situ. *ICES J. Mar. Sci.* **76**: 1959–1972.
- Ohman, M. D., R. E. Davis, J. T. Sherman, K. R. Grindley, B. M. Whitmore, C. F. Nickels, and J. S. Ellen. 2018. Zooglider: An autonomous vehicle for optical and acoustic sensing of zooplankton. *Limnol. Oceanogr. Methods* **0**. doi:10.1002/lom3.10301
- Ohman, M. D., and J.-B. Romagnan. 2016. Nonlinear effects of body size and optical attenuation on Diel Vertical Migration by zooplankton. *Limnol. Oceanogr.* **61**: 765–770.
- Olson, R. J., A. M. Chekalyuk, and H. M. Sosik. 1996. Phytoplankton photosynthetic characteristics from fluorescence induction assays of individual cells. *Limnol. Oceanogr.* **41**: 1253–1263. doi:10.4319/lno.1996.41.6.1253
- Olson, R. J., A. Shalapyonok, and H. M. Sosik. 2003. An automated submersible flow cytometer for analyzing pico- and nanophytoplankton: FlowCytobot. *Deep Sea Res. Part Oceanogr. Res. Pap.* **50**: 301–315. doi:10.1016/S0967-0637(03)00003-7
- Olson, R. J., and H. M. Sosik. 2007. A submersible imaging-in-flow instrument to analyze nano- and microplankton: Imaging FlowCytobot. *Limnol Ocean. Methods* **5**: 195–203.
- Olson, R. J., E. R. Zettler, S. W. Chisholm, and J. A. Dusenberry. 1991. Advances in Oceanography through Flow Cytometry, p. 351–399. *In* S. Demers [ed.], *Particle Analysis in Oceanography*. Springer.
- Orenstein, E. C., and O. Beijbom. 2017. Transfer Learning and Deep Feature Extraction for Planktonic Image Data Sets. *2017 IEEE Winter Conference on Applications of Computer Vision (WACV)*. Proceedings of the 2017 IEEE Winter Conference on Applications of Computer Vision (WACV). 1082–1088.
- Orenstein, E. C., D. Ratelle, C. Briseño-Avena, M. L. Carter, P. J. S. Franks, J. S. Jaffe, and P. L. D. Roberts. 2020. The Scripps Plankton Camera system: A framework and platform for in situ microscopy. *Limnol. Oceanogr. Methods* **18**: 681–695. doi:10.1002/lom3.10394
- Ortner, P. B., S. R. Cummings, R. P. Aftiring, and H. E. Edgerton. 1979. Silhouette photography of oceanic zooplankton. *Nature* **277**: 50–51.
- Ouyang, W., X. Wang, C. Zhang, and X. Yang. 2016. Factors in Finetuning Deep Model for Object Detection With Long-Tail Distribution. Proceedings of the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 864–873.
- Papandreou, G., T. Zhu, N. Kanazawa, A. Toshev, J. Tompson, C. Bregler, and K. Murphy. 2017. Towards accurate multi-person pose estimation in the wild. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 4903–4911.
- Pasternak, A. F. 1994. Gut fluorescence in herbivorous copepods: an attempt to justify the method. *Hydrobiologia* **292**: 241–248.
- Peacock, E. E., R. J. Olson, and H. M. Sosik. 2014. Parasitic infection of the diatom *Guinardia delicatula*, a recurrent and ecologically important phenomenon on the New England Shelf. *Mar Ecol Prog Ser* **503**: 1–10.
- Picheral, M., S. Colin, and J. O. Irisson. 2017. EcoTaxa, a tool for the taxonomic classification of images.
- Picheral, M., L. Guidi, L. Stemann, D. M. Karl, G. Iddaoud, and G. Gorsky. 2010. The Underwater Vision Profiler 5: An advanced instrument for high spatial resolution studies of particle size spectra and zooplankton. *Limnol. Oceanogr. Methods* **8**: 462–473.
- Platt, D. T., and K. Denman. 1977. Organisation in the pelagic ecosystem. *Helgoländer Wiss.*

- Meeresunters. **30**: 575–581. doi:10.1007/BF02207862
- Priede, I. G., P. M. Bagley, S. Way, P. J. Herring, and J. C. Partridge. 2006. Bioluminescence in the deep sea: Free-fall lander observations in the Atlantic Ocean off Cape Verde. *Deep Sea Res. Part Oceanogr. Res. Pap.* **53**: 1272–1283.
- Redmon, J., S. Divvala, R. Girshick, and A. Farhadi. 2016. You Only Look Once: Unified, Real-Time Object Detection. *Proceedings of the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.* 779–788.
- Reimann, R., B. Zeng, M. Jakopec, M. Burdukiewicz, I. Petrick, P. Schierack, and S. Rödigger. 2020. Classification of dead and living microalgae *Chlorella vulgaris* by bioimage informatics and machine learning. *Algal Res.* **48**: 101908. doi:10.1016/j.algal.2020.101908
- Rembauville, M., S. Blain, C. Manno, G. Tarling, A. Thompson, G. Wolff, and I. Salter. 2018. The role of diatom resting spores in pelagic–benthic coupling in the Southern Ocean. *Biogeosciences* **15**: 3071–3084. doi:10.5194/bg-15-3071-2018
- Remsen, A., T. L. Hopkins, and S. Samson. 2004. What you see is not what you catch: a comparison of concurrently collected net, Optical Plankton Counter, and Shadowed Image Particle Profiling Evaluation Recorder data from the northeast Gulf of Mexico. *Deep Sea Res. Part Oceanogr. Res. Pap.* **51**: 129–151. doi:10.1016/j.dsr.2003.09.008
- Rezatofghi, H., N. Tsoi, J. Gwak, A. Sadeghian, I. Reid, and S. Savarese. 2019. Generalized intersection over union: A metric and a loss for bounding box regression. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.* 658–666.
- Romagnan, J.-B., L. Legendre, L. Guidi, and others. 2015. Comprehensive Model of Annual Plankton Succession Based on the Whole-Plankton Time Series Approach. *PLOS ONE* **10**: e0119219. doi:10.1371/journal.pone.0119219
- Ronneberger, O., P. Fischer, and T. Brox. 2015. U-net: Convolutional networks for biomedical image segmentation. *International Conference on Medical image computing and computer-assisted intervention.* Springer. 234–241.
- Rothe, R., R. Timofte, and L. Van Gool. 2018. Deep expectation of real and apparent age from a single image without facial landmarks. *Int. J. Comput. Vis.* **126**: 144–157.
- Ryabov, A., O. Kerimoglu, E. Litchman, I. Olenina, L. Roselli, A. Basset, E. Stanca, and B. Blasius. 2021. Shape matters: the relationship between cell geometry and diversity in phytoplankton. *Ecol. Lett.* doi:https://doi.org/10.1111/ele.13680
- Sainmont, J. 2014. Strategies for success: Copepods in a seasonal world. National Institute of Aquatic Resources - Technical University of Denmark.
- Salter, I., A. E. Kemp, C. M. Moore, R. S. Lampitt, G. A. Wolff, and J. Holtvoeth. 2012. Diatom resting spore ecology drives enhanced carbon export from a naturally iron-fertilized bloom in the Southern Ocean. *Glob. Biogeochem. Cycles* **26**.
- Sapp, B., and B. Taskar. 2013. Modex: Multimodal decomposable models for human pose estimation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.* 3674–3681.
- Schleuter, D., M. Daufresne, F. Massol, and C. Argillier. 2010. A user’s guide to functional diversity indices. *Ecol. Monogr.* **80**: 469–484.
- Schmera, D., J. Heino, J. Podani, T. Erőss, and S. Dolédec. 2017. Functional diversity: a review of methodology and current knowledge in freshwater macroinvertebrate research. *Hydrobiologia* **787**: 27–44.
- Schmid, M. S., F. Maps, and L. Fortier. 2018. Lipid load triggers migration to diapause in Arctic *Calanus* copepods—insights from underwater imaging. *J. Plankton Res.* **40**: 311–325. doi:10.1093/plankt/fby012
- Schneider, T., G. Grosbois, W. F. Vincent, and M. Rautio. 2016. Carotenoid accumulation in copepods is related to lipid metabolism and reproduction rather than to UV-protection. *Limnol. Oceanogr.* **61**: 1201–1213.
- Schröder, S.-M., R. Kiko, and R. Koch. 2020. MorphoCluster: Efficient Annotation of Plankton Images by Clustering. *Sensors* **20**: 3060.
- Serra-Pompei, C., F. Soudijn, A. W. Visser, T. Kiørboe, and K. H. Andersen. 2020. A general size- and trait-based model of plankton communities. *Prog. Oceanogr.* **189**: 102473.
- Shaw, C. T., H. Bi, L. R. Feinberg, and W. T. Peterson. 2021. Cohort analysis of *Euphausia pacifica*

- from the Northeast Pacific population using a Gaussian mixture model. *Prog. Oceanogr.* **191**: 102495. doi:10.1016/j.pocean.2020.102495
- Sheldon, R. W., A. Prakash, and Wh. Sutcliffe Jr. 1972. The size distribution of particles in the Ocean 1. *Limnol. Oceanogr.* **17**: 327–340.
- Sieburth, J. M., V. Smetacek, and J. Lenz. 1978. Pelagic ecosystem structure: Heterotrophic compartments of the plankton and their relationship to plankton size fractions 1. *Limnol. Oceanogr.* **23**: 1256–1263. doi:10.4319/lo.1978.23.6.1256
- Sieracki, M. E., M. Benfield, A. Hanson, and others. 2010. Optical plankton imaging and analysis systems for ocean observation. *Proc. Ocean Obs* **9**: 21–25.
- Simpson, R., P. F. Culverhouse, R. Ellis, and R. Williams. 1991. Classification of Euceratium Gran. in neural networks. *IEEE International Conference on Neural Networks in Ocean Engineering, Washington, DC, USA.* 223–230.
- Sosik, H. M., and R. J. Olson. 2007. Automated taxonomic classification of phytoplankton sampled with imaging-in-flow cytometry. *Limnol. Oceanogr. Methods* **5**: 204–216. doi:10.4319/lom.2007.5.204
- Sosik, H. M., E. E. Peacock, and E. F. Brownlee. 2014. WHOI Plankton: Annotated Plankton Images - Data Set for Developing and Evaluating Classification Methods. doi:10.1575/1912/7341
- Sourisseau, M., Y. Simard, and F. J. Saucier. 2008. Krill diel vertical migration fine dynamics, nocturnal overturns, and their roles for aggregation in stratified flows. *Can. J. Fish. Aquat. Sci.* **65**: 574–587.
- Stamieszkin, K., A. J. Pershing, N. R. Record, C. H. Pilskalns, H. G. Dam, and L. R. Feinberg. 2015. Size as the master trait in modeled copepod fecal pellet carbon flux. *Limnol. Oceanogr.* **60**: 2090–2107. doi:10.1002/lno.10156
- Stemmann, L., and E. Boss. 2012. Plankton and Particle Size and Packaging: From Determining Optical Properties to Driving the Biological Pump. *Annu. Rev. Mar. Sci.* **4**: 263–290. doi:10.1146/annurev-marine-120710-100853
- Sun, C., A. Shrivastava, S. Singh, and A. Gupta. 2017. Revisiting Unreasonable Effectiveness of Data in Deep Learning Era. *Proceedings of the Proceedings of the IEEE International Conference on Computer Vision.* 843–852.
- Tarling, G. A. 2003. Sex-dependent diel vertical migration in northern krill *Meganyctiphanes norvegica* and its consequences for population dynamics. *Mar. Ecol. Prog. Ser.* **260**: 173–188. doi:10.3354/meps260173
- Tharwat, A. 2020. Classification assessment methods. *Appl. Comput. Inform.* **ahead-of-print**. doi:10.1016/j.aci.2018.08.003
- Trudnowska, E., L. Lacour, M. Ardyna, A. Rogge, J. O. Irisson, A. M. Waite, M. Babin, and L. Stemmann. 2021. Marine snow morphology illuminates the evolution of phytoplankton blooms and determines their subsequent vertical export. *Nat. Commun.* **12**: 1–13.
- Vandromme, P., L. Stemmann, C. Garcia-Comas, L. Berline, X. Sun, and G. Gorsky. 2012. Assessing biases in computing size spectra of automatically classified zooplankton from imaging systems: A case study with the ZooScan integrated system. *Methods Oceanogr.* **1**: 3–21.
- de Vargas, C., S. Audic, N. Henry, and others. 2015. Eukaryotic plankton diversity in the sunlit ocean. *Science* **348**. doi:10.1126/science.1261605
- Vilgrain, L., F. Maps, M. Picheral, M. Babin, C. Aubry, J.-O. Irisson, and S.-D. Ayata. 2021. Trait-based approach using in situ copepod images reveals contrasting ecological patterns across an Arctic ice melt zone. *Limnol. Oceanogr.* **n/a**. doi:https://doi.org/10.1002/lno.11672
- Villareal, T. A. 2004. Single-Cell Pulse Amplitude Modulation Fluorescence Measurements of the Giant Diatom *Ethmodiscus* (bacillariophyceae)1. *J. Phycol.* **40**: 1052–1061. doi:10.1111/j.1529-8817.2004.03208.x
- Villéger, S., N. W. Mason, and D. Mouillot. 2008. New multidimensional functional diversity indices for a multifaceted framework in functional ecology. *Ecology* **89**: 2290–2301.
- Viola, P., and M. J. Jones. 2004. Robust Real-Time Face Detection. *Int. J. Comput. Vis.* **57**: 137–154. doi:10.1023/B:VISI.0000013087.49260.fb
- Violle, C., M.-L. Navas, D. Vile, E. Kazakou, C. Fortunel, I. Hummel, and E. Garnier. 2007. Let the concept of trait be functional! *Oikos* **116**: 882–892. doi:https://doi.org/10.1111/j.0030-1299.2007.15559.x

- Ward, B. A., S. Dutkiewicz, O. Jahn, and M. J. Follows. 2012. A size-structured food-web model for the global ocean. *Limnol. Oceanogr.* **57**: 1877–1891. doi:<https://doi.org/10.4319/lo.2012.57.6.1877>
- Whitmore, B. M., C. F. Nickels, and M. D. Ohman. 2019. A comparison between Zooglider and shipboard net and acoustic mesozooplankton sensing systems. *J. Plankton Res.* **41**: 521–533.
- Wirtz, K. W. 2013. How fast can plankton feed? Maximum ingestion rate scales with digestive surface area. *J. Plankton Res.* **35**: 33–48. doi:[10.1093/plankt/fbs075](https://doi.org/10.1093/plankt/fbs075)
- Xu, Z., Y. Jiang, J. Ji, J. Ji, E. Forsberg, Y. Li, S. He, and S. He. 2020. Classification, identification, and growth stage estimation of microalgae based on transmission hyperspectral microscopic imaging and machine learning. *Opt. Express* **28**: 30686–30700. doi:[10.1364/OE.406036](https://doi.org/10.1364/OE.406036)
- Yosinski, J., J. Clune, Y. Bengio, and H. Lipson. 2014. How transferable are features in deep neural networks?, p. 3320–3328. In Z. Ghahramani, M. Welling, C. Cortes, N.D. Lawrence, and K.Q. Weinberger [eds.], *Advances in Neural Information Processing Systems 27*. Curran Associates, Inc.
- Zhao, Z.-Q., P. Zheng, S.-T. Xu, and X. Wu. 2019. Object Detection With Deep Learning: A Review. *IEEE Trans. Neural Netw. Learn. Syst.* **30**: 3212–3232. doi:[10.1109/TNNLS.2018.2876865](https://doi.org/10.1109/TNNLS.2018.2876865)

S1. Supplemental Information

Accuracy (*acc*) is by far the most widely used and intuitive evaluation metric for classification problems. In the binary case, it is simply the number of correctly returned labels in the “positive class” over all classified data points. For multiclass systems, *acc* can be computed as the sum of the diagonals of the confusion matrix divided by the total number of objects classified. In both cases *acc* ranges from 0 to 1 with values closer to 0 indicating poor performance and closer to 1 suggesting near perfect performance. *acc* provides a useful snapshot of classifier performance, but obscures information regarding the source of errors and is very sensitive to class distribution (Tharwat 2020).

Precision (*p*) and Recall (*r*) are related to *acc* and similarly reflect a classifier’s performance. *p* reflects how accurate the predictions are, while *r* describes the classifier’s ability to retrieve relevant data points. These two metrics are tightly coupled: when a system’s precision increases, its recall decreases since it returns more of the possible true labels at the expense of more false positives.

Average Precision (*AP*) summarizes the trade off between the precision and recall for a given class by averaging the precision for different levels of the recall. Computing *AP* requires the algorithm to return either probability of class membership or confidence values rather than hard class assignments (Everingham et al. 2010; Hoiem et al. 2012). The better the model is at prediction, the higher the precision at all values of recall thus increasing *AP*. *AP* ranges between 0 and 1, with 1 being a perfect model that returns all relevant objects with high confidence. *AP* is computed on a class-by-class basis and can be summarized for a whole model by averaging over all the classes. The mean Average Precision (*mAP*) encapsulates the system’s precision and recall over all classes in the space.

Object detection and segmentation algorithms seek to both localize and describe regions in an image. Metrics such as *AP* can be adapted to evaluate how well such systems localize a particular class. For the COCO benchmark dataset, for example, results are typically stated as *AP* as a function of Intersection over Union (*IoU*) thresholds (Lin et al. 2014; He et al. 2017). *IoU* is a pixel level computation of the ratio between the area shared correctly by the computer and the total number of pixels between the proposal and the ground truth. If the proposed region and the ground truth are perfectly overlapping, $IoU = 1$. If they are not intersecting at all, the $IoU \simeq 0$ (Rezatofighi et al. 2019).