





RESEARCH

Open Access



Plankton classification with high-throughput submersible holographic microscopy and transfer learning

Liam MacNeil^{1*} , Sergey Missan², Junliang Luo³ , Thomas Trappenberg³  and Julie LaRoche^{1*} 

Abstract

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

Results: Here we use a light and portable digital in-line holographic microscope (The HoloSea) with maximum optical resolution of 1.5 μm , intensity-based object detection through a volume, and four different pre-trained convolutional neural networks to classify > 3800 micro-mesoplankton (> 20 μm) images across 19 classes. The maximum classifier performance was quickly achieved for each convolutional neural network during training and reached F1-scores > 89%. Taking classification further, we show that off-the-shelf classifiers perform strongly across every decision threshold for ranking a majority of the plankton classes.

Conclusion: These results show compelling baselines for classifying holographic plankton images, both rare and plentiful, including several dinoflagellate and diatom groups. These results also support a broader potential for deployable holographic microscopes to sample diverse microbial eukaryotic communities, and its use for high-throughput plankton monitoring.

Keywords: Holographic microscopy, High-throughput imaging, Deep learning, Convolutional neural networks, Plankton, Classification workflow, Deployable microscope

Background

Plankton are an integral component of the global ocean. Plankton abundance and composition can be coupled to environmental conditions and yield important insights into aquatic food webs (e.g., [1, 2]). Often hugely diverse and occupying numerous trophic modes in surface ocean

*Correspondence: L.macneil@dal.ca; Julie.Laroche@dal.ca

¹ Biology Department, Dalhousie University, 1355 Oxford Street, Halifax, NS B3H 4J1, Canada

Full list of author information is available at the end of the article



© The Author(s) 2021. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

ecosystems, classifying plankton composed in a water mass is challenging, error prone, and a bottleneck of time and costs. Recent developments in imaging instruments allow biological contents to be visualized directly from bulk volumes at high image resolution, without disintegrating cell structures [3]. Imaging instruments have used a variety of optical methods including flow cytometry [4], shadowgraphs [5], holography [6], among others. Several such devices have imaged plankton size classes that collectively encompass autotrophs and heterotrophs, spanning four orders of magnitude in size from 2 μm to 10 cm [3, 7, 8]. The high sampling frequency from digital imaging also opens new ecological challenges exploring microbial eukaryotic diversity [9], alongside technical challenges to automate classification from spatial and temporally dense datasets (e.g., [10, 11]).

Digital holography is based on the diffracted light field created by interference from objects in a sample which is illuminated by a coherent light (e.g., a laser): That interference pattern is recorded by a digital sensor and composes a hologram [12]. Since their inception [13], holographic microscopes have been applied widely at micrometre scales to observe, for example, particle distributions [14], coral mucus production [15], and to differentiate cancerous pancreatic cells from healthy ones [16]. Holographic microscopes have advanced considerably with improving computational techniques for digital reconstruction and focus enhancement [17]. Digital in-line holographic microscopy (DIHM) with a point-source laser is a simple, lens-free implementation of Gabor-style holography that can capture a 3-D sample using a common path optical configuration, whereby both reference and interfered light waves copropagate and are recorded by a digital camera [18]. DIHM has several advantages for biological studies including a simple design with a larger depth of field than conventional light microscopy, allowing rapid imaging of larger volumes and 3-D numerical refocusing with no required staining of cells [19, 20]. Due to its simplicity, DIHM can easily be incorporated into various cell imaging configurations including amplitude and phase images [21] and to date, numerous studies have used holography to image marine plankton [8, 22–25]. There is increasing interest to use its advantages towards automating classification of plankton and particulates from water samples (e.g., [26–28]). A review of holographic microscopes for aquatic imaging can be found in Nayak et al. [29].

Plankton exhibit substantial morphological variation within and between major groups, are often imaged at different orientations, appear partially occluded, or damaged. Extracting features from plankton images originally relied on handcrafted feature descriptors, which are label-free and train classifiers like support vector

machines or random forest efficiently [30]. But detecting features based on predefined traits rapidly reaches its limits. Instead, deep learning algorithms have gained popularity for their state-of-the-art performance and, at least in part, because they require no domain specific knowledge or impose descriptors for pattern recognition, rather features are learned during training [31]. Deep learning involves representing features at increasing levels of abstraction and for image tasks, the most successful models have been convolutional neural networks (CNNs): a layered neural network architecture, with layers equating to depth, and where convolutions substitute as feature extractors [32]. These CNNs learn features through sequential layers connected to the local receptive field of the previous layer and the weights learned by each kernel [32]. For plankton, CNNs have improved the classification stage of automation efforts [33, 34]. But the natural imbalance in plankton datasets and frequent drifts in class distributions [35] render accuracy benchmarks for performance biased towards majority classes and poor evaluation metrics (e.g., [36, 37]).

Achieving state-of-the-art classification at scale often requires large training datasets for CNNs, but generic features can be extracted from pre-trained models and repurposed—termed transfer learning—such that CNNs have a baseline that can recognize features unspecific to any image, similar to Gabor filters or color blobs [38]. Transfer learning has achieved classification benchmarks equivalent to traditional feature descriptors (e.g., [39]). Large plankton image datasets do exist—some containing several million labelled images across hundreds of classes (e.g., [40])—but there is a current lack of easily deployable plankton imaging devices capable of rapidly sampling several litres. Holographic microscopy combined with computer vision, could bridge high throughput in-situ data production with increasingly automated classification and enumeration of major plankton groups.

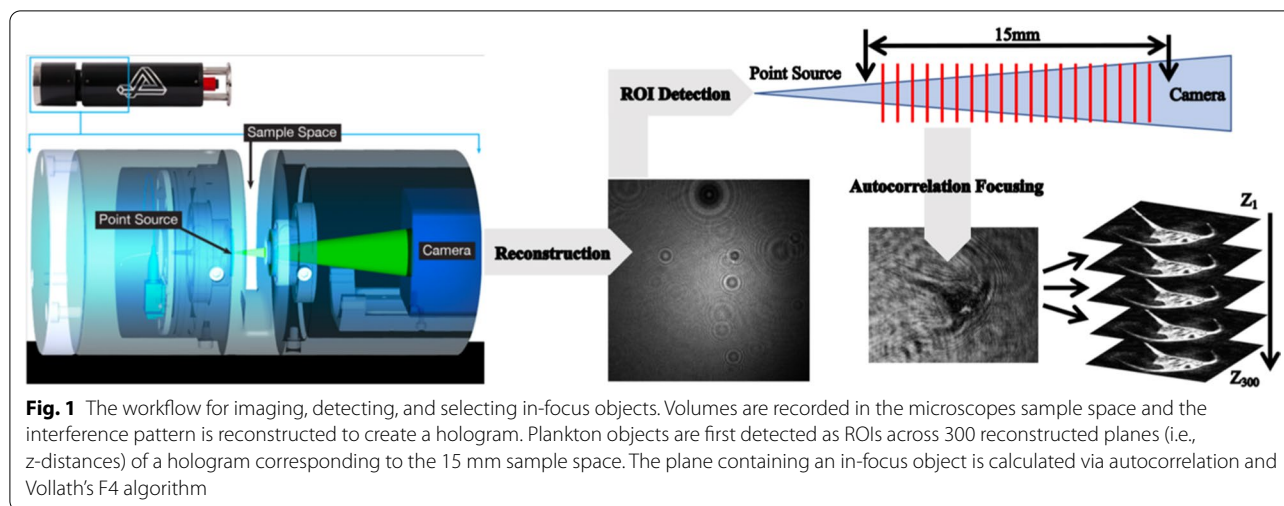
The purpose of this study is to show whether species of micro-mesoplankton can be detected in-focus from volumetric samples, classified with deep learning algorithms, and to evaluate classifiers with threshold-independent metrics—which, to date, are rarely considered for imbalanced plankton classification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

Methods

The HoloSea: submersible digital in-line holographic microscope (DIHM)

General DIHM designs for biological applications are reviewed in Garcia-Sucerquia et al. [6] and Xu et al. [20]. A similar submersible DIHM, the 4-Deep HoloSea S5¹

¹ <http://4-deep.com/>



(92 × 351 mm, 2.6 kg), first introduced by Walcutt et al. [41], was used here to image plankton cells. Its principal advantage is a simple lensless, in-flow configuration with 0.1 mL per frame and high frame rates ($>20 \text{ s}^{-1}$) that support a maximum flow rate $>130 \text{ mL min}^{-1}$. Housed in an aluminum alloy casing, the HoloSea uses a solid-state laser (405 nm) coupled to a single mode fiber optic cable acting as a point source to emit spherical light waves through a sapphire window. As light waves travel through the sampled volume, both the waves scattered by objects and reference waves copropagate until they interfere at the plane of the monochrome camera sensor (CMOS) to form an interference pattern (i.e., a 2048×2048 hologram). The camera is aligned 54 mm away from the point source and recorded holograms are stored as PNG images for further numerical reconstruction and analyses.

Numerical hologram reconstructions

Hologram reconstruction from point-source holography was first proposed by Kreuzer et al. [42], and its principles are well described [19–21]. The workflow from reconstruction to object focusing are shown in Fig. 1. In order to recover the information about objects within holograms at the specific focal distance from the point source, wave front intensity was digitally reconstructed based on a Helmholtz-Kirchhoff transformation [43] in 4-Deep Octopus software.² Each hologram was reconstructed at multiple z-distances from the point source using a 50 μm step size through the sample volume. To detect regions of interest (ROIs) in each reconstructed

plane, we used 4-Deep Stingray software³ with a globally adaptive threshold algorithm based on Otsu [44]. During the detection step, ROIs could also be discriminated based on their size, for our purposes, we defined a range of two orders of magnitude (20–2000 μm) to encompass micro-mesoplankton. Detected ROIs were clustered together across multiple z-planes based on the Euclidian distances between their centroids using the Density Based Spatial Clustering with Applications of Noise (DBSCAN) algorithm [45]. Each resultant cluster contained the same ROI tracked at multiple consecutive z-planes within the volume. To identify the plane containing an in-focus object within each cluster, we used Vollath's F4 autocorrelative algorithm [46]—the object with the highest correlation score between pixels was then stored in our database and the rest of objects within the cluster were discarded.

Holographic image dataset

The plankton for our experiments (Table 1) included monocultures grown in artificial seawater and 500 mL surface (1 m) water samples from Bedford Basin compass buoy station (44° 41' 37" N, 63° 38' 25" W). Monoculture samples were grown under f/2 nutrient replete and the recommended temperature and light conditions (Bigelow Laboratories, Maine, USA). Samples were pumped through the sample chamber using a peristaltic pump and recorded at 10 fps. The resulting image dataset was augmented by rotating each image horizontally, vertically, and translated to enlarge the number of training images, and hence the learnable features threefold [47]. All images were scaled to 128×128 pixels preserving the

² <http://4-deep.com/products/octopus-software/>

³ <http://4-deep.com/products/stingray-software/>

Table 1 Taxa identity, size ranges, and total number of images

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

Cell sizes are taken from apical cell length measurements, using 25 examples for each class

aspect ratio of the source images. Classes were randomly split approximately 50:10:40 for training, validation, and testing, respectively. Training and validation samples were divided into five stratified k-folds, where each fold retains the proportion of classes in the original training set [48]. We included a “noise” class to filter holographic artefacts [49].

Convolutional neural networks (CNNs)

The plankton detected in our holograms were classified with four different CNNs: VGG16 [50], InceptionV3 [51], ResNet50V2 [52], and Xception [53]. In terms of model depth, VGG16 is the shallowest, InceptionV3 and ResNet50 are near equal, while Xception is the deepest. Each uses convolutions as feature extractors but with different model architecture (See Additional file 1: Table S1). Due to the modest size of our plankton dataset, we used a transfer learning approach where each model was pre-trained on ~1.4 M images binned into over 1000 classes from the ImageNet dataset [54]. Pre-trained models have already learned generalizable features from the ImageNet dataset—which includes animals, sports objects, computers, and other classes very different from plankton—that provides a powerful baseline for feature recognition [38]. Classification was implemented in the Python deep learning toolbox Keras [53], which is accessible as a core component of the Tensorflow package [55].

Each model was applied in two different ways, first as a feature extractor by only retraining the deepest model layers to preserve the pre-tuned weights [38], and secondly by maintaining the first 10–20 layers and retraining the remaining layers. The second method was exploratory and involved freezing the first 10 layers in VGG16, and the first 20 layers for the other deeper models, which have presumably already learned generic features. We used dropout for each method at a probability of 0.3 to prevent overfitting [56] and added a Softmax classifier to transform the fully connected vector into a probability distribution specific to 19 classes [57]. Our images were preprocessed according to each CNNs requirements [58], and the greyscale color channel was repeated for each colored channel (i.e., RGB) that the models observed from ImageNet.

Prediction bias from our class imbalances, where the most abundant class was nearly three times greater than the least abundant, was offset by maintaining class proportions during training using stratified k-folds [59]. Combining the predictions on the validation and test sets from each fold, for each model, created an ensemble of networks to evaluate prediction variance [60]. Training was repeated for 20 epochs for each fold, where an epoch represents an entire pass of the training set. Training specifications included a batch size of 32, and momentum values of 0.9 in batch normalization layers of ResNet50, InceptionV3, and Xception [61]. The learning

algorithm minimized the log loss (cross-entropy) function through backpropagation using the Adam optimizer [62]—the learning rate was set at 0.01 and reduced by a factor of 10 if the loss function failed to improve by $1e^{-3}$ after five epochs. Holographic reconstructions, object detection and classification were implemented in the NVIDIA CUDA GPU toolkit [63] using a NVIDIA GeForce GTX960 GPU with 16 GB of RAM.

Validation measures

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

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

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

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

The balanced score between precision and recall can be represented by the F1-score, calculated using a harmonic mean (3) [64].

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

Ecologically meaningful plankton classifiers predict few false positives and a high proportion of true positives across all classes [65]. This priority favors precision, because high precision scores imply few false positives, and the F1-score as the relative balance between precision and recall, as such, high F1-score contains fewer false positives and false negatives across all labels [65]. Although both metrics are more sensitive than accuracy to the performance of minority classes, each only summarizes classifier performance at a single decision threshold: The predicted probability of an image belonging to a class

is converted to a label only when it surpasses a fixed, and often arbitrarily defined threshold [66]. To overcome this, we generated precision-recall curves at every decision threshold to visualize their trade-off—in other words, the relationship between the fraction of correctly predicted true positives (predictive positive value) and the true positive rate [67]. Precision-recall curves are robust for imbalanced classification because they are unaffected by the increasing true negatives after labels are binarized [68]. To summarize classifier performance for each class across every decision threshold, we computed the average precision of each class (4), where R_n and P_n are recall and precision at the n th threshold, respectively [48].

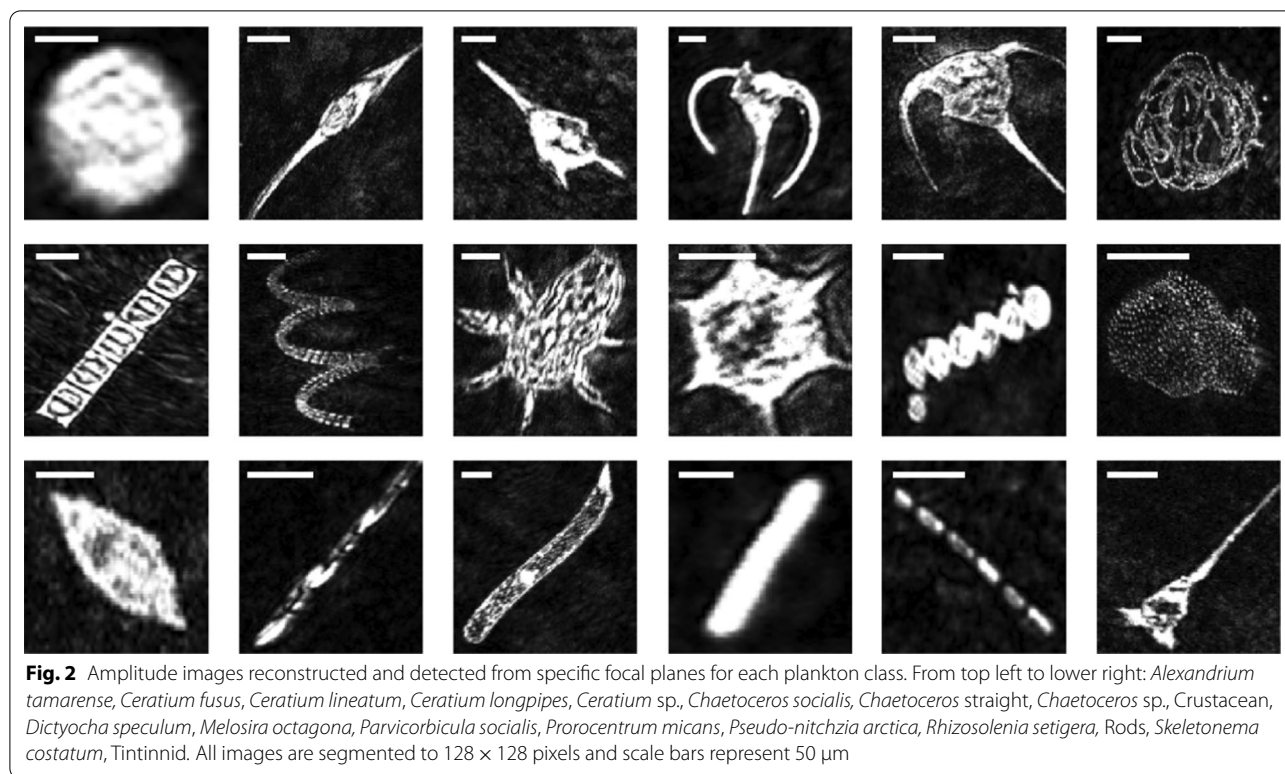
$$Average\ Precision = \sum_n (R_n - R_{n-1})P_n \quad (4)$$

Average precision is analogous to a non-linear interpolation of the area under each precision-recall curve (AUC-PR) [69]—as a rank measure, the AUC is closely related to statistical separability between classes [64, 67]. For a specific class, the performance baseline when evaluating AUC-PR is defined by the ratio of positives (P) to negatives (N) in the test set $y = \frac{P}{P+N}$, and is equal to the probability of a positive example being correctly classified over a negative example [68]. The baseline is therefore different for each class.

Results

Holographic data

In total, >17,000 holograms comprising >70 GB of data were produced from our samples. Reconstructed by Octopus software, holograms had the highest intensity in the central axis which attenuated at the hologram edges (Fig. 1). Hologram intensity was reconstructed in the order of eight milliseconds for a 2048×2048 hologram. The numerical holograms reconstruction, ROI clustering, and autofocusing that compose our multi-stage detection steps generated 3826 in-focus plankton objects from 19 classes (Fig. 2). In total, the full workflow amounted to approximately 44 h of computational time dominated by in-focus detection (>95%), and the remainder by classification. Six classes were generated from the environmental samples including *C. fusus*, *C. lineatum*, *Ceratium* sp., Crustaceans, *P. socilais*, and Tintinnids. The remaining classes derived from monoculture and represented individual plankton species. In total, the environmental classes were less abundant than classes derived from pure cultures. The size of plankton objects ranged from 20 to 640 μm , with the majority smaller than 200 μm and belonging to microplankton (Table 1). The classes proved highly imbalanced with the greatest difference between mesoplankton Crustaceans containing 13 images, and the microplankton dinoflagellate *P. micans* containing 1074



images (Table 1). After augmentation, the CNN training data contained 7215 samples which when subdivided into stratified folds contained 5772 images for training and 1443 images for validation.

Overall classification

The classification source code is publicly available on Github [70]. The feature extraction and retraining methods produced indistinguishable performance results across classification metrics, so we will consider only the feature extraction results here. For feature extraction, the overall classification performance based on the accuracy, precision, recall, and F1-score are reported in Table 2. The InceptionV3 model achieved the lowest precision values at 83% and F1-score of 81%. All the remaining three models performed comparably reaching precision scores >88%, and F1-scores >87%. The Xception model consistently outperformed every other to achieve precision and F1-scores of 89%. The underlying classification performance for each taxon is described below by their AUC-PR. Each model clearly achieved maximum precision, recall, and F1-scores quickly—in five or fewer epochs—while the mean and standard deviation for predictions across epochs was generally low (<2.5%). The log loss error showed similar model behaviour overall, with error minima in fewer than five epochs and Xception obtaining the lowest error.

Taxa-level classification

The AUC-PR values for each class are reported in Table 3. The precision-recall curves for each model (See Additional file 1: Figures S4–7) broadly showed that the highest AUC-PR values and therefore the 11 highest ranked classes included the dinoflagellates *A. tamaranse* and all four *Ceratium* taxa, along with diatoms for all three *Chaetoceros* taxa and *M. octagona*, the silicoflagelle *D. speculum.*, and our noise class. Both Xception and ResNet50 ranked the rarest class of Crustaceans highly. As the best classifier, Xception even ranked rare taxa *C. lineatum* (0.91) and Crustacean (0.86) higher than the *Chaetoceros* straight morphotype (Fig. 2), despite containing less than a seventh of examples. Classification performance deteriorated for the remaining seven taxa

Table 2 Average performance of each model across folds for each threshold metric on the test set

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

Table 3 Area under the precision-recall curves calculated using average precision for each class

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

^a Indicate rare classes with < 25 examples in the training set

to ranking only marginally better than random for the choanoflagellate *P. socialis*, and the diatoms including *P. arctica* and *S. costatum*, as well as the Rods morphotype and the ciliate Tintinnids. The dinoflagellate *P. micans* was the only taxa that was unanimously ranked worse than random in each model—that is, AUC-PR values below their class baseline—despite it containing nearly three times as many examples as the next most abundant class. No clear difference in classification performance occurred between size classes.

Discussion

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

C. lineatum and the spines in *C. longpipes* and *Ceratium* sp. resolved clearly and were conspicuous features. The dinoflagellate *P. micans* was poorly resolved and classified, it is possible that the small cell size (< 100 μm) limited any detection of its thecal plates or small (< 10 μm) apical spine [71]. For the diatoms, the chained *C. socialis*, *Chaetoceros* sp., and *M. octagona* were all distinct from each other with colonies, spirals, and straight chains that likely contributed to their reliable classification. More broadly, many chained objects showed a discernible interstitial space between cells which was especially distinct in *Chaetoceros* and *S. costatum* (Fig. 2), although the setae of the *Chaetoceros* classes was rarely visible. Comparatively, the poorly ranked diatoms *S. costatum*, *R. setigera*, and *P. arctica* all lacked morphological definition. Similar to the WHOI plankton dataset [40], the small sized choanoflagellate *P. socialis* only displayed colonies of flame bulbs and the silica loricae and flagellum cannot be seen—likely explaining its unanimously poor ranking by each CNN.

The complex morphology of plankton also presents a problem of image scale: The features available for detection in this study were limited to those that remained after objects were segmented to 128 \times 128 scale. These image sizes are different from the ImageNet images used to train each CNN—VGG16 and ResNet50 were trained on 224 \times 224 images and InceptionV3 and Xception were trained on 229 \times 229 images. This suggests encouraging transferability to our holographic plankton images. Although scaling effectively normalizes the wide variety of features and explicitly retains scale invariant features, imaged plankton features can obviously vary with size, and therefore scale invariant features only partially describe the spatial composition of any object [72]. Segmenting objects at multiple scales could capture scale-variant features, but examples of scale-variant detection are less common. Artist attribution is an example of a complex classification task where multi-scale images (256, 512, 1024, 2048 pixels) systematically improved CNN predictions using both coarse and fine grain features of digitized artworks belonging to the Rijksmuseum, at the Netherlands State Museum [73]. But currently, multi-scale CNNs lose scale invariant features that otherwise emerge during scaling and augmentation, and these features are not guaranteed to emerge during convolutional feature extraction. Further research on scale-variant feature detection could overcome this limitation and help identify the diversity of plankton features that are more or less resolvable at different scales.

Holography has certain technical challenges for capturing high-quality plankton features, owing first to the need for numerical reconstruction of a sample volume, followed by object detection and autofocusing. In

assessing the HoloSea, Walcutt et al. [41] observed two notable biases underlying particle size and density estimates, including the attenuated light intensity from the point source, both radially and axially across the sample volume and secondly, that foreground objects inevitably shade the volume background. Although this study is concerned with classification, both biases are present in this study. Several modifications offered by Walcutt et al. [41] apply here: Adjusting the point source-to-camera distance to expand sample space illumination and create a more uniform light intensity, scaling object detection probability based on pixel intensity, and local adaptive thresholding to improve ROIs detection consistency at the dimmed hologram edges—as opposed to the fixed, global thresholding algorithm used here. Because objects are less likely to be detected at the hologram edges, only a fraction of the particle field is consistently imaged. The total volume imaged, calculated as the product of the number of holograms and the volume of each hologram (maximally 0.1 mL), should be corrected by the actual illuminated proportion of the sample volume: For the HoloSea, Walcutt et al. [41] empirically derived the working image volume at 0.063 mL per hologram. The digital corrections are likely simpler and should be implemented in future quantitative assessments, unless the increasing ability of deep learning algorithms in holographic reconstruction, enhancing depth-of-field and autofocusing can outperform instrument-specific corrections [17]. Nonetheless, holography opens new opportunities for high-throughput volumetric image analysis and the robust modular casings of DIHM—which operate in the abyssopelagic zone (~6000 m) [74] and High Arctic springs [75]—make for versatile instruments to deploy in oceanic environments.

Classification tasks for almost every image domain have greatly improved with transfer learning [76], including for plankton [77]. With a transfer learning approach, our results show good classification performance for multiple groups of abundant micro and mesoplankton—encompassing the size spectra (5–50 μm) that microbial eukaryotic diversity peaks [78]. Classification performance was also high for several rare taxa including Crustaceans, *C. fusus*, *C. lineatum*, and *Ceratium* sp., all of which contained fewer than 50 training examples. Publicly shared datasets like ImageNet have been central for classification benchmarks, increasing training examples for a wider recognition of features within and across imaging modes and minimizing the imbalance of class distributions in small and large datasets [79]. For plankton, open access datasets such as the In Situ Ichthyoplankton Imaging System [5] dataset shared through the Kaggle's National Data Science Bowl competition, and the WHOI dataset captured by the Imaging Flow Cytobot

[4] are important starting points. But both image modes are quite different from holographic images: To improve transferability of feature recognition, an open database specific to the holographic domain could promote wider use and shrink the gap between its high-throughput image production and analyses. To that end, the holographic plankton images used here will be publicly available in the Cell Image library (See 'Availability of data and materials').

Although the primary concern of this work is detection and classification from holographic images, generalizing classifiers to unseen plankton populations remains challenging [35]. Plankton vary widely and are invariably observed unevenly. However rare plankton classes can be important and removing them from datasets (e.g., [34, 65]) is not desirable if imaging instruments are to be maximally effective in sampling the plankton community. Ballast water quality testing, for example, relies on presence-absence of rare, invasive taxa [80]. The proper classifier evaluation is in performance on the original imbalanced datasets, not how certain performance measures can be tuned by synthetically manipulating class balances [81]. As an alternative, optimizing decision thresholds in precision-recall curves for each class has seen revived interest, and benefits from bypassing the generated biases in common oversampling methods [82]. For evaluating classifiers of imbalanced plankton datasets, we encourage wider use of ranking metrics like AUC-PR, which summarize the trade-offs of any particular metric at every decision threshold and appear rarely used in plankton classification tasks (e.g., [83, 84]).

In machine learning, quantification is increasingly separated from classification as a different, and altogether more challenging learning task; several quantification approaches are reviewed in González et al. [85]. For in-situ plankton imaging systems, classification algorithms do not account for shifting class distributions across samples, false positive rates acquired during model training, and because most plankton studies aim to estimate total group abundance across observations in space, or through time, the learning problem then becomes at the level of the sample, not the individual image [35]. Although any classifiers false positives can be corrected for (e.g., [86]), a generalizable classifier would contain robust sample-level error, not at the taxon level [35]. The features learned by CNNs for classification, similar to those described here, can be used for plankton quantification. González et al. [87] input high-level features from pre-trained CNNs into quantification algorithms to estimate plankton prevalence throughout more than six years of the Martha's Vineyard time series collected by the Imaging Flow Cytobot and showed high correspondence—even approaching

perfect—between probabilistic quantifiers and ground-truth estimates even in rare taxa ($<1 \text{ mL}^{-1}$). These results are encouraging that even imperfect quantifiers can deliver biologically meaningful estimates of plankton.

Conclusion

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

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

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

Abbreviations

AUC: Area under the curve; AUC-PR: Area under precision-recall curve; CMOS: Complementary metal oxide semiconductor; CNN: Convolutional Neural Network; DBSCAN: Density Based Spatial Clustering with Applications of Noise; DIHM: Digital in-line holographic microscope.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12862-021-01839-0>.

Additional file 1: Figure S1. Left to right, distribution of taxa abundance for training set—where the distribution ratios are maintained during stratified cross validation—and the test set. **Figure S2.** Four classified noise objects with no resolvable features. **Figure S3.** Network architecture for basic CNN. **Figure S4.** Precision-recall curves of the InceptionV3, with iso-curves for their harmonic mean F1-score, and the area under the curve (AUC-PR). **Figure S5.** Precision-recall curves of the InceptionV3, with iso-curves for their harmonic mean F1-score, and the area under the curve (AUC-PR). **Figure S6.** Precision-recall curves of the InceptionV3, with iso-curves for their harmonic mean F1-score, and the area under the curve (AUC-PR). **Figure S7.** Precision-recall curves of the Xception model for each class, with iso-curves for their harmonic mean F1-score, and the area under the curve (AUC-PR). **Table S1.** The reference paper of four CNNs, their convolutional layers, the weighted layers that are changed during backpropagation, and broad overview of their key features. **Table S2.** Total time and memory expended for training and evaluating each model averaged for feature extraction and fine tuning. **Table S3.** Average performance of each model for each threshold metric on the test set for each fold.

Acknowledgements

The authors thank Ms. Ciara Willis for the collection of holograms from cultured phytoplankton species. LM is funded by the Alexander Graham Bell Canada Graduate Scholarship through the Natural Sciences and Engineering Research Council of Canada (NSERC). JLR is supported by a grant to Dr. Maycira Costa (PI) from Marine Environmental Observation, Prediction, and Response (MEOPAR)—partially supporting LM also—within the project: Spatiotemporal dynamics of the coastal ocean biogeochemical domains of British Columbia and Southeast Alaska—following the migration route of juvenile salmon.

Authors' contributions

LM is the main author who performed the classification experiments and drafted the manuscript. SM helped create the holographic methods including hologram reconstruction and object detection. TT and JL developed the image processing and helped with the classification approach. TT and JLR helped interpret the results. JLR, SM, and TT launched the study and contributed significantly to the manuscript. All the authors read and approved the final manuscript.

Funding

LM is funded by the Alexander Graham Bell Canada Graduate Scholarship through the Natural Sciences and Engineering Research Council of Canada (NSERC). JLR is supported by a grant to Dr. Maycira Costa (PI) from Marine Environmental Observation, Prediction, and Response (MEOPAR)—partially supporting LM also—within the project: Spatiotemporal dynamics of the coastal ocean biogeochemical domains of British Columbia and Southeast Alaska—following the migration route of juvenile salmon.

Availability of data and materials

The dataset supporting the conclusions of this article is available in the Cell Image library repository [Training data: <http://cellimagelibrary.org/groups/53406>; Test data: <http://cellimagelibrary.org/groups/53362>].

Declarations

Ethics approval and consent to participate

Not applicable.

Consent to publish

Not applicable.

Competing interests

SM declares a competing financial interest: as the chief technical officer of 4-Deep, the creators of the Octopus and Stingray software.

Author details

¹Biology Department, Dalhousie University, 1355 Oxford Street, Halifax, NS B3H 4J1, Canada. ²4Deep inwater imaging, 71 Appaloosa Run, Hammonds Plains, NS B4B 0G2, Canada. ³Department of Computer Science, Dalhousie University, 6050 University Avenue, Halifax, NS B3H 4R2, Canada.

Received: 14 January 2021 Accepted: 25 May 2021

Published online: 16 June 2021

References

- Hays G, Richardson A, Robinson C. Climate change and marine plankton. *Trends Ecol Evol*. 2005;20(6):337–44.
- Irwin AJ, Finkel ZV, Müller-Karger FE, Troccoli GL. Phytoplankton adapt to changing ocean environments. *PNAS*. 2015;112(18):5762–6. <https://doi.org/10.1073/pnas.1414752112>.
- Benfield M, Grosjean P, Culverhouse P, Irigoien X, Sieracki ME, Lopez-Urrutia A, et al. RAPID: research on automated plankton identification. *Oceanography*. 2007;20:172–87.
- Olson RJ, Sosik HM. A submersible imaging-in-flow instrument to analyze nano- and microplankton: imaging FlowCytobot: in situ imaging of nano- and microplankton. *Limnol Oceanogr Methods*. 2007;5:195–203.
- Cowen RK, Guigand CM. In situ ichthyoplankton imaging system (ISIS): system design and preliminary results: in situ ichthyoplankton imaging system. *Limnol Oceanogr Methods*. 2008;6:126–32.
- García-Sucerquia J, Xu W, Jericho SK, Klages P, Jericho MH, Kreuzer HJ. Digital in-line holographic microscopy. *Appl Opt*. 2006;45:836–50.
- Lombard F, Boss E, Waite AM, Vogt M, Uitz J, Stemmann, et al. Globally consistent quantitative observations of planktonic ecosystems. *Front Mar Sci*. 2019. <https://doi.org/10.3389/fmars.2019.00196>.
- Zetsche E, Mallahi A, Dubois F, Yourassowsky C, Kromkamp J, Meysman FJR. Imaging-in-Flow: digital holographic microscopy as a novel tool to detect and classify nanoplanktonic organisms. *Limnol Oceanogr Methods*. 2014;12:757–75.
- Colin S, Coelho LP, Sunagawa S, Bowler C, Karsenti E, Bork P, Pepperkok R, de Vargas C. Quantitative 3D-imaging for cell biology and ecology of environmental microbial eukaryotes. *Elife*. 2017;6:e26066. <https://doi.org/10.7554/eLife.26066>.
- Biard T, Stemmann L, Picheral M, Mayot N, Vandromme P, Hauss H, et al. In situ imaging reveals the biomass of giant protists in the global ocean. *Nature*. 2016;532:504–7.
- Greer AT, Cowen RK, Guigand CM, Hare JA. Fine-scale planktonic habitat partitioning at a shelf-slope front revealed by a high-resolution imaging system. *J Mar Syst*. 2015;142:111–25.
- Schnars U, Jüptner WP. Digital recording and numerical reconstruction of holograms. *Meas Sci Technol*. 2002;13:R85. <https://doi.org/10.1088/0957-0233/13/9/201>.
- Gabor D. A new microscopic principle. *Nature*. 1948;161:777–8.
- Sheng J, Malkiel E, Katz J. Digital holographic microscope for measuring three-dimensional particle distributions and motions. *Appl Opt*. 2006;45(16):3893–901.
- Zetsche EM, Baussant T, Meysman FJR, van Oevelen D. Direct visualization of mucus production by the cold-water coral *Lophelia pertusa* with digital holographic microscopy. *PLoS ONE*. 2016;11(2):e0146766. <https://doi.org/10.1371/journal.pone.0146766>.
- Kemper B, Carl D, Schneckeburger J, Bredebusch I, Schäfer M, Domschke W, von Bally G. Investigation of living pancreas tumor cells by digital holographic microscopy. *J Biomed Opt*. 2006. <https://doi.org/10.1117/1.2204609>.
- Riverson Y, Wu Y, Ozcan A. Deep learning in holography and coherent imaging. *Light Sci Appl*. 2019. <https://doi.org/10.1038/s41377-019-0196-0>.
- Kreuzer HJ, Jericho MH. Digital in-line holographic microscopy. *Imag Micro*. 2007;9:63–5.
- Jericho MH, Kreuzer HJ. Point source digital in-line holographic microscopy. In: Ferraro P, Wax A, Zalevsky Z, editors. *Coherent light microscopy*. Berlin: Springer; 2011. p. 3–30.
- Xu W, Jericho MH, Meinertzhagen IA, Kreuzer HJ. Digital in-line holography for biological applications. *Proc Natl Acad Sci USA*. 2001;98(20):11301–5.
- Jericho MH, Kreuzer HJ, Kanka M, Riesenberger R. Quantitative phase and refractive index measurements with point-source digital in-line holographic microscopy. *Appl Opt*. 2012;51(10):1503–15.
- Hobson PR, Krantz EP, Lampitt RS, Rogerson A, Watson J. A preliminary study of the distribution of plankton using hologrammetry. *Opt Laser Technol*. 1997;29(1):25–33.
- Malkiel E, Alquaddoomi O, Katz J. Measurements of plankton distribution in the ocean using submersible holography. *Meas Sci Technol*. 1999;10:1142–52.
- Rotermund LM, Samson J, Kreuzer HJ. A submersible holographic microscope for 4-D in-situ studies of micro-organisms in the ocean with intensity and quantitative phase imaging. *J Mar Sci Res Dev*. 2015. <https://doi.org/10.4172/2155-9910.1000181>.
- Sun H, Benzie PW, Burns N, Hendry DC, Player MA, Watson J. Underwater digital holography for studies of marine plankton. *Philos Trans R Soc A*. 2008;366:1789–806.
- Bianco V, Memmolo P, Carcagni P, Merola F, Paturzo M, Distanti C, Ferraro P. Microplastic identification via holographic imaging and machine learning. *Adv Intell Syst*. 2020;2(2):1900153.
- Göröcs Z, Tamamitsu M, Bianco V, Wolf P, Roy S, Shindo K, et al. A deep learning-enabled portable imaging flow cytometer for cost-effective, high-throughput, and label-free analysis of natural water samples. *Light Sci Appl*. 2018. <https://doi.org/10.1038/s41377-018-0067-0>.
- Guo B, Nyman L, Nayak AR, Milmore D, McFarland M, Twardowski MS, Sullivan JM, Yu J, Hong J. Automated plankton classification from holographic imagery with deep convolutional neural networks. *Limnol Oceanogr Methods*. 2021;19(1):21–36. <https://doi.org/10.1002/lom3.10402>.
- Nayak AR, Malkiel E, McFarland MN, Twardowski MS, Sullivan JM. A review of holography in the aquatic sciences: in situ characterization of particles, plankton, and small scale biophysical interactions. *Front Mar Sci*. 2021;7:572147. <https://doi.org/10.3389/fmars.2020.572147>.
- Gorsky G, Ohman MD, Picheral M, Gasparini S, Stemmann L, Romagnan JB, Cawood A, Pesant S, Garcia-Comas C, Prejger F. Digital zooplankton image analysis using the zooscan integrated system. *J Plankton Res*. 2010;32(3):285–303.
- LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521:436–44.
- Schmidhuber J. Deep learning in neural networks: an overview. *Neural Netw*. 2015;61:85–117.
- Dai J, Wang R, Zheng H, Ji G, Qiao X. ZooplanktoNet: deep convolutional network for zooplankton classification. *Ocean 2016—Shanghai*. 2016. <https://doi.org/10.1109/OCEANSAP.2016.7485680>.
- Luo JY, Irisson JO, Graham B, Guigand C, Sarafrax A, Mader C, Cowen RK. Automated plankton image analysis using convolutional neural networks: automated plankton image analysis using CNNs. *Limnol Oceanogr Methods*. 2018;16(12):814–27.
- González P, Álvarez E, Díez J, López-Urrutia Á, del Col JJ. Validation methods for plankton image classification systems: validation methods for plankton image classification systems. *Limnol Oceanogr Methods*. 2017;15:221–37.
- Corrêa I, Drews P, Botelho S, de Souza MS, Tavano VM. Deep learning for microalgae classification. In: *Machine learning and applications (ICMLA)*, 2017 16th IEEE international conference on machine learning and applications; 2017. p. 20–5.
- Dunker S, Boho D, Wäldchen J, Mäder P. Combining high-throughput imaging flow cytometry and deep learning for efficient species and life-cycle stage identification of phytoplankton. *BMC Ecol*. 2018;18:51. <https://doi.org/10.1186/s12898-018-0209-5>.
- Yosinski J, Clune J, Bengio Y, Lipson H. How transferable are features in deep neural networks? *arXiv preprint*. 2014; <https://arxiv.org/abs/1411.1792v1>.
- Sharif Razavian A, Azizpour H, Sullivan J, Carlsson S. CNN features off-the-shelf: an astounding baseline for recognition. *arXiv preprint*. 2014; <https://arxiv.org/abs/1403.6382v3>.

40. Orenstein EC, Beijbom O, Peacock EE, Sosik HM. WHOI-plankton-a large scale fine grained visual recognition benchmark dataset for plankton classification. arXiv preprint. 2015; <https://arxiv.org/abs/1510.00745v1>.
41. Walcutt NL, Knörlein B, Cetinić I, Ljubescić Z, Bosak S, Sgouros T, Montalbano AL, Neeley A, Menden-Deuer S, Omand MM. Assessment of holographic microscopy for quantifying marine particle size and concentration. *Limnol Oceanogr Methods*. 2020. <https://doi.org/10.1002/lom3.10379>.
42. Kreuzer HJ, Nakamura K, Wierzbicki A, Fink HW, Schmid H. Theory of the point source electron microscope. *Ultramicroscopy*. 1992;45:381–403.
43. Kanka M, Riesenberger R, Kreuzer HJ. Reconstruction of high-resolution holographic microscopic images. *Opt Lett*. 2009;34(8):1162–4.
44. Otsu N. A threshold selection method from gray-level histogram. *IEEE Trans Syst Man Cybern*. 1979;9(1):62–6.
45. Ester M, Kriegel HP, Xu X. A density-based algorithm for discovering clusters in large spatial databases with noise. *Proc Second Know Int Conf Dis Data Min*. 1996;6:226–31.
46. Vollath D. Automatic focusing by correlative methods. *J Microsc*. 1987;147(3):279–88.
47. Krizhevsky A, Sutskever I, Hinton G. ImageNet classification with deep convolutional neural networks. *Adv Neural Info Proc Sys*. 2012; Accessed 19 Aug 2019.
48. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in Python. *Mach Learn Python*. 2011;12:2825–30.
49. Garcia-Sucerquia J, Ramirez JH, Prieto DV. Improvement of the signal-to-noise ratio in digital holography. *Rev Mex Fis*. 2005;51:76–81.
50. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint. 2014; <https://arxiv.org/abs/1409.1556v6>.
51. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the inception architecture for computer vision. arXiv preprint. 2015; <https://arxiv.org/abs/1512.00567v3>.
52. He K, Zhang X, Ren S, Sun J. Identity mappings in deep residual networks. In: European conference on computer vision; 2016. p. 630–45.
53. Chollet F. Xception: deep learning with depthwise separable convolutions. arXiv preprint. 2017; <https://arxiv.org/abs/1610.02357v3>.
54. Deng J, Dong W, Socher R, Li LJ, Li K, Fei-Fei L. Imagenet: a large-scale hierarchical image database. In: 2009 IEEE conference on computer vision and pattern recognition, 2009; 248–255.
55. Abadi M, Barham P, Chen J, Chen Z, Davis A, Dean J, et al. TensorFlow: a system for large-scale machine learning. In: 12th Proceedings of the USENIX Symposium on Operating Systems Designs Implementation (OSDI). 2016; p. 21.
56. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting. *J Mach Learn Res*. 2014;15:1929–58.
57. Janocha K, Czarniecki WM. On loss functions for deep neural networks in classification. arXiv preprint. 2017; <https://arxiv.org/abs/1412.6980v9>.
58. Chollet F. Keras. 2015; <https://github.com/fchollet/keras>.
59. Yadav S, Shukla S. Analysis of k-fold cross-validation over hold-out validation on colossal datasets for quality classification. In: 2016 IEEE 6th International Conference on Advanced Computing (IACC). 2016; p. 78–83.
60. Hansen LK, Salamon P. Neural network ensembles. *IEEE Trans Pattern Anal Mach Intell*. 1990;12(10):993–1001.
61. Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. arXiv preprint. 2015; <https://arxiv.org/abs/1502.03167v3>.
62. Kingma D, Ba J. Adam: a method for stochastic optimization. arXiv preprint. 2014; <https://doi.org/10.1117/1.2204609>
63. Nickolls J, Buck I, Garland M, Skadron K. Scalable parallel programming with CUDA. *Queue. GPU Comput*. 2008;6:40–53.
64. Ferri C, Hernández-Orallo J, Modroui R. An experimental comparison of performance measures for classification. *Pattern Recognit Lett*. 2009;30(1):27–38.
65. Faillettaz R, Picheral M, Luo JY, Guigand C, Cowen RK, Irisson JO. Imperfect automatic image classification successfully describes plankton distribution patterns. *Meth Oceanogr*. 2016;15–16:60–77.
66. Tharwat A. Classification assessment methods. *Appl Comput Inf*. 2018. <https://doi.org/10.1016/j.aci.2018.08.003>.
67. Davis J, Goadrich M. The relationship between Precision-Recall and ROC curves. In: Proceedings of the 23rd International Conference on Machine Learning. 2006; 233–240.
68. Saito T, Rehmsmeier M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLoS ONE*. 2015;10(3):e0118432.
69. Boyd K, Eng KH, Page CD. Area under the Precision-Recall Curve: point estimates and confidence intervals. In: Salinesi C, Norrie MC, Pastor Ó, editors. Advanced information systems engineering, vol. 7908. Berlin, Heidelberg: Springer; 2013. p. 451–66.
70. MacNeil L. Holographic transfer learning. 2020; <https://github.com/LiamMacNeil/Holographic-Transfer-Learning>.
71. Hasle GR, Syvertsen EE. Marine diatoms. In: Tomas CR, editor. Identifying marine phytoplankton. San Diego: Academic Press; 1997. ISBN 0-12-693018-X-XV. 858pp.
72. Gluckman J. Scale variant image pyramids, In: 2006 computer vision and pattern recognition. 2006. <https://doi.org/10.1109/CVPR.2006.265>.
73. van Noord N, Postma E. Learning scale-variant and scale-invariant features for deep image classification. *Pattern Recognit*. 2017;61:583–92.
74. Bochdansky AB, Jericho MH, Herndl GJ. Development and deployment of a point-source digital inline holographic microscope for the study of plankton and particles to a depth of 6000 m: Deep-sea holographic microscopy. *Limnol Oceanogr Methods*. 2013;11:28–40.
75. Jericho SK, Klages P, Nadeau J, Dumas EM, Jericho MH, Kreuzer HJ. In-line digital holographic microscopy for terrestrial and exobiological research. *Planet Space Sci*. 2010;58(4):701–5. <https://doi.org/10.1016/j.pss.2009.07.012>.
76. Weiss K, Khoshgoftaar TM, Wang D. A survey of transfer learning. *J Big Data*. 2016. <https://doi.org/10.1186/s40537-016-0043-6>.
77. Orenstein EC, Beijbom O. Transfer learning and deep feature extraction for planktonic image data sets. *Proc IEEE Int Conf Comput Vis*. 2017. <https://doi.org/10.1109/WACV.2017.125>.
78. de Vargas C, Audic S, Henry N, Decelle J, Mahe F, Logares R, et al. Eukaryotic plankton diversity in the sunlit ocean. *Science*. 2015;348(6237):1261605–1261605. <https://doi.org/10.1126/science.1261605>.
79. Kornblith S, Shlens J, Le QV. Do better ImageNet models transfer better? arXiv preprint. 2018; <https://arxiv.org/abs/1805.08974v3>.
80. Casas-Monroy O, Linley RD, Adams JK, Chan FT, Drake DAR, Bailey SA. Relative invasion risk for plankton across marine and freshwater systems: examining efficacy of proposed international ballast water discharge standards. *PLoS ONE*. 2015;10(3):e0118267.
81. Provost F. Machine learning from imbalanced data sets 101. In: Proceedings of the AAAI-2000 Workshop on Imbalanced Data Sets. 2000.
82. Collett G, Prelec D, Patil KR. Reviving threshold-moving: a simple plug-in bagging ensemble for binary and multiclass imbalanced data. *Neuro-computing*. 2018;275:330–40.
83. Lumini A, Nanni L. Deep learning and transfer learning features for plankton classification. *Ecol Inform*. 2019;51:33–43.
84. Pastore VP, Zimmerman TG, Biswas SK, Bianco S. Annotation-free learning of plankton for classification and anomaly detection. *Sci Rep*. 2020. <https://doi.org/10.1038/s41598-020-68662-3>.
85. González P, Castaño A, Chawla NV, Coz JJD. A review on quantification learning. *ACM Comput Surv*. 2017;50(5):1–40.
86. Briseño-Avena C, Schmid MS, Swieca K, Sponaugle S, Brodeur RD, Cowen RK. Three-dimensional cross-shelf zooplankton distributions off the Central Oregon Coast during anomalous oceanographic conditions. *Prog Oceanogr*. 2020;188:102436. <https://doi.org/10.1016/j.pocan.2020.102436>.
87. González P, Castaño A, Peacock EE, Díez J, Del Coz JJ, Sosik HM. Automatic plankton quantification using deep features. *J Plankton Res*. 2019;41(4):449–63.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.