Supplement of Earth Syst. Sci. Data, 18, 945–967, 2026
https://doi.org/10.5194/essd-18-945-2026-supplement

Supplement of

Benchmark of plankton images classification: emphasizing features extraction over classifier complexity

Thelma Panaïotis et al.

Correspondence to: Thelma Panaïotis (thelma.panaiotis.pub@proton.me)

Figures

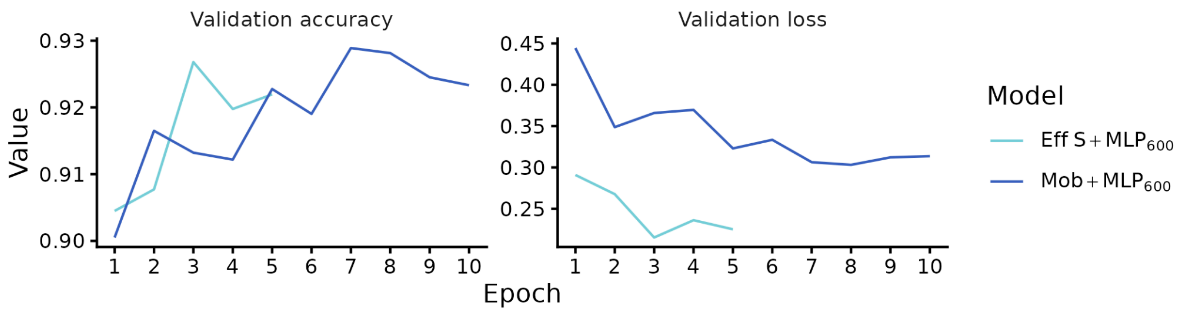

Figure S1: Illustration of training progress of the MobileNet V2 and EfficientNet S on the UVP6 dataset.

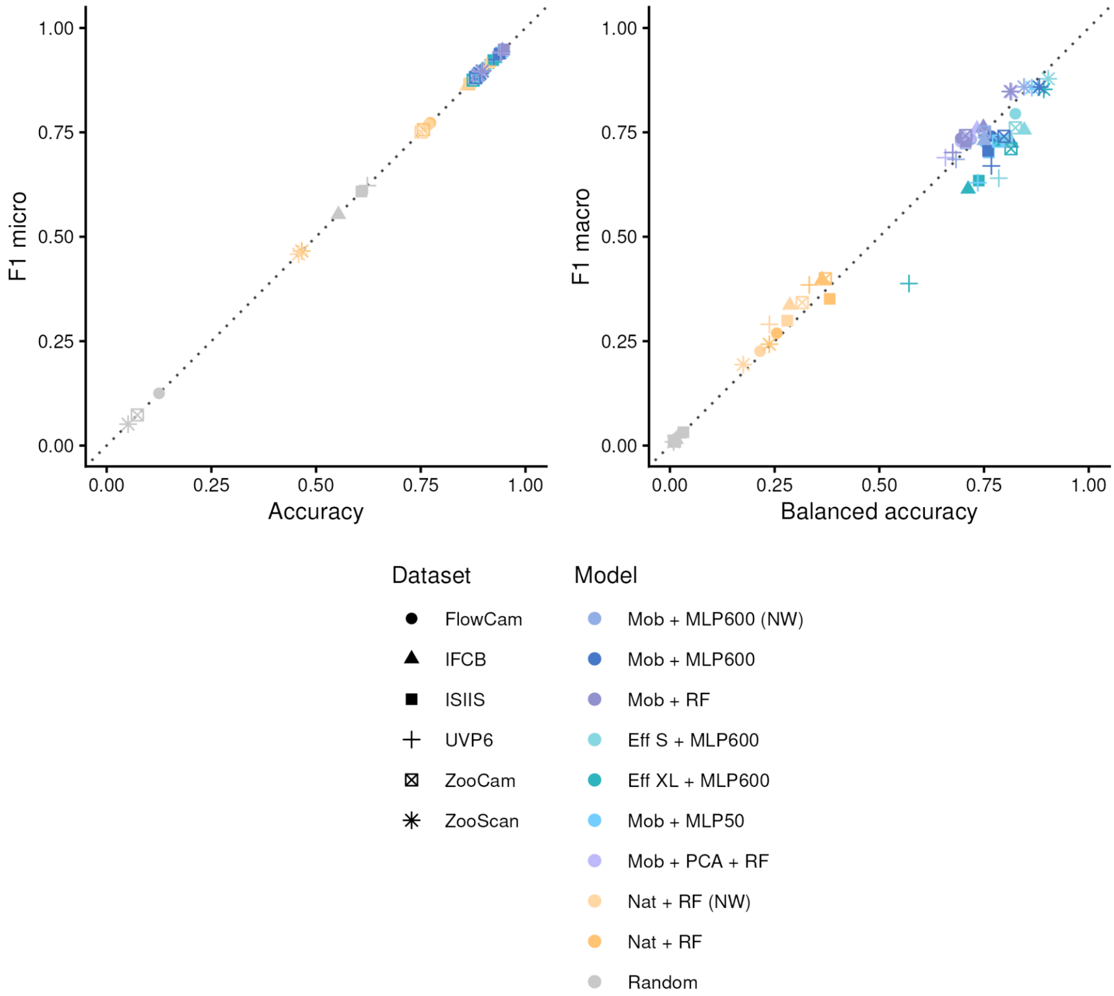

Figure S2: Relationship between F1 and accuracy-type metrics. (a) Micro-averaged F1 against accuracy and (b) macro-averaged F1 against balanced accuracy; for every model across all datasets. The dotted diagonal represents the 1:1 line, highlighting where the two metrics convey identical information.

Tables

Table S1: Reporting of per class metrics or confusion matrix (CM) in plankton image classification studies.

Reference	Classif. task	Instrument	Method	CM / per class metrics
Ten most cited papers according to Irisson et al., 2022				
Gorsky et al., 2010	Yes	Zooscan	Various	**Yes**
Sosik and Olson, 2007	Yes	IFCB	SVM	**Yes**
Benfield et al., 2007	No	-	-	-
Grosjean et al., 2004	Yes	Zooscan	Various	**No**
Luo et al., 2005	Yes	SIPPER	SVM	**No**
Blackburn et al., 1998	Yes	other	ANN	**No**
Malkiel et al., 1999	No	-	-	-
Culverhouse et al., 2006	Yes	SIPPER	ANN	**No**
Tang et al., 1998	Yes	VPR	ANN	**Yes**
Hu and Davis, 2005	Yes	VPR	SVM	**Yes**
Other plankton image classification studies cited in our paper				
Anglès et al., 2015	Yes	IFCB	SVM	**No**
Bi et al., 2015	Yes	ZOOVIS	SVM	**No**
Blaschko et al., 2005	Yes	Flowcam	Various	**No**
Callejas et al., 2025	Yes	IFCB, Zooscan	CNN / VIT	**No**
Cheng et al., 2019	Yes	ZOOVIS	CNN	**Yes**
Ciranni et al., 2025	Yes	IFCB, Zooscan	CNN-derived	**No**
Cui et al., 2018	Yes	IFCB	CNN	**No**
Dai et al., 2016	Yes	Zooscan	CNN	**No**
Dai et al., 2017	Yes	IFCB	CNN	**No**
Dieleman et al., 2016	Yes	ISIIS	CNN	**No**
Du et al., 2020	Yes	ISIIS	CNN	**No**
Eerola et al., 2024	No	-	-	-
Eftekhari et al., 2025	Yes	Plankton Imager	VIT	**Yes**
Ellen et al., 2015	Yes	Zooscan	Various	**Yes**
Ellen et al., 2019	Yes	Zooglider	CNN + Various	**Yes**
Ellen and Ohman, 2024	Yes	UVP5, Zooglider, Zooscan	CNN	**No**
Geraldes et al., 2019	Yes	ISIIS, MarinEye	CNN	**No**
Gonzalez et al., 2019	No	-	-	-
Guo et al., 2021a	Yes	ISIIS	CNN	**No**
Guo et al., 2021b	Yes	IFCB	RF	**Yes**
Guo and Guan, 2021	Yes	IFCB, ISIIS, Zooscan	CNN	**No**
Hassan et al., 2025	Yes	IFCB	CNN-derived	**Yes**

Kerr et al., 2020	Yes	Flowcam	CNN + classic	**No**
Kraft et al., 2022	Yes	IFCB	CNN	**Yes**
Kyathanahally et al., 2021	Yes	IFCB, ISIIS, Zooscan	CNN + classic	**Yes**
Kyathanahally et al., 2022	Yes	IFCB, ISIIS, Zooscan	CNN + VIT	**No**
Langeland Teigen et al., 2020	Yes	IFCB, ISIIS	CNN	**No**
Lee et al., 2016	Yes	IFCB	CNN	**Yes**
Li and Cui, 2016	Yes	ISIIS	CNN	**No**
Li et al., 2019	Yes	ISIIS	CNN	**No**
Liu et al., 2018	Yes	IFCB	CNN	**Yes**
Lumini and Nanni, 2019	Yes	IFCB, ISIIS, Zooscan	CNN	**Yes**
Luo et al., 2018	Yes	ISIIS	CNN	**Yes**
Malde and Kim, 2019	Yes	Zooscan	CNN	**Yes**
Malde et al., 2020	No	-	-	**-**
Maracani et al., 2023	Yes	IFCB, ISIIS, Zooscan	CNN + VIT	**No**
Masoudi et al., 2024	Yes	UVP6	CNN + classic	**Yes**
Orenstein et al., 2015	Yes	IFCB	CNN + RF	**No**
Orenstein and Beijbom, 2017	Yes	ISIIS, IFCB, SPC	CNN	**No**
Py et al., 2016	Yes	ISIIS	CNN	**No**
Rodrigues et al., 2018	Yes	ISIIS	CNN	**Yes**
Schmid et al., 2020	Yes	ISIIS	CNN	**Yes**
Schröder et al., 2019	Yes	UVP5	CNN	**Yes**
Uchida et al., 2018	Yes	ISIIS	CNN	**No**
Venkataramanan et al., 2021	Yes	IFCB	CNN	**No**
Yan et al., 2017	Yes	ISIIS	CNN	**No**
Zheng et al., 2017	Yes	IFCB, ISIIS, Zooscan	MKL	**Yes**
			Total yes	**22**
			Total no	**28**

Instruments acronyms: IFCB = Imaging FlowCytobot, ISIIS = In Situ Ichthyoplankton Imaging System, SIPPER = Shadowed Image Particle Profiling and Evaluation Recorder, SPC = Scripps Plankton Camera, UVP = Underwater Vision Profiler, ZOOVIS = ZOOplankton Visualization and Imaging System. Methods acronyms: ANN = Artificial Neural Network, CNN = Convolutional Neural Network, MKL = Multi Kernel Learning, RF = Random Forest, SVM = Support Vector Machine, VIT = Vision Transformer.

References from Table S1

Anglès, S., Jordi, A., and Campbell, L.: Responses of the coastal phytoplankton community to tropical cyclones revealed by high-frequency imaging flow cytometry, Limnol. Oceanogr., 60, 1562–1576, https://doi.org/10.1002/lno.10117, 2015.

Benfield, M., Grosjean, P., Culverhouse, P., Irigolen, X., Sieracki, M., Lopez-Urrutia, A., Dam, H., Hu, Q., Davis, C., Hanson, A., Pilskaln, C., Riseman, E., Schulz, H., Utgoff, P., and Gorsky, G.: RAPID: Research on Automated Plankton Identification, Oceanography, 20, 172–187, https://doi.org/10.5670/oceanog.2007.63, 2007.

Bi, H., Guo, Z., Benfield, M. C., Fan, C., Ford, M., Shahrestani, S., and Sieracki, J. M.: A Semi-Automated Image Analysis Procedure for In Situ Plankton Imaging Systems, PLOS ONE, 10, e0127121, https://doi.org/10.1371/journal.pone.0127121, 2015.

Blackburn, N., Hagström, Å., Wikner, J., Cuadros-Hansson, R., and Bjørnsen, P. K.: Rapid Determination of Bacterial Abundance, Biovolume, Morphology, and Growth by Neural Network-Based Image Analysis, Appl. Environ. Microbiol., 64, 3246–3255, https://doi.org/10.1128/AEM.64.9.3246-3255.1998, 1998.

Blaschko, M. B., Holness, G., Mattar, M. A., Lisin, D., Utgoff, P. E., Hanson, A. R., Schultz, H., Riseman, E. M., Sieracki, M. E., Balch, W. M., and Tupper, B.: Automatic In Situ Identification of Plankton, in: 2005 Seventh IEEE Workshops on Applications of Computer Vision (WACV/MOTION'05) - Volume 1, 2005 Seventh IEEE Workshops on Applications of Computer Vision (WACV/MOTION'05) - Volume 1, Citation Key: blaschko2005Automatic, 79–86, https://doi.org/10.1109/ACVMOT.2005.29, 2005.

Callejas, S., Lira, H., Berry, A., Martí, L., and Sanchez-Pi, N.: No Plankton Left Behind: Preliminary Results on Massive Plankton Image Recognition, in: High Performance Computing, Cham, 170–185, https://doi.org/10.1007/978-3-031-80084-9_12, 2025.

Cheng, K., Cheng, X., Wang, Y., Bi, H., and Benfield, M. C.: Enhanced convolutional neural network for plankton identification and enumeration, PLOS ONE, 14, e0219570, https://doi.org/10.1371/journal.pone.0219570, 2019.

Ciranni, M., Gjergji, A., Maracani, A., Murino, V., and Pastore, V. P.: In-domain self-supervised learning for plankton image classification on a budget, Proceedings of the Winter Conference on Applications of Computer Vision, 1588–1597, 2025.

Cui, J., Wei, B., Wang, C., Yu, Z., Zheng, H., Zheng, B., and Yang, H.: Texture and Shape Information Fusion of Convolutional Neural Network for Plankton Image Classification, in: 2018 OCEANS - MTS/IEEE Kobe Techno-Oceans (OTO), 2018 OCEANS - MTS/IEEE Kobe Techno-Oceans (OTO), Citation Key: cui2018Texture, 1–5, https://doi.org/10.1109/OCEANSKOBE.2018.8559156, 2018.

Culverhouse, P. F., Williams, R., Benfield, M., Flood, P. R., Sell, A. F., Mazzocchi, M. G., Buttino, I., and Sieracki, M.: Automatic image analysis of plankton: future perspectives, Mar. Ecol. Prog. Ser., 312, 297–309, https://doi.org/10.3354/meps312297, 2006.

Dai, J., Wang, R., Zheng, H., Ji, G., and Qiao, X.: ZooplanktoNet: Deep convolutional network for zooplankton classification, in: OCEANS 2016 - Shanghai, OCEANS 2016 - Shanghai, Citation Key: dai2016ZooplanktoNet, 1–6, https://doi.org/10.1109/OCEANSAP.2016.7485680, 2016.

Dai, J., Yu, Z., Zheng, H., Zheng, B., and Wang, N.: A Hybrid Convolutional Neural Network for Plankton Classification, in: Computer Vision – ACCV 2016 Workshops, Cham, Citation Key: dai2017Hybrid, 102–114, https://doi.org/10.1007/978-3-319-54526-4_8, 2017.

Dieleman, S., De Fauw, J., and Kavukcuoglu, K.: Exploiting Cyclic Symmetry in Convolutional Neural Networks, ArXiv160202660 Cs, 2016.

Du, A., Gu, Z., Yu, Z., Zheng, H., and Zheng, B.: Plankton Image Classification Using Deep Convolutional Neural Networks with Second-order Features, in: Global Oceans 2020: Singapore – U.S. Gulf Coast, Global Oceans 2020: Singapore – U.S. Gulf Coast, 1–5, https://doi.org/10.1109/IEEECONF38699.2020.9389034, 2020.

Eerola, T., Batrakhanov, D., Barazandeh, N. V., Kraft, K., Haraguchi, L., Lensu, L., Suikkanen, S., Seppälä, J., Tamminen, T., and Kälviäinen, H.: Survey of automatic plankton image recognition: challenges, existing solutions and future perspectives, Artif. Intell. Rev., 57, 114, https://doi.org/10.1007/s10462-024-10745-y, 2024.

Eftekhari, N., Pitois, S., Masoudi, M., Blackwell, R. E., Scott, J., Giering, S. L. C., and Fry, M.: Improving in Situ Real-Time Classification of Long-Tail Marine Plankton Images for Ecosystem Studies, in: Computer Vision – ECCV 2024 Workshops, Cham, 121–134, https://doi.org/10.1007/978-3-031-92387-6_8, 2025.

Ellen, J., Hongyu Li, and Ohman, M. D.: Quantifying California current plankton samples with efficient machine learning techniques, in: OCEANS 2015 - MTS/IEEE Washington, OCEANS 2015 - MTS/IEEE Washington, Citation Key: ellen2015Quantifying, 1–9, https://doi.org/10.23919/OCEANS.2015.7404607, 2015.

Ellen, J. S. and Ohman, M. D.: Beyond transfer learning: Leveraging ancillary images in automated classification of plankton, Limnol. Oceanogr. Methods, 22, 943–952, https://doi.org/10.1002/lom3.10648, 2024.

Ellen, J. S., Graff, C. A., and Ohman, M. D.: Improving plankton image classification using context metadata, Limnol. Oceanogr. Methods, 17, 439–461, https://doi.org/10.1002/lom3.10324, 2019.

Geraldes, P., Barbosa, J., Martins, A., Dias, A., Magalhães, C., Ramos, S., and Silva, E.: In situ real-time Zooplankton Detection and Classification, in: OCEANS 2019 - Marseille, OCEANS 2019 - Marseille, 1–6, https://doi.org/10.1109/OCEANSE.2019.8867552, 2019.

Gonzalez, P., Castano, A., Peacock, E. E., Diez, J., Jose Del Coz, J., and Sosik, H. M.: Automatic plankton quantification using deep features, J. Plankton Res., 41, 449–463, https://doi.org/10.1093/plankt/fbz023, 2019.

Gorsky, G., Ohman, M. D., Picheral, M., Gasparini, S., Stemmann, L., Romagnan, J.-B., Cawood, A., Pesant, S., Garcia-Comas, C., and Prejger, F.: Digital zooplankton image analysis using the ZooScan integrated system, J. Plankton Res., 32, 285–303, https://doi.org/10.1093/plankt/fbp124, 2010.

Grosjean, P., Picheral, M., Warembourg, C., and Gorsky, G.: Enumeration, measurement, and identification of net zooplankton samples using the ZOOSCAN digital imaging system, ICES J. Mar. Sci., 61, 518–525, https://doi.org/10.1016/j.icesjms.2004.03.012, 2004.

Guo, C., Wei, B., and Yu, K.: Deep Transfer Learning for Biology Cross-Domain Image Classification, J. Control Sci. Eng., 2021, 2518837, https://doi.org/10.1155/2021/2518837, 2021a.

Guo, J. and Guan, J.: Classification of Marine Plankton Based on Few-shot Learning, Arab. J. Sci. Eng., 46, 9253–9262, https://doi.org/10.1007/s13369-021-05786-2, 2021.

Guo, J., Ma, Y., and Lee, J. H. W.: Real-time automated identification of algal bloom species for fisheries management in subtropical coastal waters, J. Hydro-Environ. Res., 36, 1–32, https://doi.org/10.1016/j.jher.2021.03.002, 2021b.

Hassan, M., Salbitani, G., Carfagna, S., and Khan, J. A.: Deep learning meets marine biology: Optimized fused features and LIME-driven insights for automated plankton classification, Comput. Biol. Med., 192, 110273, https://doi.org/10.1016/j.compbiomed.2025.110273, 2025.

Hu, Q. and Davis, C.: Automatic plankton image recognition with co-occurrence matrices and Support Vector Machine, Mar. Ecol. Prog. Ser., 295, 21–31, https://doi.org/10.3354/meps295021, 2005.

Irisson, J.-O., Ayata, S.-D., Lindsay, D. J., Karp-Boss, L., and Stemmann, L.: Machine Learning for the Study of Plankton and Marine Snow from Images, Annu. Rev. Mar. Sci., 14, 277–301, https://doi.org/10.1146/annurev-marine-041921-013023, 2022.

Kerr, T., Clark, J. R., Fileman, E. S., Widdicombe, C. E., and Pugeault, N.: Collaborative Deep Learning Models to Handle Class Imbalance in FlowCam Plankton Imagery, IEEE Access, 8, 170013–170032, https://doi.org/10.1109/ACCESS.2020.3022242, 2020.

Kraft, K., Velhonoja, O., Eerola, T., Suikkanen, S., Tamminen, T., Haraguchi, L., Ylöstalo, P., Kielosto, S., Johansson, M., Lensu, L., Kälviäinen, H., Haario, H., and Seppälä, J.: Towards operational phytoplankton recognition with automated high-throughput imaging, near-real-time data processing, and convolutional neural networks, Front. Mar. Sci., 9, https://doi.org/10.3389/fmars.2022.867695, 2022.

Kyathanahally, S. P., Hardeman, T., Merz, E., Bulas, T., Reyes, M., Isles, P., Pomati, F., and Baity-Jesi, M.: Deep Learning Classification of Lake Zooplankton, Front. Microbiol., 12, 2021.

Kyathanahally, S. P., Hardeman, T., Reyes, M., Merz, E., Bulas, T., Brun, P., Pomati, F., and Baity-Jesi, M.: Ensembles of data-efficient vision transformers as a new paradigm for automated classification in ecology, Sci. Rep., 12, 18590, https://doi.org/10.1038/s41598-022-21910-0, 2022.

Langeland Teigen, A., Saad, A., and Stahl, A.: Leveraging Similarity Metrics to In-Situ Discover Planktonic Interspecies Variations or Mutations, in: Global Oceans 2020: Singapore – U.S. Gulf Coast, Global Oceans 2020: Singapore - U.S. Gulf Coast, Biloxi, MS, USA, 1–8, https://doi.org/10.1109/IEEECONF38699.2020.9388998, 2020.

Lee, H., Park, M., and Kim, J.: Plankton classification on imbalanced large scale database via convolutional neural networks with transfer learning, in: 2016 IEEE International Conference on Image Processing (ICIP), 2016 IEEE International Conference on Image Processing (ICIP), Citation Key: lee2016Plankton, 3713–3717, https://doi.org/10.1109/ICIP.2016.7533053, 2016.

Li, X. and Cui, Z.: Deep residual networks for plankton classification, in: OCEANS 2016 MTS/IEEE Monterey, OCEANS 2016 MTS/IEEE Monterey, Citation Key: li2016Deep, 1–4, https://doi.org/10.1109/OCEANS.2016.7761223, 2016.

Li, X., Long, R., Yan, J., Jin, K., and Lee, J.: TANet: A Tiny Plankton Classification Network for Mobile Devices, Mob. Inf. Syst., 2019, 6536925, https://doi.org/10.1155/2019/6536925, 2019.

Liu, J., Du, A., Wang, C., Yu, Z., Zheng, H., Zheng, B., and Zhang, H.: Deep Pyramidal Residual Networks for Plankton Image Classification, in: 2018 OCEANS - MTS/IEEE Kobe Techno-Oceans (OTO), 2018 OCEANS - MTS/IEEE Kobe Techno-Oceans (OTO), 1–5, https://doi.org/10.1109/OCEANSKOBE.2018.8559106, 2018.

Lumini, A. and Nanni, L.: Deep learning and transfer learning features for plankton classification, Ecol. Inform., 51, 33–43, https://doi.org/10.1016/j.ecoinf.2019.02.007, 2019.

Luo, J. Y., Irisson, J.-O., Graham, B., Guigand, C., Sarafraz, A., Mader, C., and Cowen, R. K.: Automated plankton image analysis using convolutional neural networks, Limnol. Oceanogr. Methods, 16, 814–827, https://doi.org/10.1002/lom3.10285, 2018.

Luo, T., Kramer, K., Goldgof, D. B., Hall, L. O., Samson, S., Remsen, A., and Hopkins, T.: Active learning to recognize multiple types of plankton, J. Mach. Learn. Res., 6, 589–613, 2005.

Malde, K. and Kim, H.: Beyond image classification: zooplankton identification with deep vector space embeddings, ArXiv190911380 Cs, 2019.

Malde, K., Handegard, N. O., Eikvil, L., and Salberg, A.-B.: Machine intelligence and the data-driven future of marine science, ICES J. Mar. Sci., 77, 1274–1285, https://doi.org/10.1093/icesjms/fsz057, 2020.

Malkiel, E., Alquaddoomi, O., and Katz, J.: Measurements of plankton distribution in the ocean using submersible holography, Meas. Sci. Technol., 10, 1142, https://doi.org/10.1088/0957-0233/10/12/305, 1999.

Maracani, A., Pastore, V. P., Natale, L., Rosasco, L., and Odone, F.: In-domain versus out-of-domain transfer learning in plankton image classification, Sci. Rep., 13, 10443, https://doi.org/10.1038/s41598-023-37627-7, 2023.

Masoudi, M., Giering, S. L. C., Eftekhari, N., Massot-Campos, M., Irisson, J.-O., and Thornton, B.: Optimizing Plankton Image Classification With Metadata-Enhanced Representation Learning, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., 17, 17117–17133, https://doi.org/10.1109/JSTARS.2024.3424498, 2024.

Orenstein, E. C. and Beijbom, O.: Transfer Learning and Deep Feature Extraction for Planktonic Image Data Sets, in: 2017 IEEE Winter Conference on Applications of Computer Vision (WACV), 2017 IEEE Winter Conference on Applications of Computer Vision (WACV), Citation Key: orenstein2017Transfer, 1082–1088, https://doi.org/10.1109/WACV.2017.125, 2017.

Orenstein, E. C., Beijbom, O., Peacock, E. E., and Sosik, H. M.: WHOI-Plankton- A Large Scale Fine Grained Visual Recognition Benchmark Dataset for Plankton Classification, ArXiv151000745 Cs, 2015.

Py, O., Hong, H., and Zhongzhi, S.: Plankton classification with deep convolutional neural networks, in: 2016 IEEE Information Technology, Networking, Electronic and Automation Control Conference, 2016 IEEE Information Technology, Networking, Electronic and Automation Control Conference, Citation Key: py2016Plankton, 132–136, https://doi.org/10.1109/ITNEC.2016.7560334, 2016.

Rodrigues, F. C. M., Hirata, N. S., Abello, A. A., Leandro, T., La Cruz, D., Lopes, R. M., and Hirata Jr, R.: Evaluation of Transfer Learning Scenarios in Plankton Image Classification., in: VISIGRAPP (5: VISAPP), Citation Key: rodrigues2018Evaluation, 359–366, 2018.

Schmid, M. S., Cowen, R. K., Robinson, K., Luo, J. Y., Briseño-Avena, C., and Sponaugle, S.: Prey and predator overlap at the edge of a mesoscale eddy: fine-scale, in-situ distributions to inform our understanding of oceanographic processes, Sci. Rep., 10, 1–16, https://doi.org/10.1038/s41598-020-57879-x, 2020.

Schröder, S.-M., Kiko, R., Irisson, J.-O., and Koch, R.: Low-Shot Learning of Plankton Categories, in: Pattern Recognition, Cham, Citation Key: schroder2019LowShot, 391–404, https://doi.org/10.1007/978-3-030-12939-2_27, 2019.

Sosik, H. M. and Olson, R. J.: Automated taxonomic classification of phytoplankton sampled with imaging-in-flow cytometry, Limnol. Oceanogr. Methods, 5, 204–216, https://doi.org/10.4319/lom.2007.5.204, 2007.

Tang, X., Stewart, W. K., Huang, H., Gallager, S. M., Davis, C. S., Vincent, L., and Marra, M.: Automatic Plankton

Image Recognition, Artif. Intell. Rev., 12, 177–199, https://doi.org/10.1023/A:1006517211724, 1998.

Uchida, K., Tanaka, M., and Okutomi, M.: Coupled convolution layer for convolutional neural network, Neural Netw., 105, 197–205, https://doi.org/10.1016/j.neunet.2018.05.002, 2018.

Venkataramanan, A., Laviale, M., Figus, C., Usseglio-Polatera, P., and Pradalier, C.: Tackling Inter-class Similarity and Intra-class Variance for Microscopic Image-Based Classification, in: Computer Vision Systems, Cham, 93–103, https://doi.org/10.1007/978-3-030-87156-7_8, 2021.

Yan, J., Li, X., and Cui, Z.: A More Efficient CNN Architecture for Plankton Classification, in: Computer Vision, Singapore, Citation Key: yan2017More, 198–208, https://doi.org/10.1007/978-981-10-7305-2_18, 2017.

Zheng, H., Wang, R., Yu, Z., Wang, N., Gu, Z., and Zheng, B.: Automatic plankton image classification combining multiple view features via multiple kernel learning, BMC Bioinformatics, 18, 570, https://doi.org/10.1186/s12859-017-1954-8, 2017.

Table S2: Classification report for detailed classes in the FlowCAM dataset. Reported values are F1-scores. The models are described in Figure 1. N test indicates the number of objects in the test set for each class. Colors reflect the values inside cells.

Taxon	Grouped	N test	Nat + RF	Mob + MLP600	Eff S + MLP600	Mob + PCA + RF
			Plankton classes			
Appendicularia	Appendicularia	19	0	84,8	76,6	88,4
Asterolamprales	Asterolamprales	23	8,3	50	55	62,5
Bacillariophyceae	Bacillariophyceae	791	38,1	78,5	75,8	81,7
cyano a	Bacteria	1096	67,7	85,1	84,9	86,9
cyano b	Bacteria	5433	88,8	95,2	94,7	95,4
Katagnymene spiralis	Bacteria	16	0	72	71,8	77,8
Richelia	Bacteria	79	22,9	71,5	77	77,8
UCYNA like	Bacteria	111	0	78	77,6	72,7
Bacteriastrum	Bacteriastrum	76	27,5	61,7	68,7	74,3
Chaetoceros	Chaetoceros	366	41,7	79,9	79,9	82,8
Chaetoceros peruvianus	Chaetoceros	19	0	42,9	66,7	70,6
Ciliophora	Ciliophora	17	0	42,9	38,9	48,8
Climacodium	Climacodium	81	22	77,1	72,5	81
Climacodium inter. Crocosphaera	Climacodium inter. Crocosphaera	52	20	68	68,8	70,8
Calanoida	Copepoda	43	44,3	72,5	75	84,4
Copepoda	Copepoda	35	5,6	41,4	29,6	55,9
Oithonidae	Copepoda	40	36,4	61,7	59,6	74,7
Oncaeidae	Copepoda	58	53,3	63,8	57,8	70,3
Coscinodiscaceae	Coscinodiscaceae	617	62,6	79,7	78,8	83,9
Cyttarocylididae	Cyttarocylididae	48	40,5	89,6	86,9	89,6
chainlarge	diatoma_chainlarge	64	19,7	85	83,8	85,9
chainthin	diatoma_chainthin	263	23,7	83	81,9	85,1
Codonaria	Dictyocystidae	13	0	66,7	72	72
Dictyocysta	Dictyocystidae	35	32,6	87,5	83,1	91,4
Dinophyceae	Dinophyceae	410	30,4	68,1	66,4	71,2
Dinophysis	Dinophysiales	23	0	50	45,8	57,1
Ornithocercus	Dinophysiales	36	5,4	81,7	83,3	88
Ditylum	Ditylum	123	31,4	92	92,7	93,9
part<Ditylum	Ditylum	183	9,3	84	84,5	90
Fragilariopsis	Fragilariopsis	13	0	76,2	88,9	85,7
Ceratocorys	Gonyaulacales	14	0	66,7	55,2	64
Cladopyxis	Gonyaulacales	13	0	72,7	80	96,3
Neoceratium	Gonyaulacales	194	33,4	88,6	90,7	93,8
Neoceratium furca	Gonyaulacales	59	18,2	86,4	82,9	93,1
Neoceratium fusus	Gonyaulacales	94	0	84,7	86,2	91,2
Neoceratium pentagonum	Gonyaulacales	394	47	92,6	93,4	95,1
Pyrocystaceae	Gonyaulacales	12	25	0	45,5	47,6
Pyrophacus	Gonyaulacales	20	38,7	78,9	72,7	87,2
Gymnodiniales	Gymnodiniales	32	0	5,6	36,4	50
Hemiaulus	Hemiaulus	32	0	61,8	67,7	70,9
Melosiraceae	Melosiraceae	760	86,7	96,1	95,8	97,4
Metacylis	Metacylididae	13	0	66,7	60	64,7
Nitzschia	Nitzschia	26	13,8	52,6	60,9	67,8
Odontella	Odontella	172	58,5	95,3	96,2	95,7
part<Odontella	Odontella	205	32,5	90,1	90,4	93,3

nauplii	other_Crustacea	803	58,2	94,7	95,3	96,9
part<Crustacea	other_Crustacea	433	37	75,6	76,5	82,6
egg<other	other_egg	25	14,3	69,6	62,5	76,9
part<Ciliophora	part_Ciliophora	23	0	79,1	66,7	83,3
pennate	pennate	142	18,5	79,4	83	91
Oxytoxum	Peridiniales	18	0	41,4	53,3	68,6
Podolampas	Peridiniales	47	8	71,9	69,4	76,3
Protoperidinium	Peridiniales	97	18,4	64,3	61,8	66,1
Planktoniella sol	Planktoniella sol	182	43,5	93	90,8	94
Pseudo-Nitzschia chain	Pseudo-Nitzschia chain	50	11,3	65,2	60,5	71,4
Acantharia	Retaria	12	0	44,4	61,5	69
Foraminifera	Retaria	35	0	72,5	67,5	72,5
Nassellaria	Retaria	78	0	69	76,5	77,9
Retaria	Retaria	205	36,6	84,6	85,2	85,8
Spumellaria	Retaria	35	19,5	56,1	62,3	69,6
Amphorides	Rhabdonellidae	20	0	78,8	76,9	87,2
Rhabdonellidae	Rhabdonellidae	164	60,3	89,1	89,5	93,2
Dactyliosolen	Rhizosoleniaceae	13	0	26,7	53,3	57,1
Guinardia	Rhizosoleniaceae	49	31,9	69,2	70	81,4
Rhizosolenia inter. Richelia	Rhizosoleniaceae	121	0	53	51,7	59,1
Rhizosoleniaceae	Rhizosoleniaceae	829	40,1	78,5	76,7	85,5
Thalassionematales	Thalassionematales	131	31,3	90,3	89,8	92,3
Codonellopsis	Tintinnidiidae	181	61,5	95,1	96,1	96,7
Eutintinnus	Tintinnidiidae	103	30,6	93,7	92,4	95,2
Poroecus	Tintinnidiidae	56	0	81,6	76	83,6
Salpingella	Tintinnidiidae	24	0	69,8	79,2	89,4
Steenstrupiella	Tintinnidiidae	61	0	84,3	87,1	86,8
Tintinnina	Tintinnina	71	21,4	52,5	50,7	59,1
Undellidae	Undellidae	103	26,2	86,1	87	92,9
Xystonellidae	Xystonellidae	54	7,1	64,8	69,5	80
	average		**22,2**	**71,8**	**73,5**	**79,7**
		Non plankton classses				
artefact	artefact	427	51,3	94,2	93,4	93,8
badfocus	artefact	65	11,6	31,8	43,5	36,9
bubble	artefact	53	53,7	91,4	89,1	89,8
darkrods	darkrods	127	52,2	68,8	62,3	65
darksphere	darksphere	256	68,7	84,9	80,8	81,1
dark	detritus	10363	94,9	96,8	96,6	97,1
detritus	detritus	10156	83,5	91,1	90,9	90,6
fiber	detritus	1331	92,5	95,4	94,6	95,7
light	detritus	2740	74,2	85,5	85,4	84,9
lightrods	lightrods	548	60,1	79,5	80,2	79,3
lightsphere	lightsphere	1971	87	93,8	92,9	93,4
ball_bearing_like	other_living	111	17,6	58,7	61,5	55,2
contrasted_blob	other_living	437	62,4	73	69,2	70,3
crumple sphere	other_living	87	15,4	76,1	70,6	77,3
dinophyceae_shape	other_living	40	0	67,5	68,4	77,6
other<living	other_living	89	12,5	49,3	57,7	67,9
polar_view<Dinophyceae	other_living	30	0	69,1	63,6	73
transparent_u	other_living	25	0	76,6	77,8	80,7
	average		**46,5**	**76,9**	**76,6**	**78,3**

Table S3: Classification report for detailed classes in the IFCB dataset. Reported values are F1-scores. The models are described in Figure 1. N test indicates the number of objects in the test set for each class. Colors reflect the values inside cells.

Taxon	Grouped	N test	Nat + RF	Mob + MLP600	Eff S + MLP600	Mob + PCA + RF
		Plankton classes				
Asterionellopsis	Asterionellopsis	153	32,1	87,8	84,2	84,7
Cerataulina	Cerataulina	1515	54	80,2	78	81,9
Chaetoceros	Chaetoceros	3416	33,3	80,4	78,7	77,6
Chaetoceros_didymus	Chaetoceros	57	0	67,9	60,6	65,7
Chaetoceros_didymus_flagellate	Chaetoceros	190	70,7	94,5	91,1	94,9
Chaetoceros_other	Chaetoceros	21	0	16	14,6	25
Chaetoceros_pennate	Chaetoceros	71	0	40,8	48,5	48,5
Corethron	Corethron	388	53,1	97,3	96,8	97,1
Cylindrotheca	Cylindrotheca	1515	63,2	95,9	95,4	95,4
DactFragCerataul	DactFragCerataul	335	36,5	62,7	58,8	61,7
Dactyliosolen	Dactyliosolen	856	70,7	95,4	93,6	94,3
Dictyocha	Dictyochophyceae	116	61,2	92,4	93,1	91,7
Pseudochattonella_farcimen	Dictyochophyceae	30	40,8	73,8	60,9	60,2
Dinobryon	Dinobryon	457	61	95,9	95,1	96,8
Ditylum	Ditylum	137	47,7	90,9	88,7	89,5
Eucampia	Eucampia	227	46,9	84,2	83,6	87,1
Euglena	Euglena	49	45,1	77,8	58,1	68,8
amoeba	Flagellates	27	0	51,2	44,2	47,8
flagellate_sp3	Flagellates	255	19,4	74,9	73,2	70,4
kiteflagellates	Flagellates	43	20,5	90,7	84,8	88,9
G_delicatula_detritus	Guinardia	58	3,3	66,7	54,3	62,5
G_delicatula_external_parasite	Guinardia	26	0	0	20	21,3
G_delicatula_parasite	Guinardia	195	30,6	67,2	61,5	61,8
Guinardia_delicatula	Guinardia	2289	74,4	92	89,6	89,5
Guinardia_flaccida	Guinardia	67	73,3	92,2	85,3	83,7
Guinardia_striata	Guinardia	167	44,4	85,8	82,1	86,7
Chrysochromulina	Haptophyta	34	0	58	43,9	53,7
Emiliania_huxleyi	Haptophyta	22	46,7	88,4	91,7	89,8
Phaeocystis	Haptophyta	122	64,8	94,5	93,7	90,2
Leptocylindrus	Leptocylindrus	6201	71	86,8	84,9	86,3
Leptocylindrus_mediterraneus	Leptocylindrus	16	11,8	82,4	66,7	82,4
Ciliate_mix	other_Ciliophora	744	49	88,8	84,8	87,5
Mesodinium_sp	other_Ciliophora	111	21,6	82,3	81,7	84,1
Tintinnid	other_Ciliophora	64	53,4	82	80,9	88,9
Tontonia	other_Ciliophora	34	66,7	82,4	73,8	64,2
Cerataulina_flagellate	other_diatom	14	0	47,6	62,5	63,2
Coscinodiscus	other_diatom	39	73,2	79,5	76,6	81
Delphineis	other_diatom	42	0	67,4	56,6	61,9
diatom_flagellate	other_diatom	56	0	35	37,1	40,4
Ephemera	other_diatom	25	56,5	94,1	88	90,9
other_diatom	other_diatom	14	0	35,3	51,3	71,4
Paralia	other_diatom	22	29,6	77,8	85	89,4
pennate	other_diatom	307	19,2	75,9	72,2	76,2

pennate_morphotype1	other_diatom	35	82,4	94,3	89,2	94,3
pennates_on_diatoms	other_diatom	65	0	40,4	45,8	50,7
Pleurosigma	other_diatom	75	66,7	92	91,2	92,6
Amphidinium_sp	other_Dinoflagellates	25	20	63,4	46,4	35,6
Ceratium	other_Dinoflagellates	31	63,2	96,8	96,9	95,4
dino30	other_Dinoflagellates	3563	61,4	76,3	72,5	74,9
Gyrodinium	other_Dinoflagellates	47	55	77,6	65	82,4
Katodinium_or_Torodinium	other_Dinoflagellates	28	28,6	63,6	62,2	63,9
other_dino	other_Dinoflagellates	21	30,8	63,4	47,8	56,7
Proterythropsis_sp	other_Dinoflagellates	41	13,3	66,7	55,3	62,7
Peridiniales	Peridiniales	120	59,8	81,4	76,6	77,4
Prorocentrum	Prorocentrum	188	68	89,7	87	91,5
Pseudonitzschia	Pseudonitzschia	364	16,6	67,4	64,6	70,7
Pyramimonas_longicauda	Pyramimonas_longicauda	51	24,6	90,9	85,2	96
Rhizosolenia	Rhizosolenia	2571	73	94,4	93,5	92,5
Strombidium	Strombidium	64	48,7	66,7	58,7	64,7
Thalassionema	Thalassionema	117	46,3	89,2	88,9	88
Skeletonema	Thalassiosirales	1171	38,6	85	80,9	84,1
Thalassiosira	Thalassiosirales	861	47	80,8	78,3	80
Thalassiosira_dirty	Thalassiosirales	158	45,4	69,2	61,5	64,2
	average		**38,7**	**75,6**	**72,3**	**75,4**
		Non plankton classses				
bad	bad	919	96,5	98,4	96,8	97,3
detritus	detritus	26442	59,7	87,8	87,1	87,8
other_interaction	other_interaction	248	11,7	38,3	37,1	43,2
other_living	other_living	176625	94,8	97,8	97,5	97,6
other_living_elongated	other_living_elongated	4481	44,9	72,1	69,5	70,7
spore	spore	28	0	76,6	55,4	63,2
	average		**51,3**	**78,5**	**73,9**	**76,6**

Table S4: Classification report for detailed classes in the ISIIS dataset. Reported values are F1-scores. The models are described in Figure 1. N test indicates the number of objects in the test set for each class. Colors reflect the values inside cells.

Taxon	Grouped	N test	Nat + RF	Mob + MLP600	Eff S + MLP600	Mob + PCA + RF
		Plankton classes				
Acantharea	Acantharea	317	40,9	75,9	76,5	79,5
Actinopterygii	Actinopterygii	42	28,9	59,6	55,7	62,5
Annelida	Annelida	20	0	64,5	58,8	62,2
Appendicularia	Appendicularia	63	28,4	62,5	54,7	58,8
body<Appendicularia	Appendicularia	490	10,4	78,2	76,5	79,8
house	Appendicularia	14	11,4	14,3	11,1	20,4
like<body<Appendicularia	Appendicularia	126	0	35,3	39,9	43,8
Aulacanthidae	Aulacanthidae	293	63,8	84,1	77,7	82,6
Bacillariophyceae	Bacillariophyceae	279	44,8	80,5	71,4	75,8
Chaetognatha	Chaetognatha	37	12,2	56	47,6	54,2
Cnidaria	Cnidaria	20	0	42,9	37,5	51,9
colonial<Collodaria	colonial_colodaria	44	47,1	76,9	66,7	80,9
Copepoda	Copepoda	1811	34,7	86	85,6	86,1
Harpacticoida	Copepoda	97	9,4	87,7	80,2	78,8
like<Copepoda	Copepoda	2880	45,6	82,6	82,3	81,5
Crustacea	Crustacea	148	31,4	65,5	65,8	68,7
Ctenophora	Ctenophora	205	51,9	81,7	79,7	78,6
Doliolida	Doliolida	1062	72,8	94,9	94,6	95,8
ephyra<Scyphozoa	ephyra	73	48	83,6	80,3	83,6
Eumalacostraca	Eumalacostraca	15	16	56	71,4	61,1
like<Acantharea	like_Acantharea	14	0	66,7	72	74,1
Mollusca	Mollusca	28	6,7	69,2	58,3	56,4
part<Cnidaria	part_Cnidaria	193	94,4	95,8	94,4	96,7
Pyrocystis	Pyrocystis	386	36,1	92,5	92,6	93
Rhizaria	Rhizaria	310	10,9	64	63,7	65
Rhopalonematidae	Rhopalonematidae	64	38,1	79,7	80	78,4
Siphonophorae	Siphonophorae	255	28,3	80,7	80,4	82,3
solitaryblack	solitaryblack	69	25,5	85,9	82,4	77,4
average			**29,9**	**71,5**	**69,2**	**71,8**
		Non plankton classses				
detritus	detritus	48201	94,1	97,7	97,2	97,4
other<living	other_living	149	0	33	40,4	39,1
streak	streak	3526	92,3	88,6	86,6	87,2
vertical line	vertical line	11	100	100	95,7	100
average			**71,6**	**79,8**	**79,9**	**80,9**

Table S5: Classification report for detailed classes in the UVP6 dataset. Reported values are F1-scores. The models are described in Figure 1. N test indicates the number of objects in the test set for each class. Colors reflect the values inside cells.

Taxon	Grouped	N test	Nat + RF	Mob + MLP600	Eff S + MLP600	Mob + PCA + RF
			Plankton classes			
fuzzy	Acantharia	28	43,2	90,6	88,5	79,4
spiky<Acantharia	Acantharia	31	6,1	88,9	95,4	69
star<Acantharia	Acantharia	62	61,3	98,4	99,2	96,1
Actinopterygii	Actinopterygii	17	48,3	47,4	48,8	52,4
Amphipoda	Amphipoda	28	22,7	46,5	56,6	56,2
Hyperiidea	Amphipoda	14	0	40	38,7	32,7
Annelida	Annelida	38	13,6	28,1	32,9	36,4
Phyllodocida	Annelida	20	46,7	78,9	63,3	66,7
Swima	Annelida	36	18,6	71,9	78,5	68,1
Appendicularia	Appendicularia	22	66,7	85,7	78,3	73,1
dead<house	Appendicularia	137	41,3	69	59,7	52,6
Aulacanthidae	Aulacanthidae	45	59,7	82,6	79,6	78,5
Aulatractus	Aulosphaeridae	17	11,1	75	59,6	80
Aulosphaeridae	Aulosphaeridae	55	33,3	86,8	83,5	75,4
Chaetognatha	Chaetognatha	86	55,1	72,2	65,8	60,4
other<Cnidaria	Cnidaria	62	11,8	40	45,6	39,7
small<Cnidaria	Cnidaria	140	55,7	79,7	74	74,5
Coelodendridae	Coelodendridae	32	27	74,6	71,4	73,5
spiky<Coelodendridae	Coelodendridae	17	45,5	83,9	84,8	72,2
cloud	Collodaria	37	10,3	82,2	69,6	68
Collodaria	Collodaria	99	49,3	87,8	79,8	75,8
solitaryblack	Collodaria	23	13,3	48,8	48,5	44,7
solitaryglobule	Collodaria	38	82,7	93,5	92,5	92,7
Calanidae	Copepoda	132	58,4	80	78,8	78,5
Calanoida	Copepoda	2026	63,6	83,5	81,9	82,5
Copepoda	Copepoda	918	8,2	44,6	44,8	45,8
copepoda_eggs	Copepoda	30	0	44,4	44,2	46,5
like<Copepoda	Copepoda	408	0	14,4	26,3	32,5
Ctenophora	Ctenophora	31	6,2	41,7	46,6	52,6
Echinodermata	Echinodermata	16	0	64	63,2	51
Eumalacostraca	Eumalacostraca	197	44,7	72,5	70,2	63,9
Foraminifera	Foraminifera	29	52,4	84,2	79,4	83,3
Creseis acicula	Mollusca	23	69,1	88,4	79,3	54,1
Thecosomata	Mollusca	28	17,1	65,4	68,9	55,9
Ostracoda	Ostracoda	82	17,6	73,5	64,7	71,6
like<Rhizaria	Rhizaria	82	6,5	24,4	36,7	28,3
Rhizaria	Rhizaria	331	67,2	76,6	75,6	72,4
chain<Salpida	Salpida	23	37,5	72,7	73,1	59,7
Salpida	Salpida	32	60,6	72,5	73,2	55,1
Siphonophorae	Siphonophorae	19	25	38,5	43,8	23,5
Narcomedusae	Trachylina	17	35,7	64,9	64,9	70
Trachymedusae	Trachylina	14	11,1	30	34,8	40,9
puff	Trichodesmium	1115	61,6	90,6	89,6	85,5
tuff	Trichodesmium	730	32,7	84,1	82,5	81,8
	average		**34,1**	**67,3**	**66,3**	**62,6**

Non plankton classses

artefact	artefact	4521	92,1	95,2	93,2	93,2
crystal	crystal	566	90,6	89,6	87,5	83,4
detritus	detritus	76323	95,2	97,2	96,4	96,4
fiber	fiber	4889	78,6	82,5	80,6	80,5
filament	filament	496	56,5	68,9	61,4	59,1
darksphere	other<living	52	20	61,4	36,4	53,8
other<living	other<living	238	0,8	7,5	18,4	15,2
reflection	reflection	689	69,1	78,7	71,4	77,8
t004	t004	38	0	84,6	70,5	70,7
t011	t011	17	75,9	93,8	82,9	73,9
	average		**57,9**	**75,9**	**69,9**	**70,4**

Table S6: Classification report for detailed classes in the ZooCAM dataset. Reported values are F1-scores. The models are described in Figure 1. N test indicates the number of objects in the test set for each class. Colors reflect the values inside cells.

Taxon	Grouped	N test	Nat + RF	Mob + MLP600	Eff S + MLP600	Mob + PCA + RF
		Plankton classes				
Actinopterygii	Actinopterygii	16	22,2	61,5	41,9	73,2
Annelida	Annelida	59	0	47,8	49,6	50,7
larvae<Annelida	Annelida	56	0	34,7	51,4	51,4
Appendicularia	Appendicularia	1963	61,1	83,1	81,9	84,6
tail<Appendicularia	Appendicularia	280	11,7	54,5	52,6	54
chainlarge	Bacillariophyceae	2157	89,9	94,8	94,7	94,5
Diatoma	Bacillariophyceae	5738	85	97	96,5	97,2
Diatoma tenuis<Diatoma	Bacillariophyceae	678	67	88,4	87,3	89,4
Thalassionema	Bacillariophyceae	42	13,3	80,5	76,6	80
Chaetognatha	Chaetognatha	473	67,1	85,9	86,2	88,6
Cirripedia	Ciripedia	1244	44,7	88,6	88,3	88
Cladocera	Cladocera	1001	45,1	75,5	75,2	75,5
Evadne	Cladocera	303	43,3	63,7	67,1	67,5
Penilia	Cladocera	60	3,2	75,2	79	77,5
Podon	Cladocera	29	0	64	63,5	67,8
Acartiidae	Copepoda	5087	57,5	92,7	92,8	92,8
Calanidae	Copepoda	1789	71,4	87,2	87,1	88,6
Calanoida	Copepoda	29663	79,5	94,2	94	94,1
Calocalanus	Copepoda	159	0	65,8	67,4	65,5
Candaciidae	Copepoda	26	0	51,2	57,1	61
Centropagidae	Copepoda	931	32,5	81,2	81,8	81,9
Copepoda	Copepoda	2784	39,7	63,2	63,6	63,8
Corycaeidae	Copepoda	549	16,1	88,3	87,9	88,4
Cyclopoida	Copepoda	20877	83,7	96	96	96,1
empty<Copepoda	Copepoda	2587	5,2	61,5	62,3	65,3
empty<Harpacticoida	Copepoda	93	0	65,5	54,2	59,3
Euchaeta	Copepoda	683	56,7	87,2	86,3	86,7
Euterpina	Copepoda	218	27,2	74,9	74,8	73,6
Harpacticoida	Copepoda	97	0	62,7	64,4	65,4
Isias	Copepoda	43	51,7	90,2	82,1	85,7
Metrinidae	Copepoda	375	38,4	83,5	81,1	83,2
Microstella	Copepoda	48	22,2	75,9	79,6	85,7
multiple<Copepoda	Copepoda	863	42,9	71,5	71,9	74,7
Oncaea	Copepoda	3501	37,1	89,8	89,9	89,4
Poecilostomatoida	Copepoda	104	0	10,9	27,8	21,3
Pontellidae	Copepoda	36	35,6	95,7	94,4	94,4
Temoridae	Copepoda	2196	31,8	90,9	91	91,7
comb<Ctenophora	Ctenophora	41	4,8	44,4	51,9	46,2
Ctenophora	Ctenophora	18	0	19	40	35
cyphonaute	cyphonaute	239	0	86,9	85,8	84,7
Decapoda	Decapoda	73	16,7	48,1	48,9	65,2
megalopa	Decapoda	215	84,4	94,8	94,6	95,9
Porcellanidae	Decapoda	133	77,5	89,3	91,2	92,7
zoea<Brachyura	Decapoda	2804	85,8	96,9	97,1	97,2
larvae<Echinodermata	Echinodermata	74	14,8	62,9	60,5	64,2
Luidiidae	Echinodermata	17	68,8	73,3	68,4	77,3

pluteus<Echinodermata	Echinodermata	362	13,8	88,1	84,7	88
egg<Actinopterygii	egg_Actinopterygii	771	87,4	96,3	95,6	96,9
egg<Engraulidae	egg_Engraudilae	6219	96,6	99,6	99,5	99,7
egg<other	egg_other	36	0	84,5	81,9	90
egg<Sardina	egg_Sardina	1164	72,5	97	96,8	97,3
Halosphaera	Halosphaera	77	72,2	85,5	73,1	77,1
Bivalvia	Mollusca	373	1,6	67,7	67,8	68,4
Cavoliniidae	Mollusca	55	3,6	84,5	86,5	85,2
egg<Mollusca	Mollusca	39	37,5	66,7	68,8	69,5
Gastropoda	Mollusca	18	46,7	72,7	70,6	70
Limacinidae	Mollusca	1534	73,6	85,8	84,8	86,2
veliger	Mollusca	20	0	29,4	40	41,4
Neoceratium	Neoceratium	14653	95,6	98,1	98,1	98,1
Noctiluca<Noctilucaceae	Noctiluca	2141	70,2	80,9	79,3	80,2
Hydrozoa	other_Cnidaria	373	37,4	75,3	75,1	75,1
Obelia	other_Cnidaria	83	0	85,9	83,1	88,8
other<Cnidaria	other_Cnidaria	24	36,4	48,8	56	60,4
Amphipoda	other_Crustacea	46	32,1	67,4	63,5	76,3
larvae<Crustacea	other_Crustacea	534	27,7	90	89,6	89,7
Nannosquillidae	other_Crustacea	13	37,5	83,3	82,8	92,3
nauplii<Crustacea	other_Crustacea	696	18,8	86,6	85,4	87,8
other<Crustacea	other_Crustacea	71	13,2	68,3	72,5	69,2
part<Crustacea	part_Crustacea	190	42,3	72,2	74	75,6
Phoronida	Phoronida	15	0	75,9	52	70,3
Rhizaria	Rhizaria	4326	80,6	92,1	91,1	91,6
Rhizosolenids	Rhizosolenids	3627	85,8	92,8	92,6	93,3
shrimp_like	shrimp_like	535	66,2	86,5	86,3	88
bract<Diphyidae	Siphonophorae	277	38,7	74,9	73,8	74,2
eudoxie<Diphyidae	Siphonophorae	108	16,7	75,6	68,7	74,9
gonophore<Diphyidae	Siphonophorae	892	53,9	73,2	71,7	72,1
nectophore<Diphyidae	Siphonophorae	211	57,3	74,9	74,5	77,7
other<Siphonophorae	Siphonophorae	107	0	11	26,9	26,9
Physonectae	Siphonophorae	27	31,2	52	63,3	60,3
Doliolida	Thaliacea	14	0	78,3	58,8	66,7
Thaliacea	Thaliacea	76	66,7	74,1	74,4	76,8
Trichodesmium	Trichodesmium	798	79,9	88,3	87,4	87,1
	average		**38,2**	**75,1**	**75**	**77,2**
			Non plankton classses			
artefact	artefact	2029	93,9	94,7	93,7	93,5
bubble	bubble	8059	90,4	97,3	97,2	97,2
detritus	detritus	30620	69,3	82,1	80,6	81,1
feces	feces	128	36,6	46,8	43,7	41,5
fiber<detritus	fiber_detritus	8304	89,1	89	88,2	88,9
gelatinous	gelatinous	178	35,3	40	39,7	43,5
light<detritus	light_detritus	7992	74,8	79,7	79,4	79,8
medium<detritus	medium_detritus	320	4,8	66,6	66,4	68,5
other<living	other_living	3819	33,8	51,1	49,7	50
fiber<plastic	plastic	28	0	37	35,6	38,3
other<plastic	plastic	635	57,5	63,3	63,6	64,7
	average		**53,2**	**68**	**67,1**	**67,9**

Table S7: Classification report for detailed classes in the Zooscan dataset. Reported values are F1-scores. The models are described in Figure 1. N test indicates the number of objects in the test set for each class. Colors reflect the values inside cells.

Taxon	Grouped	N test	Nat + RF	Mob + MLP600	Eff S + MLP600	Mob + PCA + RF
		Plankton classes				
Actinopterygii	Actinopterygii	289	23,8	87,9	91,6	94,5
egg<Actinopterygii	Actinopterygii	689	35,3	88,3	88,3	90,5
Neoceratium	Alveolata	53	0	92,3	89,5	92,7
Noctiluca	Alveolata	980	54,6	92,7	90,2	92,5
Amphipoda	Amphipoda	125	0	82,7	86,1	90,1
Cumacea	Amphipoda	78	30,4	91,2	94	94,8
Hyperiidea	Amphipoda	289	26,1	90,2	93,4	92,8
Annelida	Annelida	349	21,3	85	85,9	87,5
larvae<Annelida	Annelida	50	0	72,9	75,2	75
part<Annelida	Annelida	149	35,7	86,2	85,4	88,2
Tomopteridae	Annelida	83	7	92,1	91,8	89,6
Fritillariidae	Appendicularia	1820	28,1	89,7	88,9	90,5
Oikopleuridae	Appendicularia	4967	39,4	94,2	94,5	95
tail<Appendicularia	Appendicularia	1243	48,6	85,2	84,4	86,9
trunk	Appendicularia	193	0	67,3	67,1	72,4
Chaetognatha	Chaetognatha	7859	75,4	97,3	97,6	97,9
head<Chaetognatha	Chaetognatha	190	0	56,9	69,8	72,4
tail<Chaetognatha	Chaetognatha	555	15,3	73	75	77,6
cirrus	Cirripedia	60	9,1	68,5	59,5	68,6
cypris	Cirripedia	147	0	87,9	92,8	91,8
nauplii<Cirripedia	Cirripedia	649	0	92,2	92,4	94,3
Evadne	Cladocera	5003	17,1	96,8	97,1	97,4
Penilia	Cladocera	3592	39,9	96,8	97	97,7
Podon	Cladocera	292	0	88,3	87,8	87,6
Acartiidae	Copepoda	8853	24,2	95,5	95,4	95,9
Calanidae	Copepoda	6190	33	96,3	96,4	97
Calanoida	Copepoda	22713	57,6	94,3	94,3	94,9
Calocalanus pavo	Copepoda	71	2,7	84,2	85,5	89,9
Candaciidae	Copepoda	1767	11,9	95,5	95,1	95,5
Centropagidae	Copepoda	6890	32,8	94,6	94,6	95,1
Copilia	Copepoda	99	0	88,5	94,2	95,1
Corycaeidae	Copepoda	3576	28,5	96,3	96,6	97,2
Eucalanidae	Copepoda	183	16,8	88,4	90,2	91,3
Euchaetidae	Copepoda	1019	21,3	94,2	94,1	96,2
Haloptilus	Copepoda	407	31,8	95,6	95,4	96,5
Harpacticoida	Copepoda	832	0,2	90,7	92,7	93,1
Heterorhabdidae	Copepoda	355	0	87,6	86,2	89,3
Metridinidae	Copepoda	2439	14,7	94,6	94,6	95,7
Oithonidae	Copepoda	9847	59,2	96,6	96,6	97
Oncaeidae	Copepoda	3070	9,1	93,4	94,2	94,8
Pontellidae	Copepoda	1080	54,8	97	96,5	98,6
Rhincalanidae	Copepoda	35	52	70,2	78,3	85,3
Sapphirinidae	Copepoda	162	0	91,8	91,2	91,9
Temoridae	Copepoda	4549	23,4	96	96	96,9
Ctenophora	Ctenophora	137	0	67	72,3	81,1
cyphonaute	cyphonaute	1334	29,8	98,4	98,5	98,4

larvae<Luciferidae	Decapoda	98	16,4	95,2	95,4	97,9
larvae<Porcellanidae	Decapoda	748	64,2	96,2	97,4	98,3
megalopa	Decapoda	213	27,9	95,9	95,2	96,7
protozoea<Penaeidae	Decapoda	59	0	84,2	87,6	92,3
protozoea<Sergestidae	Decapoda	89	0	78,5	71,7	81
zoea<Brachyura	Decapoda	1750	40	95,7	96,7	97,5
zoea<Galatheidae	Decapoda	759	1,3	88,1	88,3	89,3
Doliolida	Doliolida	1461	37,7	93,2	92,4	93,8
larvae<Echinodermata	Echinodermata	76	0	80,6	76,6	84
pluteus<Echinoidea	Echinodermata	361	26,8	86,7	87,8	89,7
pluteus<Ophiuroidea	Echinodermata	542	13,4	91	92,5	92
Eumalacostraca	Eumalacostraca	3453	61,3	91,4	91,7	92,4
Eumalacostraca potentially protozoea	Eumalacostraca	225	26,1	83	81,4	83,8
larvae<Mysida	Eumalacostraca	14	0	72,7	88,9	82,8
Mysida	Eumalacostraca	120	76,5	86,4	91,6	94,4
Harosa	Harosa	244	1,6	76,7	75,1	74,2
Isopoda	Isopoda	83	67,1	98,8	97,6	98,2
Atlanta	Mollusca	68	0	84,8	83,9	90,9
Bivalvia<Mollusca	Mollusca	777	12,6	95	95,5	95,8
Cavolinia inflexa	Mollusca	662	58,2	97,5	96,2	97,2
Creseidae	Mollusca	767	47,4	93,7	94	94,2
Creseis acicula	Mollusca	1294	67,6	94,5	94,4	94,9
Cymbulia peroni	Mollusca	14	0	80	72,7	76,5
egg<Mollusca	Mollusca	129	1,5	76,7	77	75,7
Gymnosomata	Mollusca	79	60,4	92,8	95,7	95,6
Limacinidae	Mollusca	2113	25,3	96,1	96,3	96,9
part<Mollusca	Mollusca	255	2,2	61,9	55,3	60,9
Actiniaria	other_Cnidaria	22	16,7	93	93,3	89,8
ephyra	other_Cnidaria	179	36,7	86,4	91,5	91,3
Hydrozoa	other_Cnidaria	579	13,6	74,6	75,1	78,4
Obelia	other_Cnidaria	147	18,2	85,9	85,7	88,5
part<Cnidaria	other_Cnidaria	125	0	14,8	44	44,6
calyptopsis	other_Crustacea	1205	12,2	93,5	94,3	93,3
larvae<Stomatopoda	other_Crustacea	245	46,5	95,6	96,5	98,4
metanauplii<Crustacea	other_Crustacea	37	0	81,8	85,3	93,7
nauplii<Crustacea	other_Crustacea	845	4,6	91,5	91,8	93,3
Ostracoda	other_Crustacea	1169	46,4	96,4	96,7	97,6
part<Crustacea	other_Crustacea	3065	2,6	63,2	65,3	68,2
Pyrosomatida	Pyrosomatida	75	22,2	93,9	95,4	94,8
Foraminifera	Rhizaria	469	25,7	89,7	89,8	90,4
Phaeodaria	Rhizaria	8106	55,1	96,6	96,2	96,7
endostyle	Salpida	135	16	60,4	58,2	61,4
juvenile<Salpida	Salpida	67	0	82,3	84	81,9
nucleus	Salpida	222	11,5	68,6	71,4	74,7
Salpida	Salpida	2460	42,1	92,9	92,3	93,4
Bassia	Siphonophorae	15	0	57,1	50	56
bract<Abylopsis tetragona	Siphonophorae	185	34,9	91,2	89	89,9
bract<Diphyidae	Siphonophorae	2185	12	85,9	86	87,9
eudoxie<Abylopsis tetragona	Siphonophorae	98	0	90,3	92,1	89,6
eudoxie<Diphyidae	Siphonophorae	525	2,9	84,3	86,9	89,9

gonophore<Abylopsis tetragona	Siphonophorae	199	12,1	90,9	90,2	93,5
gonophore<Diphyidae	Siphonophorae	2460	30	93,2	93,4	94,2
nectophore<Abylopsis tetragona	Siphonophorae	173	20,7	88,6	87,6	91,7
nectophore<Diphyidae	Siphonophorae	4417	63,1	92,9	92,2	93,1
nectophore<Hippopodiidae	Siphonophorae	17	18,2	73,3	81,1	85,7
nectophore<Physonectae	Siphonophorae	1386	59,5	87,4	81,8	84,7
part<Siphonophorae	Siphonophorae	412	0	66,8	67,4	69,5
Physonectae	Siphonophorae	16	0	43,5	48,5	66,7
siphonula	Siphonophorae	144	19,2	90,3	86,1	89
Coscinodiscus	Stramenopiles	1075	41,2	97,3	96,8	97,2
actinula<Solmundella bitentaculata	Trachylina	19	0	68,8	78,9	82,4
Aglaura	Trachylina	455	57,9	91,8	91,7	93
Liriope<Geryoniidae	Trachylina	34	0	52	73	78,7
Rhopalonema velatum	Trachylina	373	49,1	85,6	85,2	87,2
Solmundella bitentaculata	Trachylina	56	3,5	67,4	70,6	73,4
average			**22,9**	**85,5**	**86,6**	**88,5**
		Non plankton classses				
artefact	artefact	7718	76,7	80,8	80	79,8
badfocus<artefact	badfocus	6046	19,6	63,1	62,9	63,1
bubble	bubble	2432	19	92,2	91	91,2
detritus	detritus	36260	55,2	82,9	81,4	81,6
fiber<detritus	fiber	6708	62,9	74,6	74,7	74,8
Insecta	Insecta	169	27,1	84,3	86,9	89,6
egg<other	other_egg	2015	59,7	92,2	91	92,4
other<living	other_living	40	16,3	39,2	59,3	73,7
seaweed	seaweed	1272	35,3	68	68,2	66,3
average			**41,3**	**75,2**	**77,3**	**79,2**

Table S8: Detailed and grouped metrics for all models across all datasets. Acc. = accuracy, Bal. acc. = balanced accuracy, Mac. F1 = macro-averaged F1, Mic. F1 = micro-averaged F1, Pk. pr. = plankton precision, Pk. rec. = plankton recall.

		Detailed metrics						Grouped metrics					
Dataset	Model	Acc.	Bal. acc.	Mac. F1	Mic. F1	Pk. pr.	Pk. rec.	Acc.	Bal. acc.	Mac. F1	Mic. F1	Pk. pr.	Pk. rec.
FlowCAM	Nat + RF	0.77	0.26	0.27	0.77	0.62	0.61	0.81	0.31	0.33	0.81	0.67	0.64
FlowCAM	Nat + RF (NW)	0.77	0.22	0.23	0.77	0.62	0.59	0.80	0.26	0.28	0.80	0.68	0.62
FlowCAM	Mob + MLP600	0.89	0.77	0.74	0.89	0.86	0.86	0.93	0.79	0.77	0.93	0.88	0.89
FlowCAM	Mob + MLP600 (NW)	0.90	0.72	0.73	0.90	0.87	0.87	0.93	0.76	0.77	0.93	0.90	0.89
FlowCAM	Mob + MLP50	0.89	0.75	0.73	0.89	0.85	0.87	0.93	0.79	0.76	0.93	0.88	0.89
FlowCAM	Eff S + MLP600	0.91	0.82	0.79	0.91	0.89	0.89	0.94	0.85	0.82	0.94	0.91	0.92
FlowCAM	Eff XL + MLP600	0.89	0.78	0.73	0.89	0.85	0.87	0.93	0.81	0.74	0.93	0.88	0.90
FlowCAM	Mob + RF	0.90	0.69	0.74	0.90	0.87	0.87	0.94	0.74	0.77	0.94	0.90	0.89
FlowCAM	Mob + PCA + RF	0.90	0.69	0.73	0.90	0.86	0.86	0.93	0.74	0.77	0.93	0.89	0.89
FlowCAM	Random	0.13	0.01	0.01	0.13	0.05	0.05	0.32	0.02	0.02	0.32	0.01	0.01
IFCB	Nat + RF	0.86	0.36	0.40	0.86	0.60	0.58	0.86	0.48	0.50	0.86	0.63	0.59
IFCB	Nat + RF (NW)	0.86	0.29	0.34	0.86	0.63	0.55	0.86	0.41	0.46	0.86	0.66	0.56
IFCB	Mob + MLP600	0.94	0.82	0.72	0.94	0.80	0.85	0.94	0.86	0.80	0.94	0.81	0.87
IFCB	Mob + MLP600 (NW)	0.94	0.75	0.73	0.94	0.83	0.84	0.94	0.84	0.82	0.94	0.84	0.86
IFCB	Mob + MLP50	0.94	0.81	0.73	0.94	0.80	0.84	0.94	0.85	0.80	0.94	0.82	0.85
IFCB	Eff S + MLP600	0.94	0.85	0.76	0.94	0.81	0.87	0.94	0.87	0.82	0.94	0.82	0.89
IFCB	Eff XL + MLP600	0.93	0.71	0.61	0.93	0.76	0.80	0.93	0.80	0.72	0.93	0.77	0.82
IFCB	Mob + RF	0.95	0.75	0.76	0.95	0.84	0.86	0.95	0.83	0.84	0.95	0.85	0.87
IFCB	Mob + PCA + RF	0.95	0.73	0.76	0.95	0.84	0.85	0.95	0.82	0.84	0.95	0.86	0.86
IFCB	Random	0.55	0.01	0.01	0.55	0.01	0.01	0.55	0.03	0.03	0.55	0.01	0.01
ISIIS	Nat + RF	0.87	0.38	0.35	0.87	0.50	0.42	0.87	0.42	0.39	0.87	0.61	0.48
ISIIS	Nat + RF (NW)	0.87	0.28	0.30	0.87	0.57	0.29	0.87	0.31	0.33	0.87	0.68	0.32
ISIIS	Mob + MLP600	0.94	0.76	0.71	0.94	0.79	0.85	0.95	0.79	0.74	0.95	0.85	0.92
ISIIS	Mob + MLP600 (NW)	0.95	0.71	0.72	0.95	0.82	0.83	0.96	0.74	0.75	0.96	0.88	0.90
ISIIS	Mob + MLP50	0.94	0.76	0.70	0.94	0.78	0.85	0.95	0.79	0.73	0.95	0.84	0.92
ISIIS	Eff S + MLP600	0.94	0.81	0.73	0.94	0.79	0.86	0.95	0.84	0.76	0.95	0.85	0.93
ISIIS	Eff XL + MLP600	0.92	0.74	0.63	0.92	0.73	0.84	0.93	0.75	0.66	0.93	0.79	0.91
ISIIS	Mob + RF	0.95	0.71	0.73	0.95	0.83	0.83	0.96	0.74	0.76	0.96	0.89	0.89
ISIIS	Mob + PCA + RF	0.95	0.71	0.73	0.95	0.82	0.83	0.96	0.74	0.76	0.96	0.88	0.89
ISIIS	Random	0.61	0.03	0.03	0.61	0.02	0.03	0.61	0.04	0.04	0.61	0.04	0.05
UVP6	Nat + RF	0.91	0.33	0.38	0.91	0.62	0.38	0.91	0.40	0.47	0.91	0.73	0.40
UVP6	Nat + RF (NW)	0.91	0.24	0.29	0.91	0.65	0.29	0.91	0.30	0.38	0.91	0.78	0.30
UVP6	Mob + MLP600	0.92	0.77	0.67	0.92	0.67	0.78	0.93	0.82	0.71	0.93	0.74	0.88
UVP6	Mob + MLP600 (NW)	0.94	0.68	0.69	0.94	0.74	0.72	0.95	0.74	0.74	0.95	0.84	0.80
UVP6	Mob + MLP50	0.92	0.74	0.63	0.92	0.65	0.76	0.93	0.78	0.67	0.93	0.74	0.85
UVP6	Eff S + MLP600	0.92	0.79	0.64	0.92	0.65	0.79	0.93	0.83	0.68	0.93	0.71	0.89
UVP6	Eff XL + MLP600	0.89	0.57	0.39	0.89	0.52	0.64	0.90	0.63	0.44	0.90	0.64	0.74
UVP6	Mob + RF	0.94	0.68	0.70	0.94	0.75	0.72	0.95	0.72	0.75	0.95	0.86	0.80
UVP6	Mob + PCA + RF	0.94	0.66	0.69	0.94	0.75	0.72	0.95	0.71	0.74	0.95	0.86	0.79
UVP6	Random	0.62	0.02	0.02	0.62	0.01	0.01	0.62	0.03	0.03	0.62	0.03	0.03
ZooCAM	Nat + RF	0.76	0.37	0.40	0.76	0.77	0.75	0.83	0.54	0.55	0.83	0.89	0.86
ZooCAM	Nat + RF (NW)	0.75	0.32	0.34	0.75	0.78	0.73	0.83	0.50	0.52	0.83	0.91	0.84
ZooCAM	Mob + MLP600	0.88	0.80	0.74	0.88	0.91	0.92	0.91	0.84	0.79	0.91	0.95	0.96
ZooCAM	Mob + MLP600 (NW)	0.89	0.75	0.75	0.89	0.91	0.92	0.92	0.81	0.80	0.92	0.96	0.96
ZooCAM	Mob + MLP50	0.88	0.78	0.73	0.88	0.91	0.92	0.91	0.83	0.78	0.91	0.95	0.96
ZooCAM	Eff S + MLP600	0.89	0.82	0.76	0.89	0.91	0.93	0.92	0.87	0.81	0.92	0.95	0.97
ZooCAM	Eff XL + MLP600	0.88	0.81	0.71	0.88	0.91	0.92	0.91	0.86	0.77	0.91	0.95	0.96
ZooCAM	Mob + RF	0.89	0.71	0.74	0.89	0.92	0.91	0.92	0.78	0.80	0.92	0.96	0.96
ZooCAM	Mob + PCA + RF	0.89	0.71	0.74	0.89	0.92	0.91	0.92	0.79	0.80	0.92	0.96	0.96
ZooCAM	Random	0.07	0.01	0.01	0.07	0.07	0.07	0.18	0.03	0.03	0.18	0.28	0.28
ZooScan	Nat + RF	0.47	0.24	0.24	0.47	0.52	0.41	0.61	0.38	0.39	0.61	0.73	0.61
ZooScan	Nat + RF (NW)	0.46	0.18	0.19	0.46	0.55	0.38	0.58	0.31	0.34	0.58	0.77	0.55
ZooScan	Mob + MLP600	0.89	0.88	0.86	0.89	0.93	0.94	0.91	0.90	0.88	0.91	0.96	0.97
ZooScan	Mob + MLP600 (NW)	0.90	0.85	0.86	0.90	0.93	0.93	0.92	0.87	0.88	0.92	0.96	0.96
ZooScan	Mob + MLP50	0.89	0.86	0.86	0.89	0.93	0.93	0.91	0.88	0.87	0.91	0.96	0.97
ZooScan	Eff S + MLP600	0.90	0.90	0.88	0.90	0.94	0.95	0.92	0.91	0.89	0.92	0.96	0.97
ZooScan	Eff XL + MLP600	0.89	0.89	0.85	0.89	0.93	0.94	0.91	0.90	0.88	0.91	0.95	0.97
ZooScan	Mob + RF	0.90	0.81	0.85	0.90	0.94	0.93	0.92	0.85	0.87	0.92	0.97	0.96
ZooScan	Mob + PCA + RF	0.90	0.82	0.85	0.90	0.94	0.93	0.92	0.85	0.87	0.92	0.97	0.96
ZooScan	Random	0.05	0.01	0.01	0.05	0.03	0.03	0.16	0.03	0.03	0.16	0.02	0.02