Dagmar Kainmueller, Integrative Imaging Data Sciences (HPI, Helmholtz Imaging, MDC Berlin)
Project partners: Rainer Kiko (GEOMAR), Klas-Ove Möller (hereon), David Greenberg (hereon)

Contact: dagmar.kainmueller@mdc-berlin.de

HELMHOLTZ IMAGING · GEOMAR · Helmholtz-Zentrum hereon · MAX DELBRÜCK CENTER · HPI Hasso Plattner Institut — Digital Engineering · Universität Potsdam

AqQua — **The Aquatic Life Foundation Project:** Quantifying Life at Scale in a Changing World

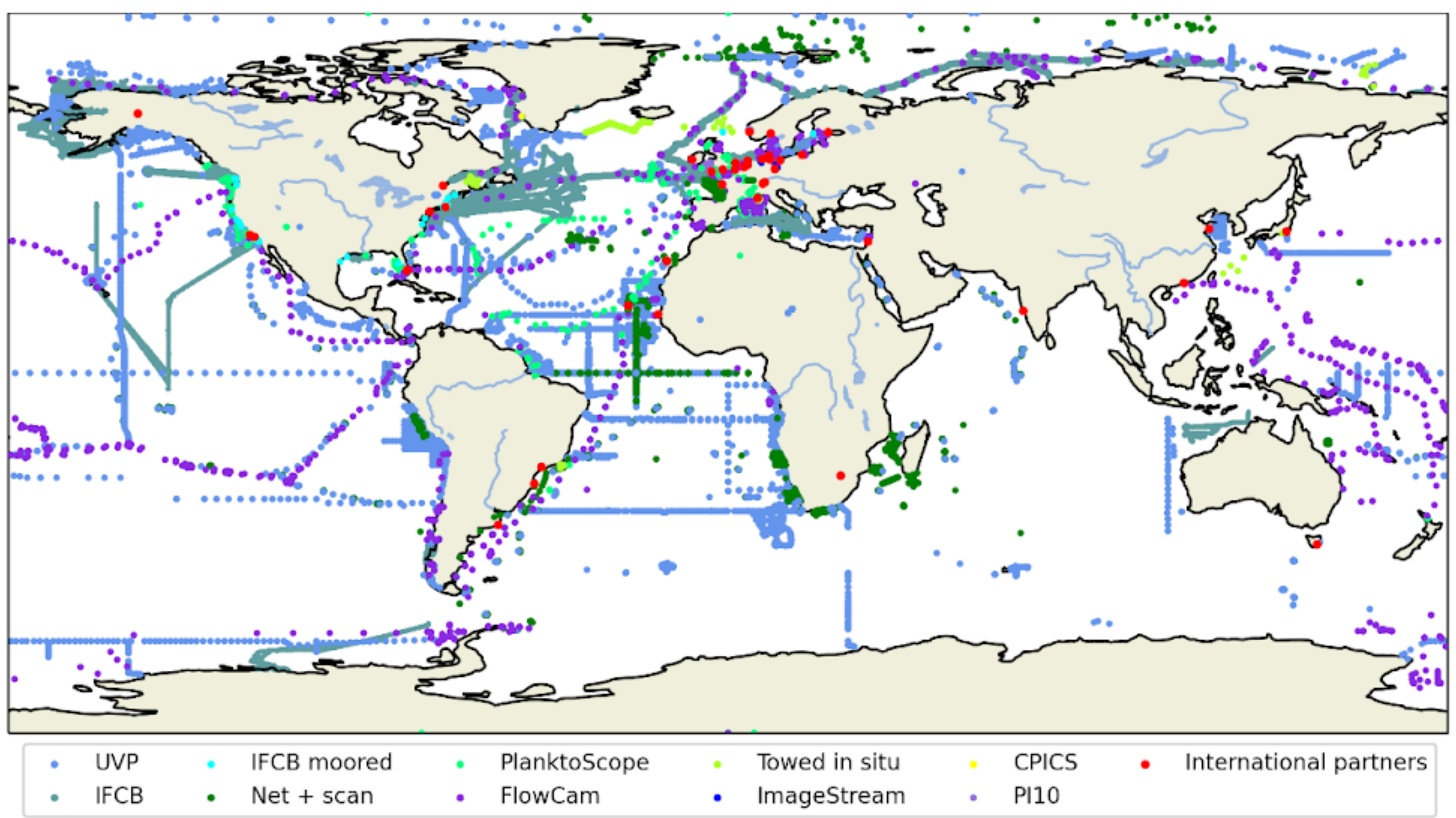
Human well-being to a large extent depends on aquatic life

Organic matter formed by plankton sustainably sequesters vast amounts of carbon from the atmosphere in the ocean's interior. The efficiency of this so-called biological carbon pump is likely altered by climate changes that affect planktonic diversity. Furthermore, climate change-induced reduction of food web efficiency directly impacts oceanic fish resources that play a major role in human nutrition. The critical role of aquatic life in biogeochemical cycles, for climate regulation, conservation of aquatic biodiversity and human nutrition mandates precise mapping and monitoring, to provide essential knowledge and to serve decision making in the face of global change, e.g. with respect to emerging carbon dioxide removal techniques. To this end, distributed *pelagic imaging* techniques enable the sustained observation of aquatic life and its debris, comprehensively covering the earth's water bodies down to the bottom of the deep sea.

The Aquatic Life Foundation Project will, for the first time, combine more than three billions of images acquired with a variety of devices across the globe for large-scale training of a foundational pelagic imaging model, which will be fine-tuned for species classification, trait extraction and particulate organic carbon estimation. The model will allow to establish global maps of species biodiversity, ecosystem status, and carbon flux at unprecedented accuracy and granularity, thereby generating a fundamental understanding of marine and freshwater life in times of global change.

Billions of images of aquatic organisms are acquired around the globe in a distributed fashion

Operational global monitoring of aquatic life requires full automation and a consistent approach across imaging systems. A range of different imaging systems are in use, and respective image characteristics and information content differ across modalities. Yet data from different systems are in principle compatible as all systems image aquatic life. The vast and highly diverse amounts of available data, together with the inherent scarcity of the available labels and the distinct information content of pelagic imaging compared to standard computer vision data, calls for a dedicated, large-scale domain-specific self-supervised learning approach.



Compilation of globally existing, accessible pelagic imaging datasets (blue dots) and location of project partner institutions (red)

References

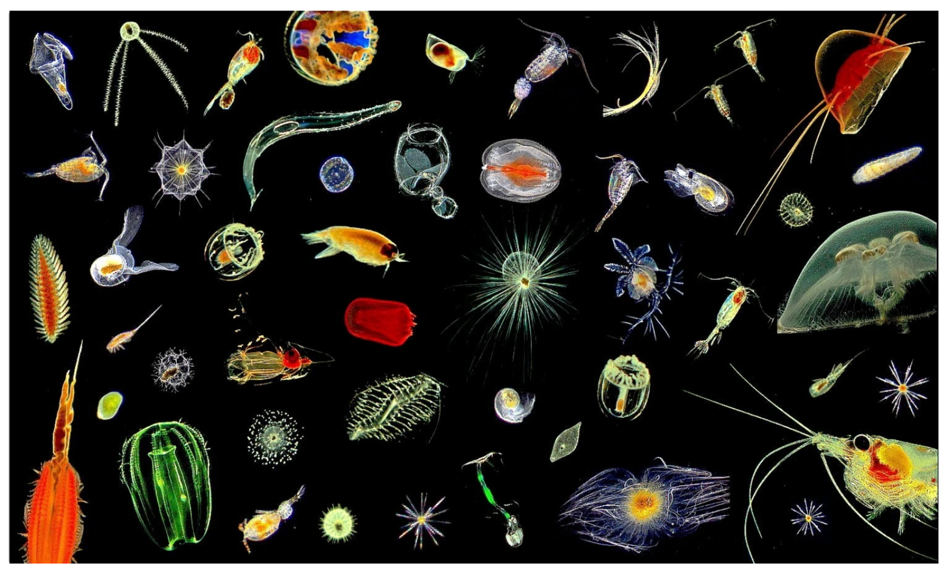
Oquab et al.: *DINOv2: Learning Robust Visual Features without Supervision*. TMLR, 2024.

Dehghani et al.: *Patch n'pack: Navit, a vision transformer for any aspect ratio and resolution.* NeurIPS, 2024.

Reed et al.: *Scale-mae: A scale-aware masked autoencoder for multiscale geospatial representation learning.* CVPR 2023.

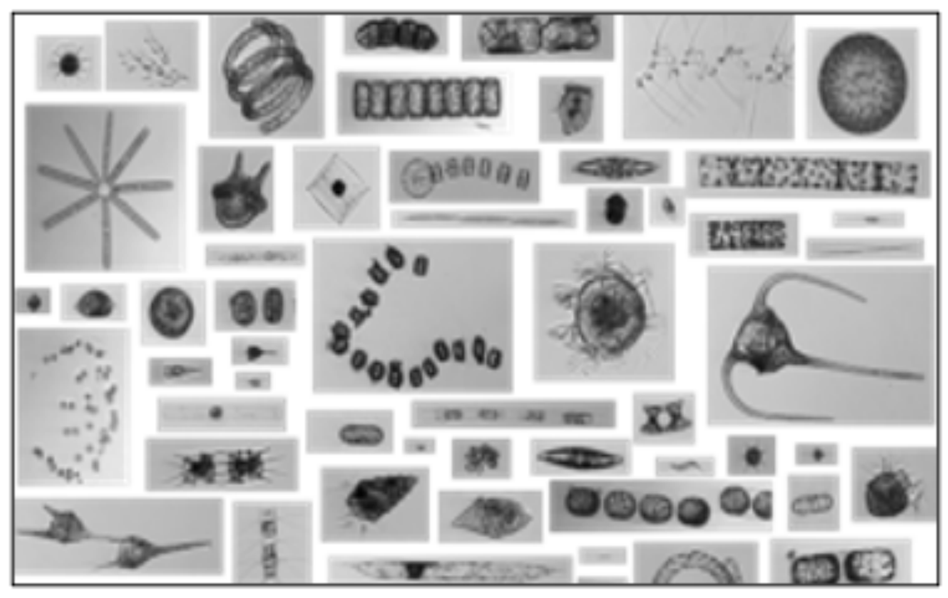
Three billion images are available for this project, with high diversity across modalities and geolocations

Continuous Plankton Identification and Classification System (CPICS)



Underwater Vision Profiler (UVP)

Imaging Flow Cytobot (IFCB)



Exemplary pelagic imaging systems and their outputs

Project Objective: Optimize Large-scale Self-supervised Learning for Plankton Classification

Foundation model architecture and training: The project builds upon the DINOv2 self-supervised learning pipeline (Oquab et al. 2024), which integrates contrastive learning and masked image modeling for Vision Transformers (ViTs), with demonstrated success across domains. Several unique aspects of pelagic imaging call for enriching DINOv2 with domain knowledge to optimize performance: (1) Pelagic organisms and particles vary greatly in size. It is advisable to extend DINOv2 respectively, e.g. with a Native Resolution ViT (Dehghani et al. 2024). (2) Image resolution varies across imaging modalities. This information could be harvested e.g. via resolution-aware position embeddings (Reed et al. 2023). (3) Data curation and balancing are vital for the success of DINOv2, yet specifics of pelagic imaging call for a tailored approach towards optimal data diversity. (4) Image augmentations need to be optimally tailored to the data. Your task is to explore these and potential further-reaching adaptations to devise a self-supervised learning pipeline that is optimally tailored to pelagic imaging data.

Model exploration and understanding: Interactive data exploration leveraging UMAP-style dimensionality reduction allows to visually relate the structure of learned image embedding spaces with available information on plankton species, functional traits, particulate organic carbon content and image modality. This is vital for targeted model optimization. A prototype data exploration tool developed by Helmholtz Imaging at DKFZ is available for the project and to be expanded as needed.

Benchmarking: Available manual annotations of plankton species and -traits need to be leveraged to form a benchmark / leaderboard for quantifying and tracking model performance.

Compute: A budget of 200.000 GPU-h at the HoreKa HPC-system at KIT is available for the project. A baseline DINOv2 pipeline is operational on this system and scales to multi-node training.

Your Skill Set

- Curiosity and genuine interest in conducting research, incl. publication of results
- Strong programming skills in python; experience with pytorch
- Attendance of HPI course *Machine Learning for Image Analysis* (or equivalent)
- Degree program: IT Systems Engineering