

Making sense of underwater particle images with PyOPIA



SINTEF

Raymond Nepstad^{1,*}, Arsalan Mostaani², Dana King¹, Tor Nordam¹, Emlyn Davies¹

¹ SINTEF Ocean, Trondheim, Norway

² Department of Physics, Norwegian University of Science and Technology, Trondheim, Norway

*raymond.nepstad@sintef.no

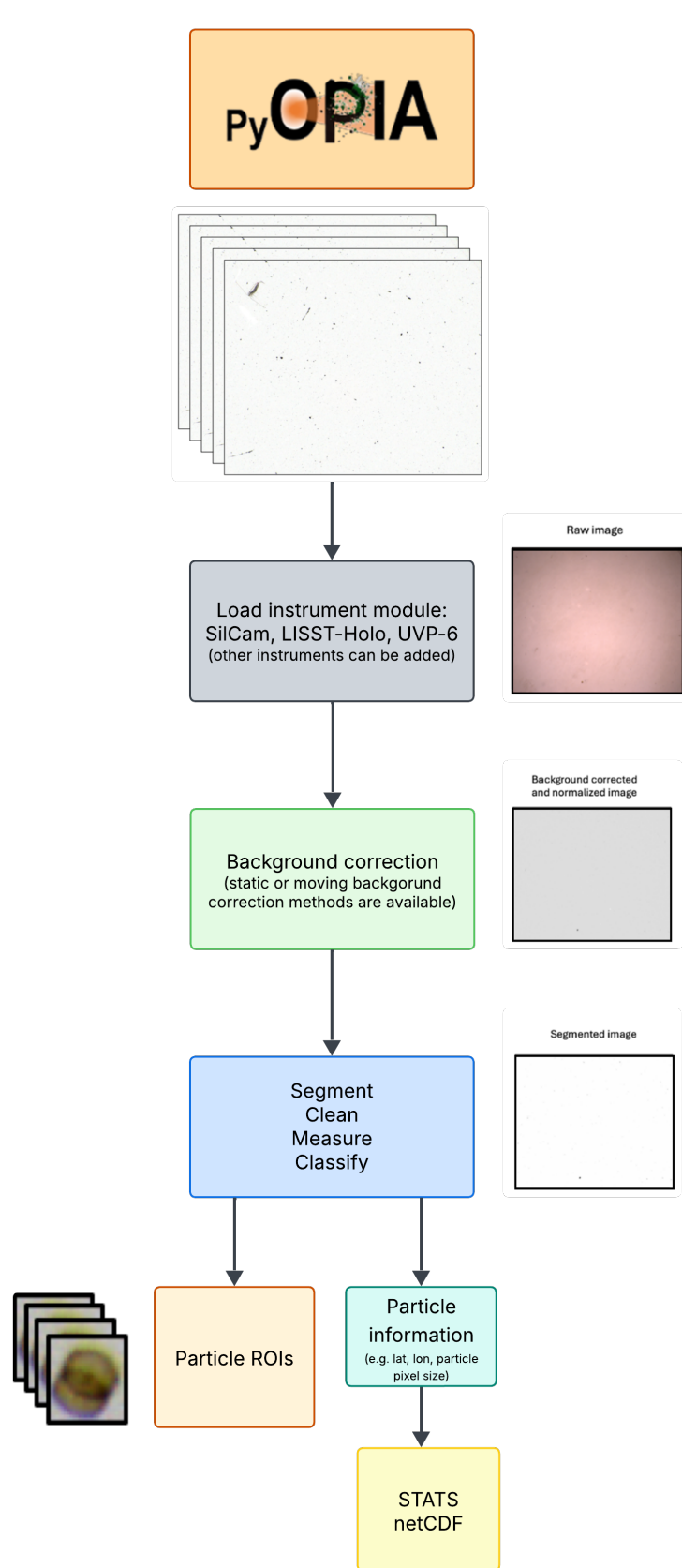
Abstract

PyOPIA (Python Ocean Particle Image Analysis) [1, 2] is an open-source Python library for combing through large amounts of particle image data collected by camera instruments such as SilCam [5], LISST-Holo and UVP-6. Using a standardized and user-customizable pipeline, images are cleaned and segmented, before individual particles are geometrically measured and classified (CNN-based). The resulting information is stored in standardized netCDF files for further use. Example applications of PyOPIA include the study of oil droplet and bubble formation under breaking waves in a laboratory flume [3], data analysis pipeline component in a Digital Twin of the Ocean [4] and estimating copepod concentrations and sizes from ship-based casts.

Challenges

1. Segmentation (under/over), lighting, focus
2. Particle classifier robustness
3. Auxillary data cleaning and normalization
4. Particle overlap and high concentrations

Image processing and analysis



The steps in the PyOPIA pipeline can easily be customised for different applications, and description of the pipeline in a toml format ensures reproducibility. There is also functionality for tagging each identified particle with auxillary data from other instruments (e.g., CTD to associate each particle with a temperature and salinity).

Multi-instrument analysis

Image data from multiple instruments can be run through the same analysis pipeline, facilitating direct comparison of particle statistics.

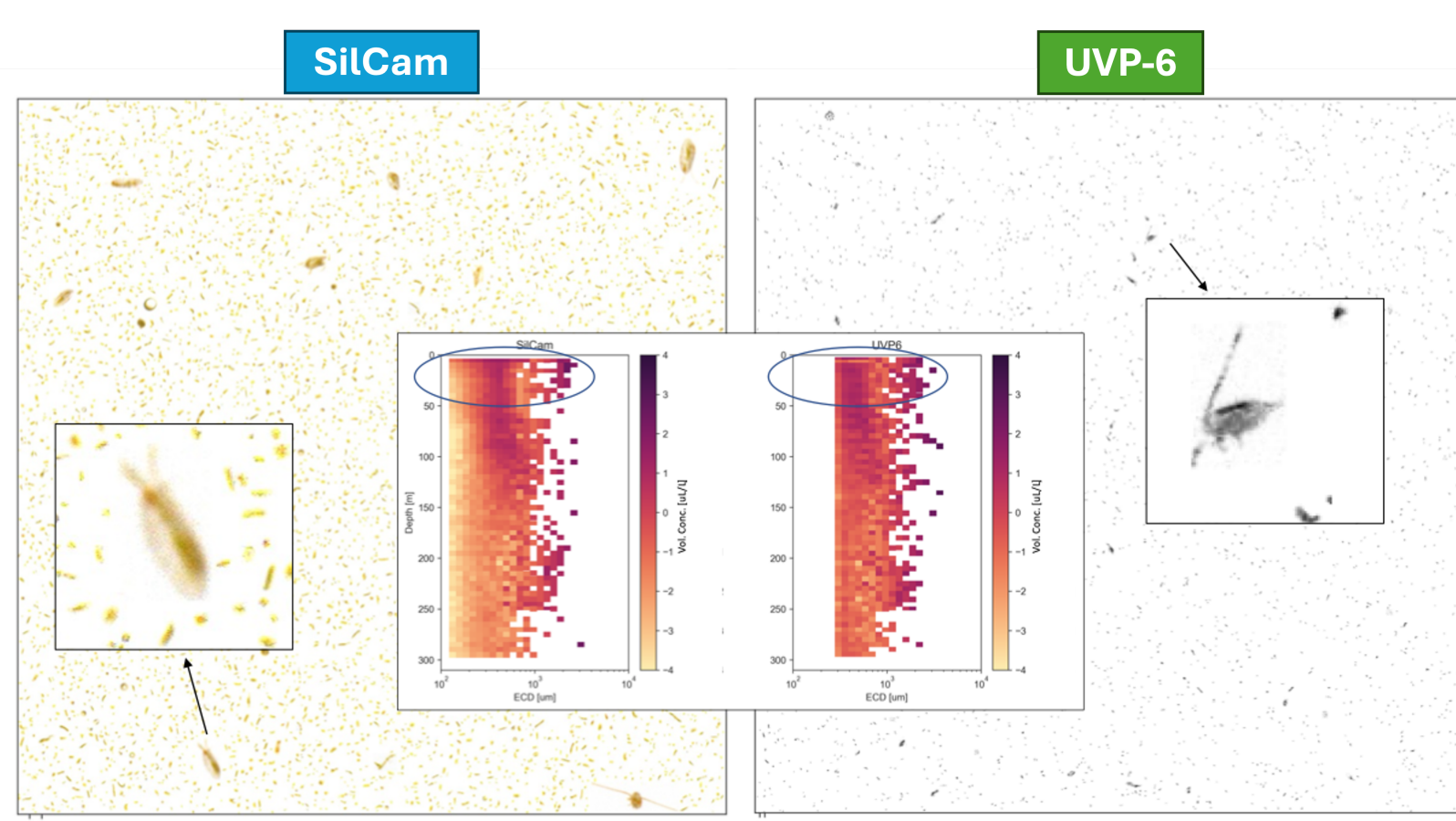


Figure credit: Karoline Barstein (NTNU) [2].

Testing and benchmarking

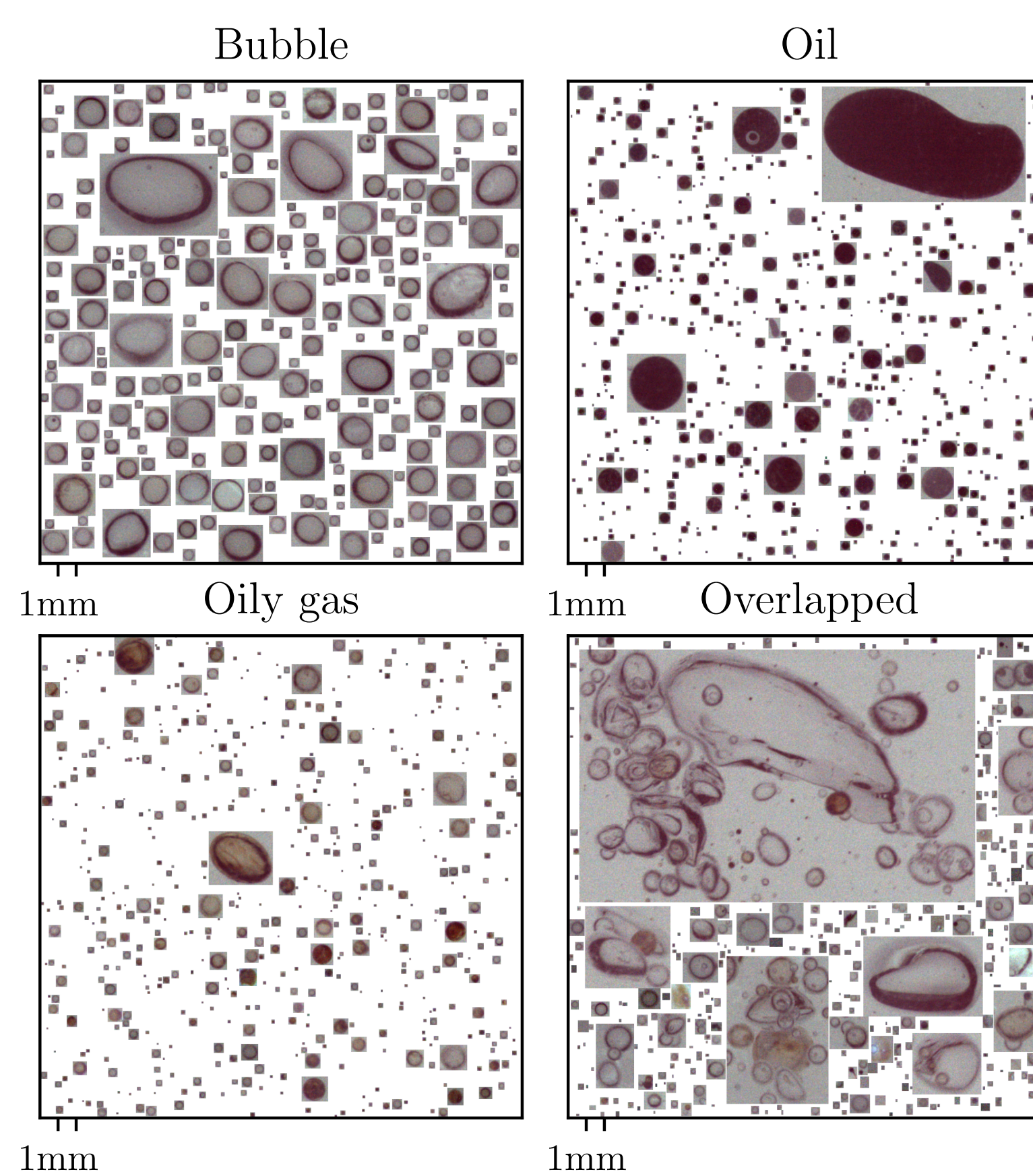
Consistency and correctness over time is supported by automated tests of basic functionality and expected analysis results (Github actions). You can also simulate processing of synthetic images from specified size distribution and camera parameters, using the simulation module.

Classifying particles

- PyOPIA includes a small (5 layer, 200 k parameters) convolutional neural network particle classifier as default, trained on 11 617 images belonging to 7 classes (bubble, copepod, diatom chain, faecal pellets, oil, oily gas, other)
- Works well for particles similar to training data, but does not generalize well - retraining may be needed on specific datasets
- Future directions: Self-supervised pre-training, distillation from foundation models

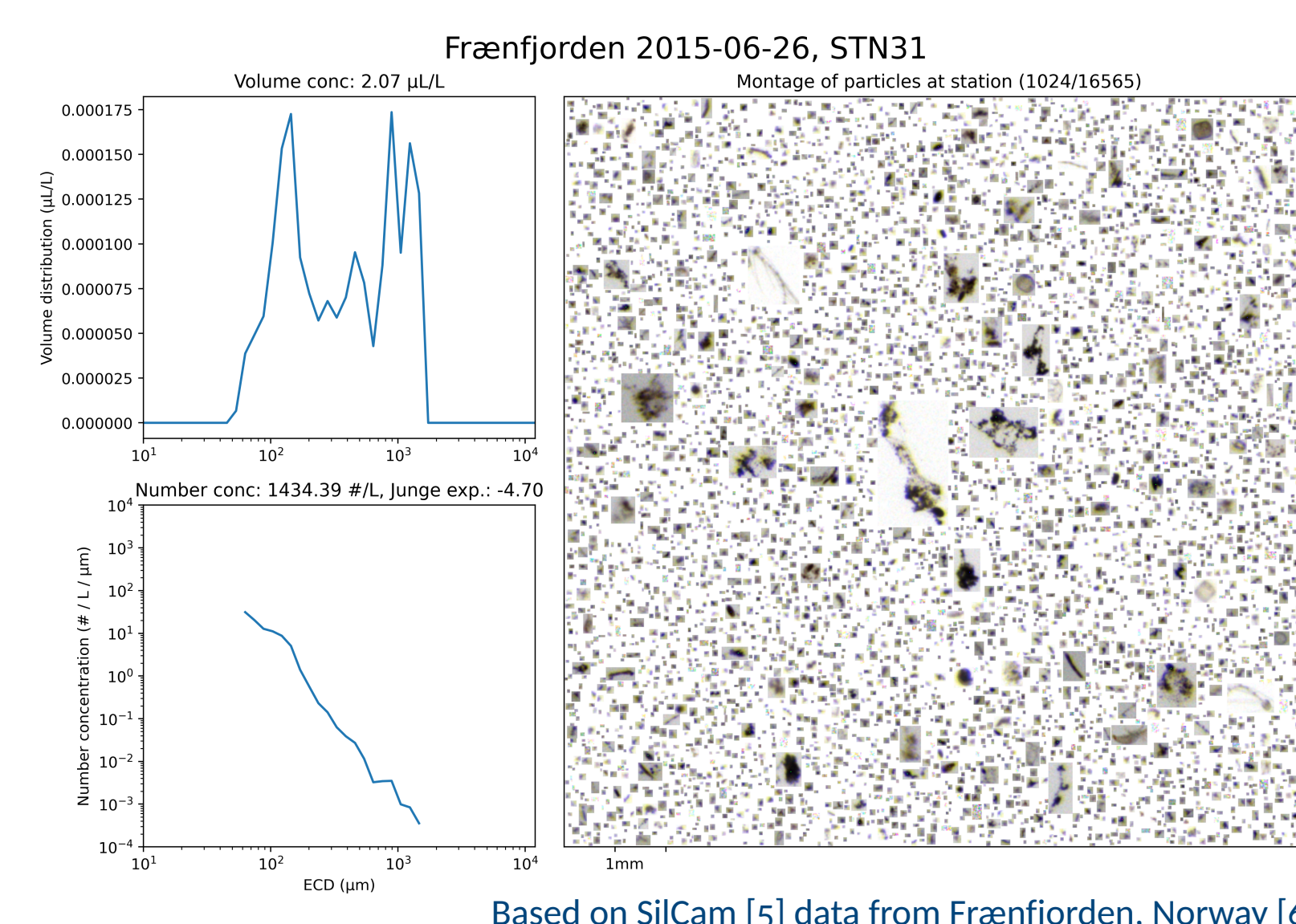
Re-training for oil droplet classification

Retrained network for separating bubble, oil and oily gas from overlapping particles, ca. 10 000 labelled images per class (Mostaani *et al.*, submitted 2026).



Particle statistics

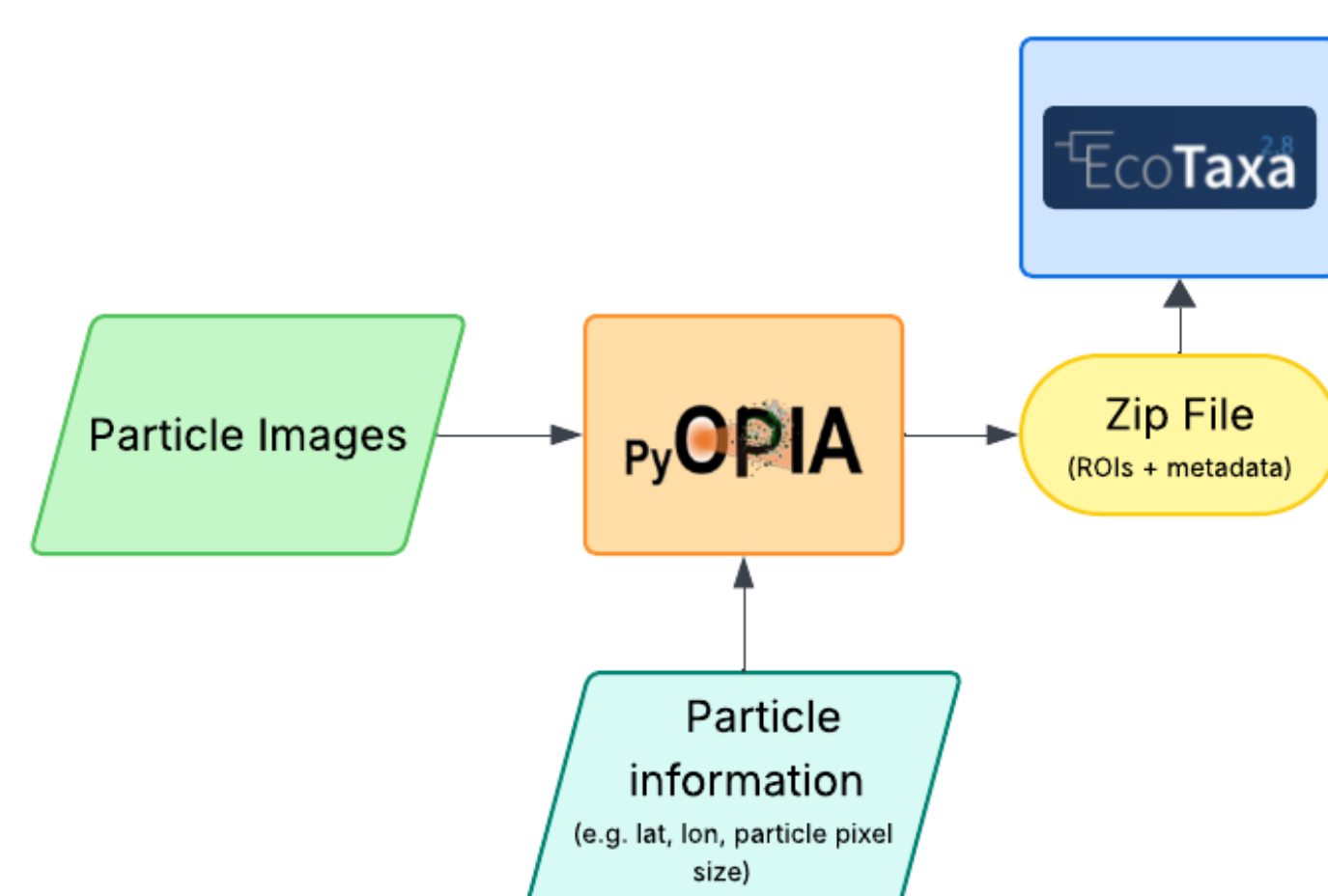
Summary particle statistics, such as number and volume distributions can be easily made with PyOPIA. Montages of randomly selected subset of particle images from different scales provides a visual impression and check of the data.



Based on SilCam [5] data from Frøenfjorden, Norway [6].

From PyOPIA to EcoTaxa

Import particle data from PyOPIA into EcoTaxa



PyOPIA output

- Annotated netCDF file with metadata: per particle statistics and per image statistics
- Auxillary data: depth, temperature, position etc.
- Reproducibility: PyOPIA version and pipeline config included



Run PyOPIA on EDITO

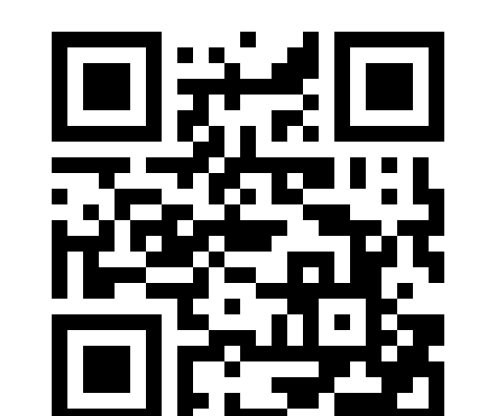
A small PyOPIA tutorial is available on the EDITO Datalab platform. This demonstrates an end-to-end analysis of a small example image dataset.



<https://div.eedito.eu/training?search=pyopia>

Get PyOPIA

Source code and documentation is available on Github (github.com/SINTEF/PyOPIA).



PyOPIA Contributors

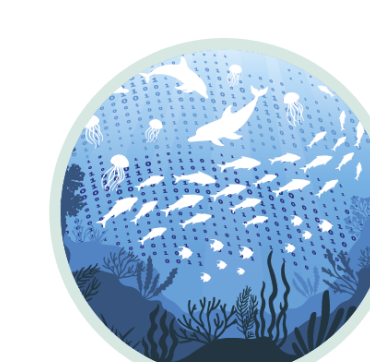
Several people from different institutions have contributed to PyOPIA development (Github users ids): emlynjdavies, animmosmith, nepstad, arsalanmstn, nordam, DanaMaKing, EE-Shuan-Song, klarriou, sarigiering.

References

1. <https://github.com/SINTEF/PyOPIA>
2. E Davies, A Nimmo-Smith, *et al.* "PyOPIA" Particles in Europe (PIE) 2023.
3. A Mostaani, *et al.* "Clouds of bubbles and droplets; formation and distribution under breaking waves." EGU 2024.
4. A Vasiljevic *et al.* "A digital twin of the Trondheim fjord for environmental monitoring—a pilot case." JMSE 2024.
5. E Davies *et al.* "The use of wide-band transmittance imaging to size and classify suspended particulate matter in seawater" MPB 2017.
6. E Davies, R Nepstad "In situ characterisation of complex suspended particulates surrounding an active submarine tailings placement site in a Norwegian fjord" RSIMS 2017.

Funding

This work was partially funded by the European Union through the DTO-BioFlow project. PyOPIA has been developed with support from multiple projects (including SFI Harvest) and contributors.



Integration of biodiversity monitoring data into the Digital Twin Ocean

