ASTRAL

All Atlantic Ocean Sustainable Profitable and Resilient Aquaculture

A data integration pipeline towards reliable monitoring of phytoplankton and early detection of harmful algal blooms

Bruna Guterres[1], Sara Khalid[2], Marcelo Pias[1], Silvia Botelho[1]

[1] Federal University of Rio Grande
[2] University of Oxford

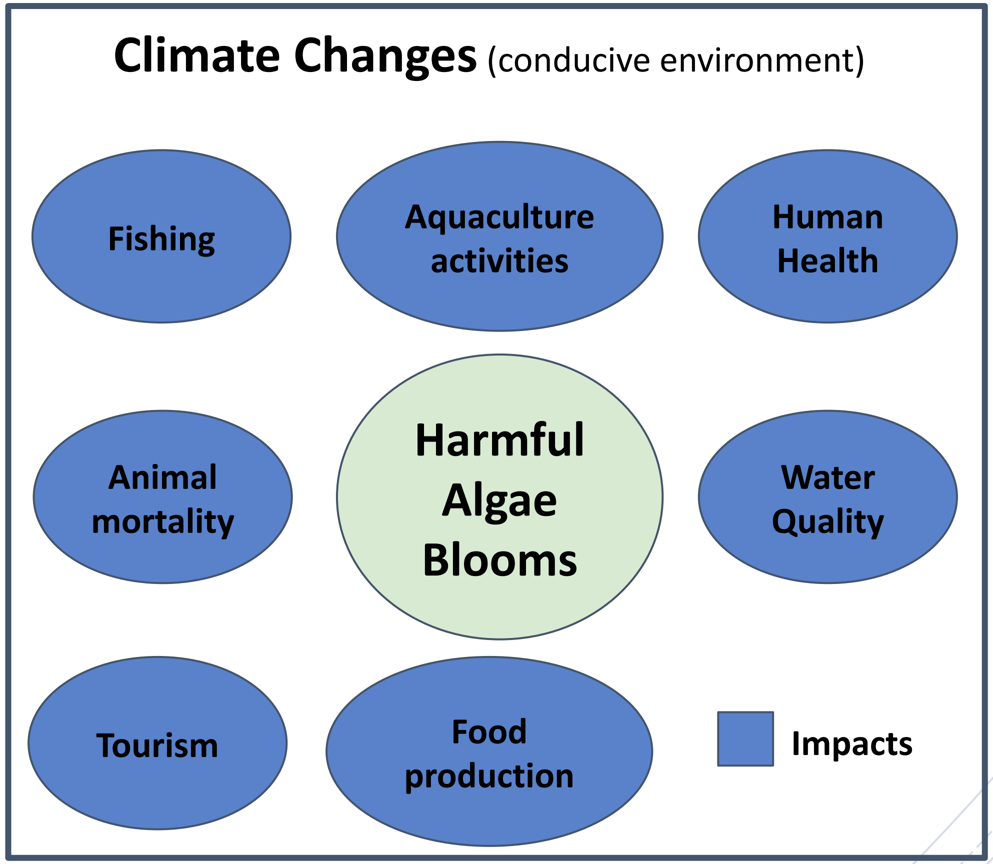
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 863034.

Introduction

- Climate change is causing progressive warming, acidification, and de-oxygenation of oceans [1].
- Phytoplankton monitoring is vital for protecting marine life and developing climate-resilient economies
- Public phytoplankton image databases have several limitations that prevent the practical usage of artificial intelligent models.

Proposed Solution and Climate Impact

- Pipeline for integration and standardization of image databases;
- It can be applied for curation of real-world data and training of scalable AI models (e.g. early detection of HAB outbreaks), ultimately contributing to climate resilience and adaptation;



Methodology

Genus	Aquaculture farm
Alexandrium	Argentina, UK
Anabaena	Brazil
Azadinium	Ireland, UK
Centric	Ireland
Chaetoceros	Argentina, Ireland, Brazil
Cilliates	Ireland
Dinophysis	UK, South Africa
Euglena	Ireland
Fragilaria	Ireland
Gonyaulax	UK, South Africa

Genus	Aquaculture farm
Karenia	Ireland, UK, South Africa
Katodinium	Ireland, UK
Leptocylindrus	Argentina, Ireland
Lingulodinium	UK, South Africa
Mesodinium	Ireland
Nematodinium	Ireland
Nodularia	Brazil
Paralia	Ireland
Pennate	Ireland
Prorocentrum	Ireland, UK

Genus	Aquaculture farm
Protoceratium	South Africa
Pseudo-nitzschia	Argentina, Ireland, UK, South Africa
Rhizosolenia	Ireland
Scrippsiella	Ireland
Skeletonema	Ireland, Brazil
Tetraselmis	Brazil
Thalassiosira	Argentina, Brazil
Tripos	Ireland

*Argentina ( ), Brazil ( ), Ireland ( ), South Africa ( ) and UK ( )

Figure 1 - Target phytoplankton organisms within aquaculture farms of Brazil, South Africa, Argentina, Ireland, South Africa and Scotland. The information is organized by genus.

Methodology

Data Integration Pipeline

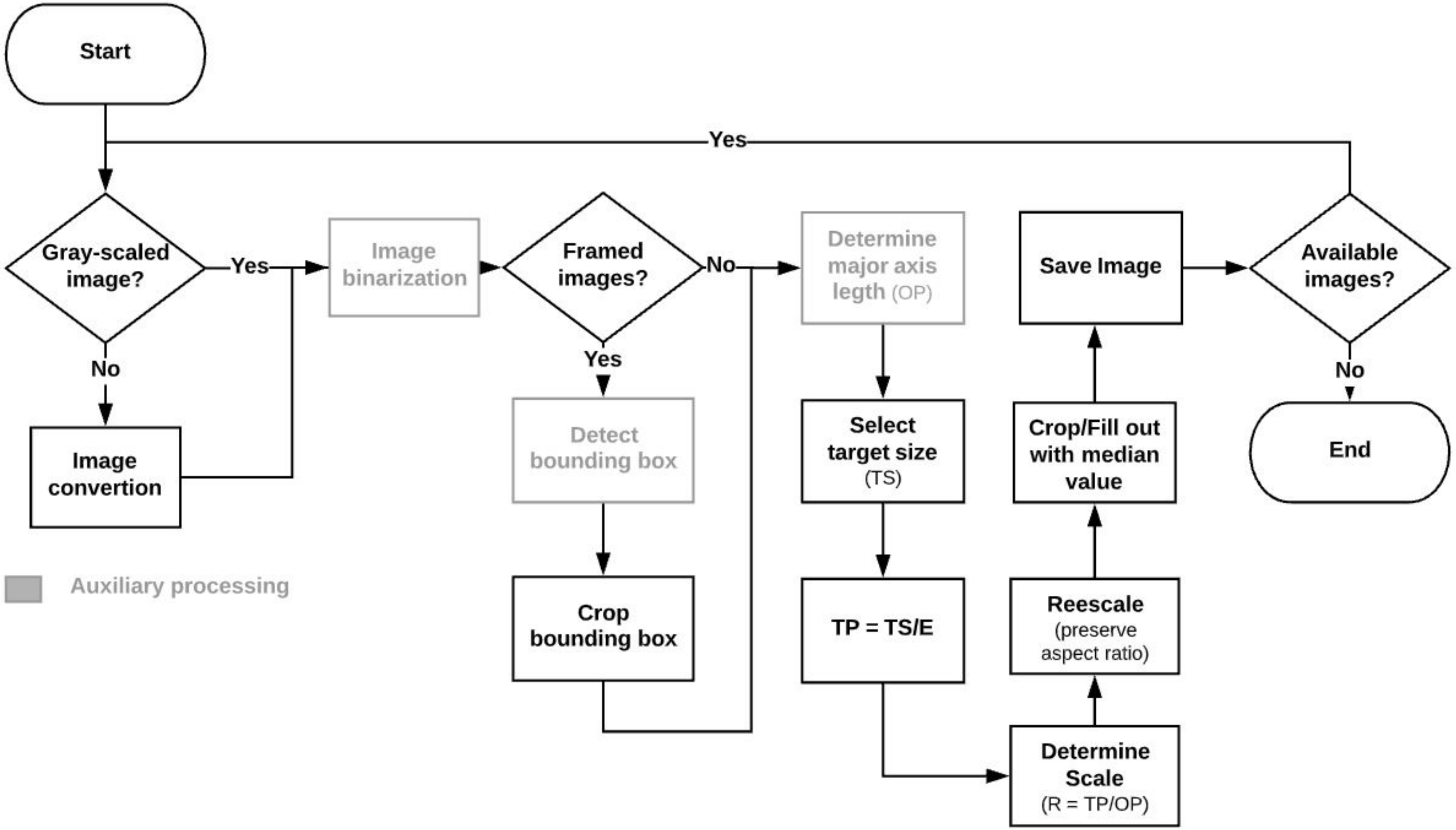


Figure 2 - Pipeline for dataset integration. Target Size (TS) is a random selected value between minimum and maximum expected size of each phytoplankton genus. TP represents the target pixel size considering an output scale E (um / pixel).

Results and Discussions

- Fourteen public phytoplankton image datasets were identified from the literature.
- Some databases (29%) do not contain genus-level images for any target phytoplankton.

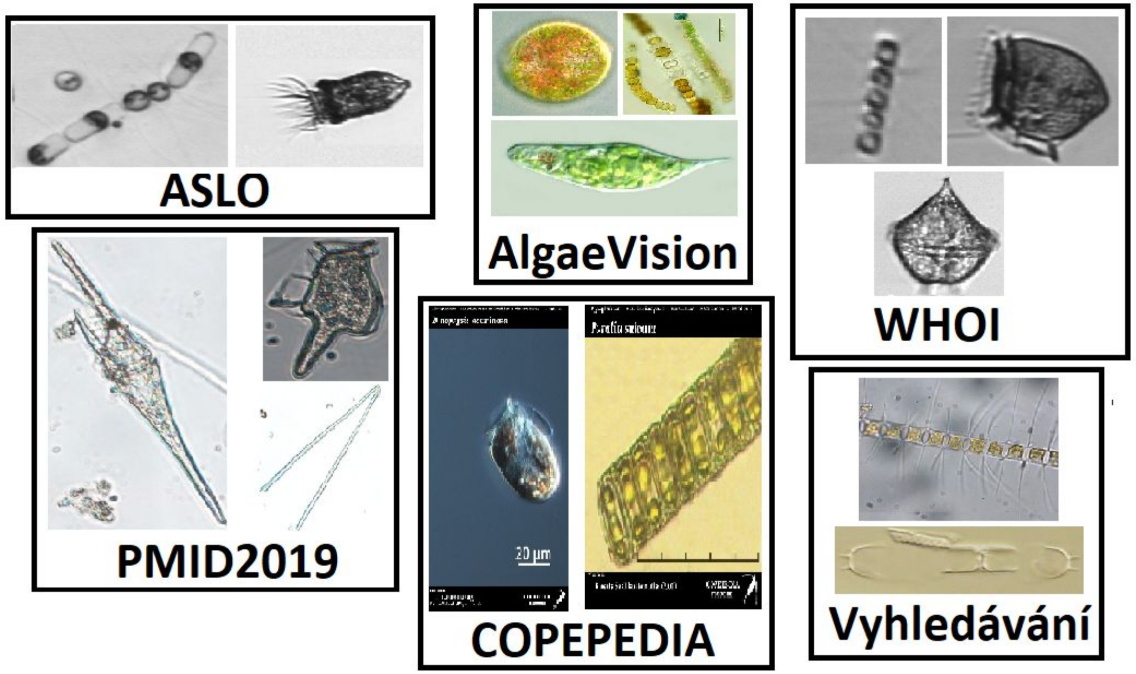


Figure 3 - Images of some phytoplankton species identified in six different public databases.

Results and Discussions

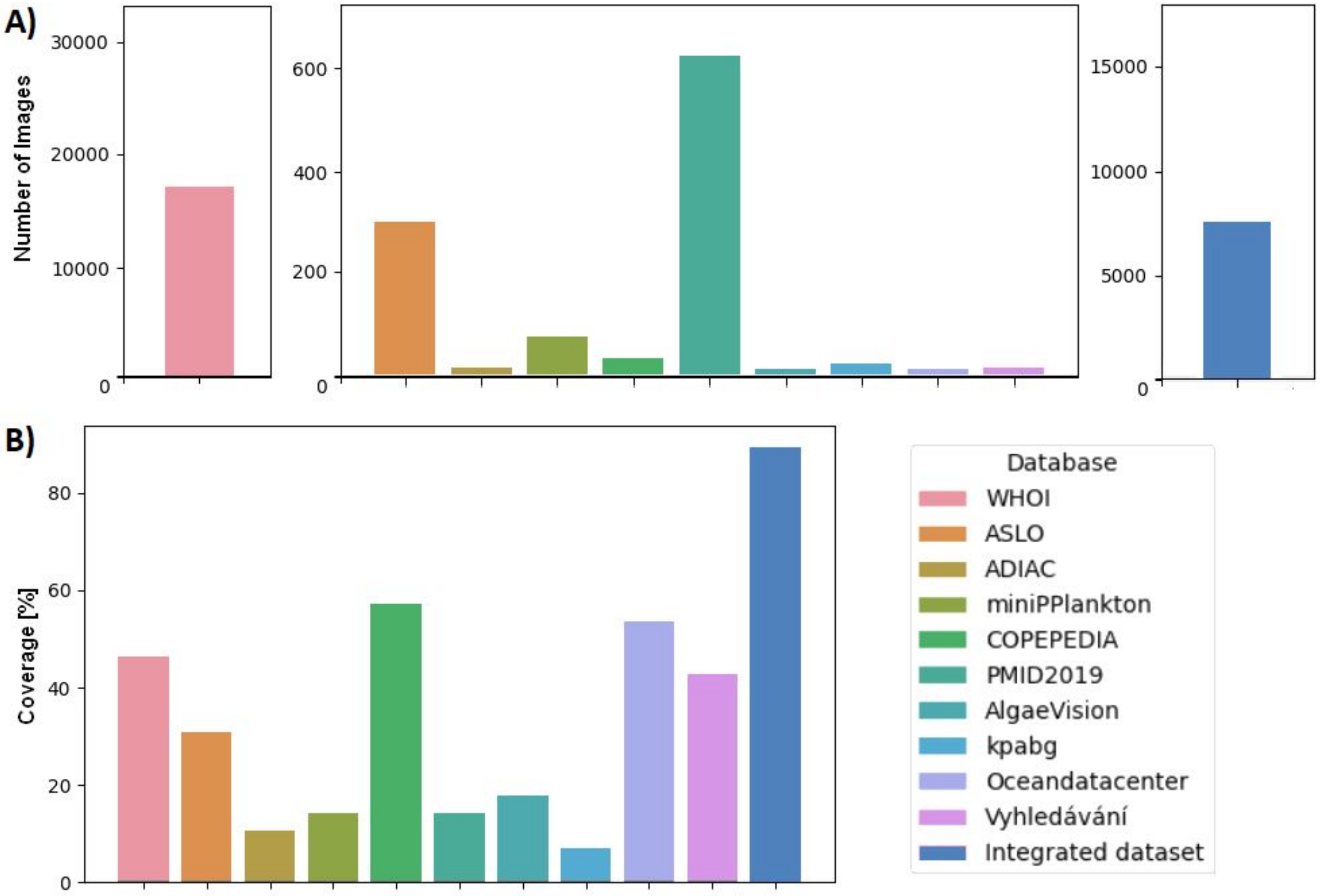


Figure 4 - Coverage (A) and number of images (B) within original and integrated databases. The number of images is presented as the average and standard deviation of the number of images within covered genus.

Conclusions

- Integration pipeline for phytoplankton image-based datasets;
- Unified, benchmark database covering publicly available databases;
- Increased coverage from an average of 26% to 89% considering species in the natural marine environment;
- Important tool towards versatile machine learning models;
  - Planning protection and resilience of marine ecosystems in the face of climate change;

ASTRAL

Thank you!

Bruna Guterres

PhD Student

astral@norceresearch.no

astral-project.eu

@ASTRAL_H2020

ASTRALH2020

astral-h2020project

ASTRAL_H2020

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 863034.

Bibliography

[1] Christopher J Gobler. Climate change and harmful algal blooms: insights and perspective. Harmful algae, 91:101731, 2020

[2] Planktonmkl. https://github.com/zhenglab/PlanktonMKL

[3] Algaevision. http://algaevision.myspecies.info

[4]Heidi M Sosik and Robert J Olson. Automated taxonomic classification of phytoplankton sampled with imaging-in-flow cytometry.Limnology and Oceanography: Methods, 5(6):204–138216, 2007

[5] Qiong Li, Xin Sun, Junyu Dong, Shuqun Song, Tongtong Zhang, Dan Liu, Han Zhang, and Shuai Han. Developing a microscopic image dataset in support of intelligent phytoplankton detection using deep learning.ICES Journal of Marine Science, 77(4):1427–1439, 2020

[6]Copepedia: The database of taxonomy, distribution maps, photos, and biometric traits. https://www.st.nmfs.noaa.gov/copepod/about/about-copepedia.html

[7] Vyhledávání.http://galerie.sinicearasy.cz/galerie.