Semi-automatic identification of phytoplankton using image classification techniques

Amadei Martínez Luz, Debusschere Elisabeth, Mortelmans Jonas, Deneudt Klaas and Hernandez Francisco

Flanders Marine Institute (VLIZ), InnovOcean site, Wandelaarkaai 7, 8400 Oostende, Belgium
E-mail: luz.amadei.martinez@vliz.be

Phytoplankton is the base of the marine food web. Its changes in abundance and species composition are indicators of the health of the marine ecosystem and has an effect on trophic levels of economic importance. In addition, harmful algal blooms may cause losses to aquaculture, fisheries, tourism and human health. Furthermore, phytoplankton has an influence on the global climate, due to its role in the carbon cycle and in the production of precursors of clouds.

Long-term phytoplankton data series are necessary to understand changes and pressures on an ecosystem; therefore, they are relevant to achieve the Sustainable Development Goals from United Nations. As traditional methodologies to obtain these type of data are labor intensive and time consuming, the Flanders Marine Institute (VLIZ), as part of the Belgian LifeWatch observatory, decided to apply image analysis for the taxonomic classification of such samples.

From 2017 onwards, VLIZ is sampling phytoplankton on 17 stations in the Belgian Part of the North Sea (BPNS). Those samples are processed with a Flow Cytometer And Microscope (FlowCAM®), resulting in a collection of pictures. Each picture captures a particle and a set of measurements associated to that particle. These produced images are then further analysed with the software VisualSpreadsheet to annotate taxonomic names to these particles.

Using VisualSpreadsheet, we have created digital libraries of the phytoplankton taxa for the BPNS. Based on those libraries, we generate filters for each taxon that allow us to run an auto-classification for the pictures of each sample. The output of the auto-classification is manually validated by an experienced operator, to correct the errors of the automatic prediction. The quality-controlled datasets generated are freely available through the LifeWatch Data Explorer (http://www.lifewatch.be/en/lifewatch-data-explorer), on the condition that the dataset is cited.

Keywords: Biodiversity; Image analysis; Phytoplankton; Belgian Part of the North Sea; FlowCAM