From microplastics to pixels: Testing the robustness of two machine learning approaches for automated, Nile red-based marine microplastic identification.

Nelle Meyers*[1,2,3], Bavo De Witte[2], Natascha Schmidt[4,5], Dorte Herzke[6,7], Jean-Luc Fuda[8], David Vanavermaete[2], Colin Janssen[3,9], and Gert Everaert[1]

[1]Flanders Marine Institute (VLIZ) – Innovocean Site, Jacobsenstraat 1, 8400 Oostende, Belgium
[2]Flanders Research Institute for Agriculture, Fisheries and Food (ILVO) – InnovOcean Campus, Jacobsenstraat 1, 8400 Oostende, Belgium
[3]Ghent University (Ghent University) – Faculty of Bioscience Engineering, Department of Animal Sciences and Aquatic Ecology, Blue Growth Research Lab, 9000 Gent, Belgium
[4]Norwegian Institute for Air Research (NILU) – The FRAM Centre, P.O. Box 6606, 9296 Tromsø, Norway
[5]Mediterranean Institute of Oceanography (MIO) – Aix Marseille Université (Aix-en-Provence), Toulon University, CNRS, Institut de recherche pour le développement [IRD] – Mediterranean Institute of Oceanography, UM 110, Marseille, France
[6]Norwegian Institute for Air Research (NILU) – The FRAM Centre, P.O. Box 6606, 9296 Tromsø, Norway
[7]Norwegian Institute of Public Health [Oslo] (NIPH) – PO Box 222 Skøyen, 0213 Oslo, Norway
[8]Mediterranean Institute of Oceanography (MIO) – Aix Marseille Université (Aix-en-Provence), Toulon University, CNRS, Institut de recherche pour le développement [IRD] – Mediterranean Institute of Oceanography, UM 110, Marseille, France
[9]Ghent University (Ghent University) – Blue Growth Research Lab, Bluebridge, Wetenschapspark 1, 8400 Oostende, Belgium

Abstract

Despite the pressing need for accurate and robust observations of microplastics in the marine environment to assess current and future environmental risks, existing procedures remain labour-intensive. Additionally, microplastic analysis is challenged by environmental weathering, affecting the reliability of studies relying on pristine plastics. This study addresses these knowledge gaps by testing the robustness of two automated analysis techniques which combine machine learning algorithms with fluorescent colouration of Nile red-stained particles. The study evaluated and compared the accuracy of models based on decision tree and random forest classifiers in detecting and identifying microplastic polymers weathered under semi-controlled surface water and deep-sea conditions, for a duration of one year. Furthermore, the analysis time, model complexity, lower size limit, and interoperability of the machine learning based approach were assessed. The decision tree and random forest models demonstrated a comparable accuracy in detecting and identifying pristine plastic polymers

*Speaker

(both $>$ 90%). For the detection of weathered microplastics, both models achieved sufficiently high detection accuracies ($>$ 77%), but only the random forest model was reliable for polymer identification ($>$ 70%), except for PET particles. The random forest models showed an accuracy $>$ 90% for particle predictions based on 12-30 pixels, corresponding to microplastics sized $<$ 10 μm. The results indicated that the knowledge rules generated by the random forest classifier did not produce consistent results across different labs. However, the inherent flexibility of the method allows for a rapid adaptation and optimisation of the models, ensuring the possibility to fine-tune the method to specific research goals through customised datasets, thereby enhancing its robustness.

Keywords: Microplastics, Machine learning, Automation, Fluorescence, Nile red, Weathered plastics, Monitoring, Marine pollution