

EGU2020-7606

<https://doi.org/10.5194/egusphere-egu2020-7606>

EGU General Assembly 2020

© Author(s) 2022. This work is distributed under the Creative Commons Attribution 4.0 License.



## First-order machine learning based detection and classification of foraminifera in marine sediments from Arctic environments

**Steffen Aagaard-Sørensen**<sup>1</sup>, Thomas Haugland Johansen<sup>2</sup>, and Juho Junttila<sup>1</sup>

<sup>1</sup>Department of Geoscience, UiT The Arctic University of Norway, Tromsø, Norway

<sup>2</sup>Department of Mathematics and Statistics, UiT The Arctic University of Norway, Tromsø, Norway

Foraminifera are microscopic single-celled organisms, ubiquitous to the marine realm, that construct shells during their life cycle. The shells, in general, fossilize well in the sediment and they are diagnosable due to inter-species morphology and ornamentation variability. Classifying and counting foraminiferal shells is an important tool in assessing and reconstructing past and present environmental, oceanographic and climatological conditions. However, the present day manual identification procedure, performed with a microscope and a needle/brush, is a very time consuming. Circumventing this manual procedure, using machine learning, promises to dramatically lower the time consumption related to generating foraminiferal data records.

The first step towards that end is developing a deep learning model that can detect and classify microscopic foraminifera from 2D digital microscope pictures. The work is based on a VGG16 model implementation that has been pre trained on the ImageNet dataset and employing transfer learning techniques to adapt the model to the foraminifera task. The 2D photographic training data input was constructed by combining objects representative of and extracted from Arctic marine sediments (100µm-1mm size fraction) from the Barents Sea region. Four object groups, including 1) calcareous and 2) agglutinated benthic foraminifera, 3) planktic foraminifera and 4) sediments were used in the training data construction. With the initial set-up the algorithms were able to identify adherence to one of the four groups correctly ~90% of the time and with further fine-tuning and refinement reaching 98% correct identifications.

The second step is to use machine learning for classification of individual benthic calcareous foraminiferal species within the sediment. The work will focus on the 20 most common species that comprise ca. ≥ 80% of the total benthic calcareous foraminiferal fauna in the Arctic. The training of the algorithms will be done using targeted species-specific 2D photographic and 3D CT scanning data in addition to potentially using hyperspectral imaging.