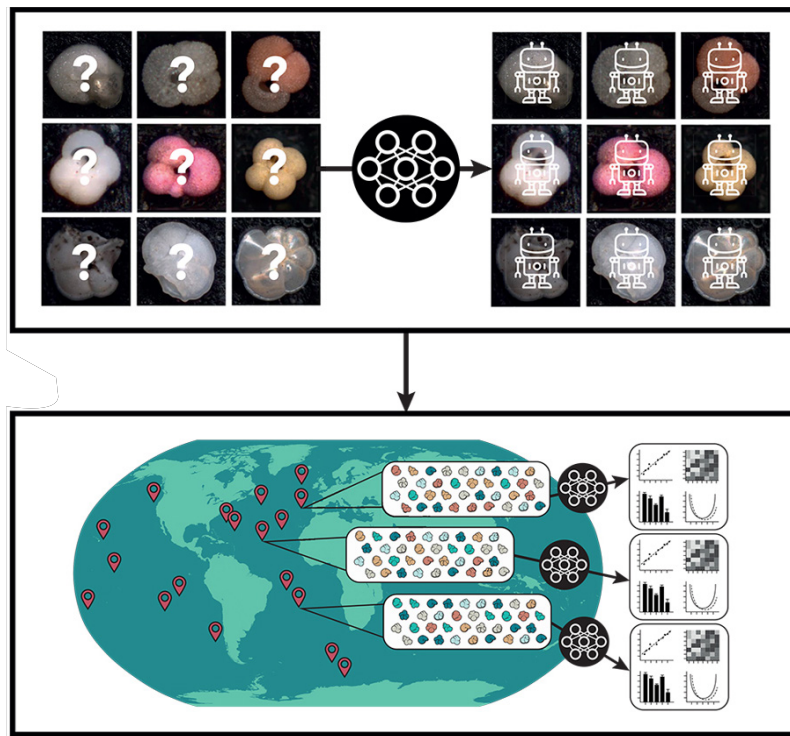




Endless Forams: Automating large-scale geobiological studies using deep learning

Allison Hsiang, Researcher at the Department of Geological Sciences

► October 28, 2021 at 15h00 in William-Olsson lecture hall and by Zoom

<https://stockholmuniversitet.zoom.us/j/69573004265>

Planktonic foraminiferal species identification is central to many paleoceanographic studies, from selecting species for geochemical research to elucidating the biotic dynamics of microfossil communities relevant to physical oceanographic processes and interconnected phenomena such as climate change. However, few resources exist to train students in the difficult task of discerning amongst closely related species, resulting in diverging taxonomic schools that differ in species concepts and boundaries. This problem is exacerbated by the limited number of taxonomic experts. Here we document our initial progress toward removing these confounding and/or rate-limiting factors by generating the first extensive image library of modern planktonic foraminifera, providing digital taxonomic training tools and resources, and automating species-level taxonomic identification of planktonic foraminifera via machine learning using convolution neural networks. Experts identified 34,640 images of modern (extant) planktonic foraminifera to the species level. These images are served as species exemplars through the online portal Endless Forams (endlessforams.org) and a taxonomic training portal hosted on the citizen science platform Zooniverse (zooniverse.org/projects/ahsiang/endless-forams). A supervised machine learning classifier was then trained with ~27,000 images of these identified planktonic foraminifera. The best-performing model provided the correct species name for an image in the validation set 87.4% of the time and included the correct name in its top three guesses 97.7% of the time. Together, these resources provide a rigorous set of training tools in modern planktonic foraminiferal taxonomy and a means of rapidly generating assemblage data via machine learning in future studies for applications such as paleotemperature reconstruction.

28 | IGV
OCT | Seminar
Series

Department of Geological Sciences

Contact information: wei-li.hong@geo.su.se, paola.manzotti@geo.su.se & christian.stranne@geo.su.se



**Stockholms
universitet**