

The impact of experts' inter- and intra-observer agreements on computational object classification in images from the deep seafloor of the HAUSGARTEN observatory

Timm Schoening¹, Melanie Bergmann², Autun Purser³, Julian Gutt², Jennifer Dannheim², James Taylor⁴, Tim W. Nattkemper¹, Antje Boetius⁵

¹ Biodata Mining Group, Faculty of Technology, Bielefeld University, Bielefeld, Germany

² HGF-MPG Group for Deep-Sea Ecology and Technology, Alfred Wegener Institute for Polar and Marine Research, Bremerhaven, Germany

³ School of Engineering and Science, Jacobs University, Bremen, Germany

⁴ University of Glasgow, Glasgow, United Kingdom

⁵ Max Planck Institute for Marine Microbiology, Bremen, Germany

Far-sighted marine research institutions around the globe are capturing images from the seafloor at a scale of hundreds of thousands. Only a small part of these data have been accessed to date, as manual analyses are time-consuming and automated evaluation approaches are still under development.

Machine learning and neural networks have been identified as a promising algorithmic approach to automate analysis of images from the seafloor. These algorithms need ground-truth data about the objects to be detected. As the information provided by one human expert lacks reproducibility, the expertise of a group of individuals has to be employed to collect training data as well as to evaluate the performance of an automated detection.

In this paper we show that the inter- and intra-observer agreements of these human experts is a critical factor for the training of a learning architecture and has shown to be conditional to image quality for some object classes. A supervised automated detection approach is evaluated where five experts marked the positions of eight distinct object classes within seventy images taken at the HAUSGARTEN observatory (eastern Fram Strait, Arctic). Support Vector Machines were trained to detect and classify objects in the images with an overall sensitivity of 0.87 and precision of 0.67. A detailed comparison of the human expert agreements showed interesting correlations with the system's performance and pointed us towards new strategies for (semi-) automated underwater image analysis.