How to Empower the Geologist's Expert Knowledge with Machine Learning in Drill Core Logging?

Christian Günther[1], Filip Siman[1], Nils Jansson[1], Foteini Simistira-Liwicki[1], Marcus Liwicki[1], Tobias Hermansson[2], Paul McDonnell[2]
1. Luleå University of Technology, Luleå, Sweden, 2. Boliden Mines, Boliden, Sweden

Drill core logging by geologists is essential in mineral exploration, which in turn is essential for securing the substantial amount of mineral resources needed to meet current societal goals, such as the net zero goal. However, when several geologists are involved in the core logging process, subjectivity commonly leads to inconsistent data sets, complicating further analysis and geological modelling. Additionally, core logging is time-consuming and cost-intensive. Machine learning (ML) could be a solution to more objective and efficient drill core logging. Still, current approaches in research are in the beginning and generally specific to their used context given by training and test data sets. Therefore, new innovative and flexible solutions are needed to aid core logging.

We present a solution where we integrate drill core logging with machine learning methods that, due to the complexity of the task, utilise a human-in-the-loop (HITL) approach to transform and empower the geologist's workflow. First, the geologist's visual analysis can be enhanced using ML labelling tools based on computer vision principles, allowing the geologist to label findings on drill core photos directly. The labels are then learned by existing ML models via transfer learning, so that these can predict labels themselves. Next, the geologist relabels incorrect ML labels to improve the ML model, with the potential to create a digital assistant in the analysis. With this approach, several ML models can be combined based on different labels and contexts (from joints to lithologies) for more complex analyses. Furthermore, since the analysis is based on digital drill core photos, a digital collaboration of several geologists is easy. This results again in several ML models that can be combined for better objectivity.

The current work focuses on the first results of the HITL approach using the open-source labelling software CVAT and ML Model Detecron2.