Automated Analysis of Borehole Core Imagery from Oman Drilling Project Hole GT2A

Jin Wei Lim[1], Peter Hopkinson[2], Michelle Harris[4], Blair Thornton[1], Adam Prugel-Bennett[3], *Damon A H Teagle[5], Phase 1 Science Party Oman Drilling Project

1. Engineering and the Environment, University of Southampton, SO16 7QF, Southampton, UK, 2. Mathematical Sciences, University of Southampton, Southampton, SO17 1BJ, Southampton, UK, 3. Electronics and Computer Science, University of Southampton, SO17 1BJ, Southampton, UK, 4. School of Geography, Earth and Environmental Sciences, Plymouth University, PL4 8AA, Plymouth, UK, 5. Ocean & Earth Science, National Oceanography Centre Southampton, University of Southampton, SO14-3ZH, Southampton, UK

Rapid and reliable observations and consequent data interpretations are essential to achieving greatest knowledge return from scientific drilling. However, most scientific drilling campaigns to date rely on manual approaches to describe cores directly or analyze and annotate imagery scans of cores extracted from boreholes. Such approaches are extremely time consuming, labour intensive, and even with highly trained observers, prone to error. Recent advancements in machine learning enable new approaches of using automated tools that can rapidly analyze visual data with well characterized levels of accuracy. One of the most successful methods available for recognizing complex patterns in data with reasonable performance, which matches or even exceeds human performance in some domains, are Convolutional Neural Networks (CNN). A CNN is a computational model loosely inspired by the learning capabilities of a human brain, where layers of artificial neurons are connected to perform a set of mathematical tasks which can learn from large labeled datasets and is ultimately able to classify unlabeled data. Commonly, the architecture of a CNN consists of stacked convolution and pooling layers, but for this work a novel CNN architecture was developed, with parallel convolutions and image pre-processing that can achieve accurate training within a small amount of time.

Here, the application of such an approach is investigated for the automated analysis of core scan imagery of drilled rocks recovered from Oman Drilling Project Hole GT2A, which has a near 100% rate of core recovery. The rock and mineral types in image scans of cores obtained from the GT2A borehole were classified and quantified using the CNN developed. By disassembling the original core scans into half a million tiny images, they were individually fed into the CNN to be classified and reassembled back to the original position. With this new map data containing classified rock types and other features of geological interest, further analysis can be carried out efficiently by scientists such as generating density graphs or highlighting different rock types for visual inspection. The comparison of machine and human labels was used as a metric to estimate the accuracy of the automated analysis approach. This ensures that the work remains grounded and that results are understandable. Although the performance of this CNN on real world data is difficult to characterize due to absence of ground truth for the whole core scans, results from cross-validation (85-98% accuracy) of the labeled data indicates that this approach is promising and that it can reduce human effort without significantly compromising accuracy. However, this is a supervised learning approach where the performance heavily relies on the quality of training data. Often, the nature of images in new drill cores is unknown, and new unexpected features will be present which were not captured during training on prior dataset. This leads to poor classifications, requiring training data to be manually generated and the CNN re-trained, thus reducing the efficacy of the automated techniques in reducing human effort. This research demonstrates that CNNs have the potential to automatically analyze core imagery to produce quantitative data. However, further work is required to overcome limitations in training data generation to achieve full scalability.

Keywords: Automated core characterisation, Convolutional Neural Networks, Scientific Drilling

