Assessing the Dependability of Machine Learning Outcomes in Geosciences: Insights from Explainable Artificial Intelligence in Image Analysis

Korhan Ayranci[1], Isa E. Yildirim[1], Umair Waheed[2], Umut E. Yildirim[1], James MacEachern[3]

[1]King Fahd University of Petroleum & Minerals
[2]King Fahd University of Petroleum and Minerals
[3]Simon Fraser University

Abstract

Applications of machine learning in geology have received significant attention in the last decade. Given that visual observations play a vital role in geology (e.g., lithology identification and trace fossils, etc.) the role of automation using machine learning yielded numerous benefits for geoscientists and industry, such as reducing human error and increasing productivity. Although machine learning practices may offer highly accurate results, particularly in image analysis, it is still unclear how machine learning algorithms make certain decisions and provide “black box” answers. To increase the reliability of the results and level of confidence in predictions, it is crucial to add one more step to demonstrate how algorithms “think” and come up with certain decisions.

Perhaps the only step available today is the application of explainable artificial intelligence (EAI) where the end- user receives an output image along with a heat map pointing out the zones on which the algorithm mainly focuses when making the decision. This technique is relatively new and it has rarely been applied in geosciences. To test the application of EAI, a pre-trained model for detecting bioturbation intensity is used. This model uses core and outcrop images as input and provides whether they are (1) unbioturbated, (2) moderately bioturbated, and (3) intensely bioturbated as output. The overall accuracy of the pre-trained model reaches a minimum of 85%. The results of EAI in this pre-trained model show striking similarities with the human approach. In unbioturbated images, the heat maps are predominantly uniformly distributed throughout the images suggesting that the algorithm effectively focused on the preserved primary fabric. Interestingly, objects in some unbioturbated images (e.g., mud clasts) are characterized by low-degree heat maps indicating that the algorithm ignored those objects and was not confused with possible bioturbation. Moderately bioturbated images are perhaps the most challenging ones as these images include bioturbated and unbioturbated intervals. The heat maps suggest that the algorithm successfully recognized zones with preserved primary fabric intervals, as well as bioturbated intervals with both preserved individual trace fossils and disruption in the primary fabric. In intensely bioturbated images the heat maps are distributed uniformly throughout the image, similar to unbioturbated images, with a slight increase in concentered individual trace fossil zones. Identification of bioturbation intensity is a challenging task in geosciences and requires consideration of several parameters such as preserved primary fabric, laminae disruption, the presence of individual trace fossils, and more. The results in this study suggest that EAI can be applied in geology can provide valuable information and can increase the reliability of automated results. Furthermore, it can be used for educational purposes where certain expertise (e.g., ichnology) is not readily available in academia or industry.