Identifying Invertebrate Fossils in the Field Utilizing Machine Learning with a Novel 3D Augmentation Approach

Nichols, Christian (School: Brockton High School)

The study of invertebrate paleontology has been of great significance in the understanding of extant species and their evolutionary patterns. Paleontologists spend months in harsh conditions attempting to find fossils on their field expeditions and have not made significant technological advancements in several hundred years, hindering new scientific discoveries. Machine learning has proved to be accurate and reliable and can make the work of paleontologists more productive. In this project, automated searches of online public image repositories yielded more than 2,500 photos of 5 classes of invertebrates. TensorFlow 2 was used along with Object Detection API, OpenCV, and Blender-generated 3D data augmentation to simulate an aerial view of a drone. My hypothesis is if a TensorFlow-based neural network model was applied to the process of surveying large-scale invertebrate sites using ODAPI/localization classification in a near-real-time deployed system, then this approach could be used to survey potential and existing sites more efficiently. A Keras-based classification mode was run with an accuracy of over 85% with 200 iterations. I then trained a ResNet-fine-tuned localization model, with custom orientation, alpha and scaling data augmentation. This enhanced the quality of the dataset, with a loss of 0.007 over 170,000 iterations. The data supported my hypothesis. With unobstructed images, the model demonstrated 95% accuracy for general-classification and 91% accuracy for localization. These results outperform recently published models with similar approaches. This novel approach has the potential to contribute greatly to a more efficient, AI-enhanced future for paleontology, and may have broad crossover potential for other geoscience fields.