Contents

1 Abstract

Digitizing and uniformizing the structure of handwritten fossil records exhibits a great potential for increasing the accuracy of paleontological data analysis by increasing sample sizes. Approximately 90, 000 of such samples reside in the archives of the National Museum of Kenya, and an ongoing effort is to store this data in a digital format for better accessibility. A previous project utilized a commercial optical character recognition service for automated reading of these catalogues. This generalist handwriting detection model lacked the ability to detect special characters used to denote tooth samples, and could not utilize prior knowledge of the vocabulary that is more likely to be present in the data, leading to loss of information and detection mistakes.

This thesis aims to build a specialist character recognition model to increase the accuracy of the bone or tooth type specifying column of the digitized data by fine-tuning a state-of-the-art optical character recognition model with few-shot transfer learning. This is performed by first finding most accurate recognition models, variants of convolutional neural networks or vision transformers, and most successful transfer learning methods for adapting a model to a new character set. Then, the character recognition accuracy of combinations of these methods are benchmarked using handlabeled image segments from the fossil catalogues. The final aim of this work is to use the best-performing model to obtain an accurate reading of the catalogues of the National Museum of Kenya, and publish the final model to be used by the paleontological community for further digitization efforts.

Keywords: Optical character recognition, Few-shot transfer learning, Vision transformers, Paleontological databases