



## **USE OF MACHINE LEARNING TO AUTOMATE THE LABELING OF OUTCROP IMAGES FROM THE LUSITANIAN BASIN - PORTUGAL- PT**

Alex Willian Bessane, graduation student, Federal University of Pampa, Campus  
Caçapava do Sul

Dr. Ítalo Gomes Gonçalves, teacher, Federal University of Pampa, Dr Felipe  
Guadagnin, University of Pampa

e-mail: - [alexnessane.aluno@unipampa.edu.br](mailto:alexnessane.aluno@unipampa.edu.br)

Imagery represents a large and efficient source of information in the exploration and prospecting branches of geological research, both academic and corporate. In order to always obtain new and better images, the improvement of imaging techniques has become increasingly sought after and widely applied. In this, the acquisition of images from remotely piloted aircraft (RPAs) has gained great prominence due to its cost-benefit, reach and practicality, so it can be said that the generation of virtual models of outcrops (VMOs) can be considered as the current state of the art, because they can safely exercise the issues listed. Taking into account the economic and academic importance in the identification and correct classification, to better analyze them, several machine learning (ML) algorithms are used, with the purpose of extracting data such as mineralogy, porosity, permeability, among several other information minimizing human bias, reducing interpretation errors. However, the use of machine learning requires some relevant factors, among them trained professionals, computational capacity, and an extensive database about the target of interest, as well as previous knowledge about the object to help build the algorithm. Taking this problem into consideration, the optimization of time in the process of identification and categorization of the rocks in outcroppings means directly saving money. The main objective of this work is the creation of an VMO and the elaboration of an ML algorithm that systematically labels the point cloud used for the construction of the VMO of the Lusitanica Basin, Portugal-PT, making use of 10% to 20% of the database for training with the support of the pixels of the photographs, reducing the processing time and computational capacity, automating the lithological classification of the photos. The methodology applied for the realization of this work was divided into two stages: generation of an VMO of the Lusitanian Basin outcrop and, in parallel, the creation of the ML algorithm for the labeling of the point cloud. Initially a bibliographic survey of the region and the outcrop was done. Then photographs were taken of the target outcrop using a Phantom IV model RPA, then the images were imported into the program Agisoft Photoscan where in turn were removed photos that were of low quality or off the target of interest. The alignment of the images was performed according to the following parameters: "high" accuracy, i.e. the most accurate estimates in the positioning of the cameras; pre-selection of "reference" pairs (correspondence of features detected in the photos), selecting a subset of pairs of images to be matched; upper limit of points and feature in each image to be considered during processing; limit of tie points (upper limit of corresponding points for each image). Finally alignment optimization was performed. The result of this process is a sparse point cloud. With the sparse cloud it was possible to perform the

georeferencing of the project, therefore the generation of the dense point cloud was performed, then the triangulated mesh and texturing, this had the basic parameters "generic" mapping (determines how the texture of the object will be compacted in the texture atlas); "mosaic" blending mode (selects the way the pixel values from different photos will be combined in the final texture) thus finalizing the creation of the VMO. The Google Colaboratory platform was used for the algorithm creation, this being a Jupyter notebook environment that requires no configuration and runs in the cloud. The language chosen for the algorithm creation was Python, an Open-Source general-purpose language commonly used in data science and ML. Based on this, we proceeded to create image segmentation models using discrete convolutional filters and masks ensuring the correct identification of each lithology. Later, neural networks were used to classify and finally, several configurations using different models were tested. As a partial result of this work, the VMO of the Lusitanian Basin outcrop was built with a high degree of resolution compatible with the real model. At the end of the second stage of this work, it is expected to create an algorithm that allows the automation of the labeling of the point cloud generated by the VAM, classifying the outcrop in a fast and efficient way. Finally, we conclude that the interaction of VMOs and ML can become a new state of the art in geosciences due to the time optimization in rock identification and classification. Another advantage is the portability in which, once the algorithm is done, following the same methodology it is feasible to apply it to other outcrops and regions of interest.

**Acknowledgments:** CNPq, FAPERGS, UNIPAMPA.

**Keywords:** Geology; Remotely Piloted Aircraft; Machine Learning; Labeling; Point Cloud.