Analysis of Methodologies for Creating Mars Topography Classifier

Michael Porter
Department of Mathematics and Computer Science
Alabama A&M University
Normal, AL
mporter9@bulldogs.aamu.edu

Dr. Kenneth Sartor
Department of Mathematics
Alabama A&M University
Normal, AL
kenneth.sartor@aamu.edu

Abstract

The purpose of this research effort is to investigate mathematical techniques that can be related to creating a Mars topography classifier. There are many different methodologies to choose between when creating a process that can be used to classify Mars topography. For example, numerical regression techniques excel at fitting data trends while matrix similarity heuristics can be used to classify more complex objects based on their features. Machine learning uses structures such as neural networks or decision trees to create the most flexible classification method. Since machine learning methods can be complex and can require large amounts of data, this methodology is generally reserved for applications where standard modeling methods are not well-suited.

Introduction

A Mars topographic classifier is a process that takes Mars surface elevation data as input and maps it to one of a distinct number of topographic types. Astronomers have already identified and labeled many of the significant topographic features of Mars. However, a classifier could be useful for identifying smaller-scale features and could be extended to other planets should sufficient data exist. Definition of a Mars classifier as a function *f* is as follows:

$$S = \{\{\text{Mons}\}, \{\text{Crater}\}, \{\text{Planum}\}, \dots\}$$
$$f : X \to Y$$
$$Y \subset S$$
$$X = \text{Elevation data}$$

Methodology

The elevation data for this study comes from NASA's MOLA dataset, approximately 2GB in size. It is encoded as a rasterized greyscale image, with a pure black pixel representing the lowest point on Mars (approximately -8,200m) and a pure white pixel representing the highest point (approximately 21,000m) [1]. The pixel color can then be converted to the actual elevation in meters using a formula:

$$H(x,y) = Color(x,y) * P_h + P_l$$
$$P_h = \text{Highest elevation}$$
$$P_l = \text{Lowest elevation}$$

Where Color(x, y) gives the greyscale color value ranging from 0 to 1 for the corresponding pixel.

Different cross-sections of Mars topography can then be extracted by recording the elevation of each pixel across a line from the original image. These different cross-sections can then be recorded for various types of known Mars topography and used as a reference for a classifier.

Definition 1. Archetype Function

$$A_i(x) = mean(C_1(x), C_2(x), C_3(x), \dots, C_j(x))$$
$$C_j(x) = \text{corresponding y-value for the j-th cross-section}$$

Definition 2. Classification Function

$$L = \min(H_1, H_2, H_3, \dots, H_n)$$
$$n = \text{number of archetypes}$$
$$H_i = mean(|A_i(x) - C(x)|) : 0 \le x \le 1$$
$$H_i = \text{error function between archetype and cross section}$$
$$C(x) = \text{function representing unknown cross-section}$$

Figure 1. MOLA dataset encoded as greyscale image showing elevation of Mars [1].

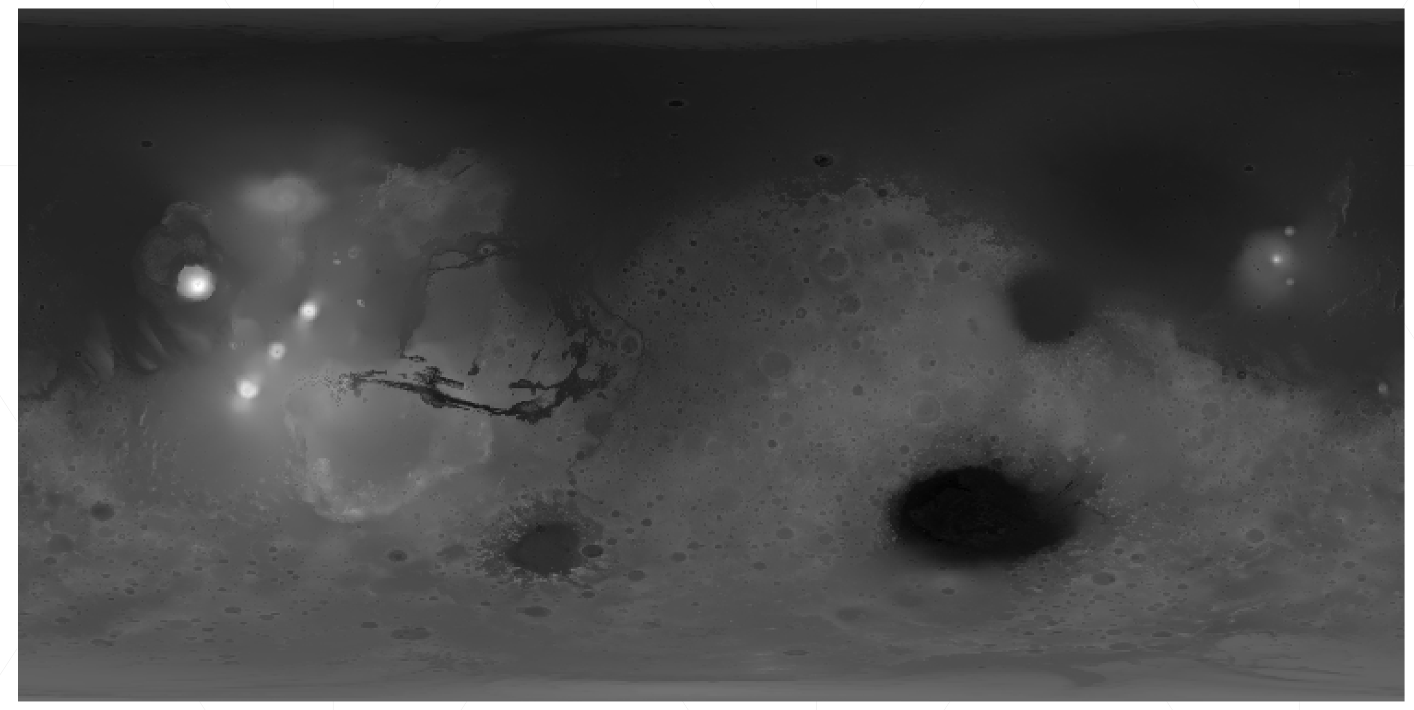


Figure 2. Longitudinal cross-section of Olympus Mons extracted from above image.

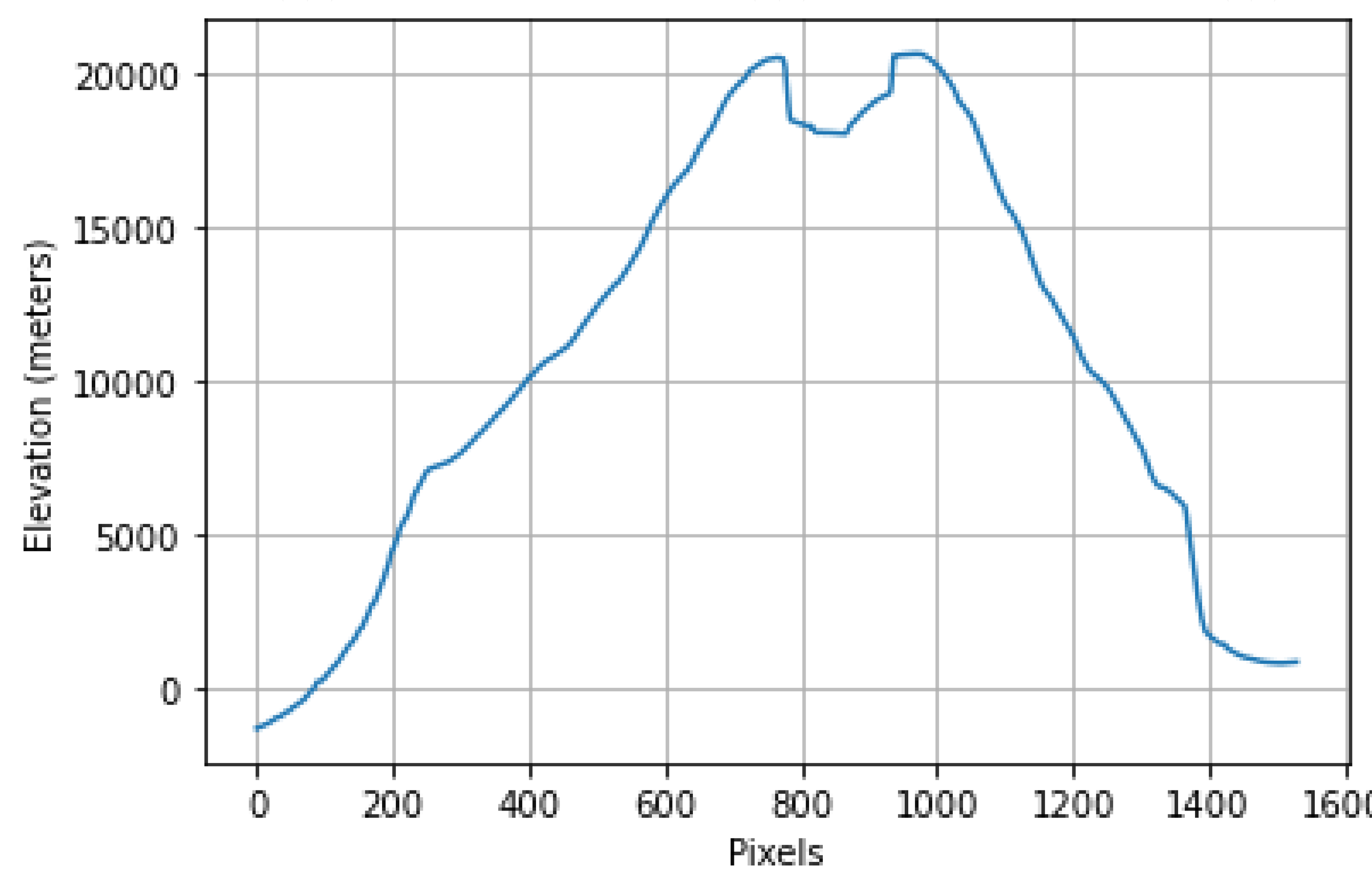
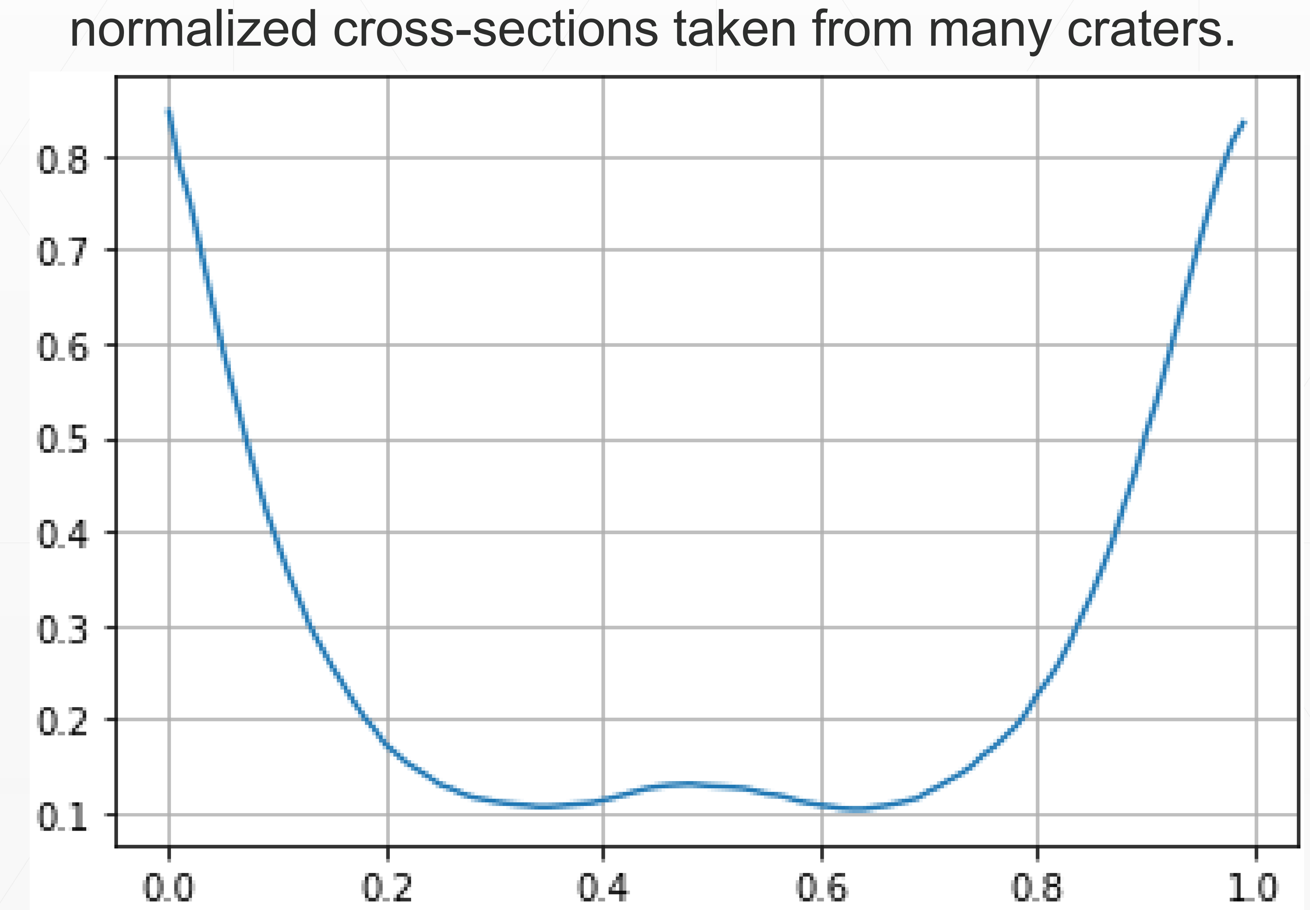


Figure 3. Archetype of crater feature generated from various normalized cross-sections taken from many craters.

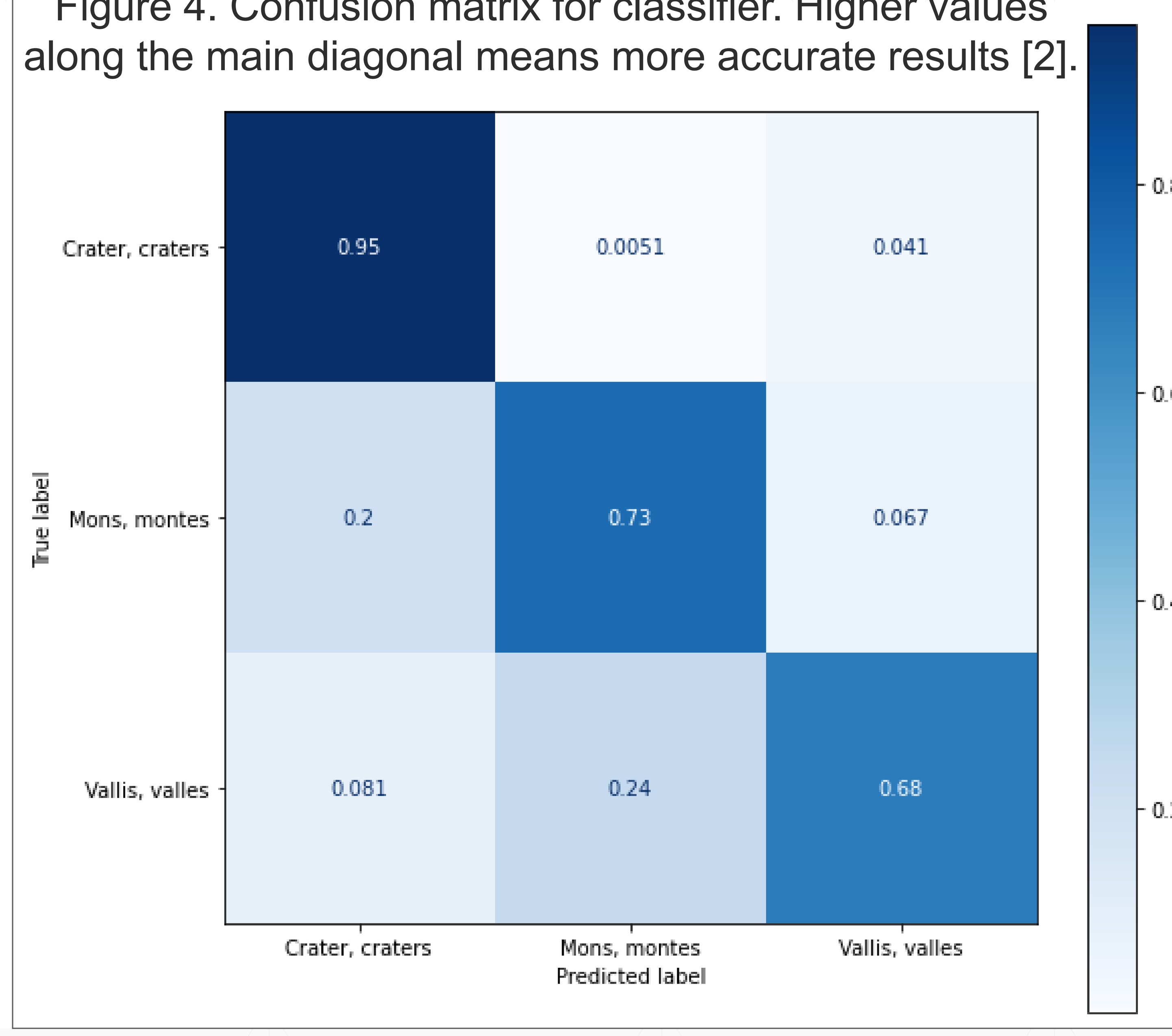


Results & Discussion

A classifier was built using the archetype and classification definitions 1 and 2 for the crater, mons, and vallis feature types. Archetypes were generated from 1,005 different examples of the three features. Afterwards, accuracy of the classifier was tested by comparing the predicted class and the actual class of an additional 248 unseen features. Accuracy can be visualized using the confusion matrix in Figure 4. Craters were correctly identified with an accuracy of 95%, 73% for mons, and 68% for vallis.

Since the cross-sections used for the classifiers are 2-dimensional, this technique of classification will generally work better for symmetrical feature types such as mons or craters. For asymmetrical feature types such as a canyon or ridge, orientation is more important so it may be necessary to use three dimensions or use an entirely different approach. The archetype-based classification suffers from reduced accuracy when trying to distinguish between two very similar feature types, such as between mons and tholus.

Figure 4. Confusion matrix for classifier. Higher values along the main diagonal means more accurate results [2].



Conclusion

This research has investigated how a Mars topographic classifier can be constructed using a novel application of mathematical techniques. The results have demonstrated the efficacy of the classifier as well as discussing its successes and limitations.

References

[1] MOLA Team, "Mars MGS Mola Dem 463M V2: USGS Astrogeology Science Center," *Mars MGS MOLA DEM 463m v2*. [Online]. Available: https://astrogeology.usgs.gov/search/details/Mars/GlobalSurveyor/MOLA/Mars_MGS_MOLA_DEM_mosaic_global_463m/cub. [Accessed: 23-Mar-2022].

[2] "Confusion matrix," *scikit-learn*. [Online]. Available: https://scikit-learn.org/stable/auto_examples/model_selection/plot_confusion_matrix.html. [Accessed: 23-Mar-2022].