NOAH-H, a deep-learning, terrain classification system for HiRISE: Results for Oxia Planum and Mawrth Vallis.

Alexander M. Barrett¹, Matthew R. Balme¹, Mark Woods², Spyros Karachalios², Danilo Petrocelli², Elliot Sefton-Nash³, and Luc Joudrier³

¹Open University, Walton Hall, Milton Keynes, MK7 6AA, UK
²SCISYS Ltd, Methuen Park, Chippenham, SN14 0GB, UK
³ESTEC, European Space Agency, Noordwijk, The Netherlands.

Planetary remote sensing (RS) missions are returning an ever increasing volume of data from across the Solar System. This wealth of data, much of it with a high spatial resolution, presents major challenges as well as opportunities. It is becoming increasingly difficult to fully interrogate large RS datasets such as the High Resolution Imaging Science Experiment (HiRISE) images of Mars (McEwen et al. 2010). The time required to survey all relevant images at full resolution can be daunting for all but the largest teams.

Advances in machine learning provide a way to overcome these challenges, by automating the initial surveying of planetary RS data, and providing a more accessible dataset, which highlights textural features of interest to the project.

The Novelty or Anomaly Hunter – HiRISE employs a deep learning convolutional neural network (DNN) to classify HiRISE images (LeCun et al. 2015; Simonyan and Zisserman 2015; He et al. 2016). A set of ontological classes was designed, which covered the complete range of textures at the site. These consisted of broad “terrain types” rather than formal geomorphological units. The aim of the project was to classify the Exo-Mars Rosalind Franklin Rover (Vago et al. 2017) landing site, and identify features such as aeolian bedforms or blockfields which might present localised hazards to rover operations (e.g. Rothrock et al. 2016). This would provide a useful input to traversability analysis. The focus was thus on detection, rather than digitisation. Producing a formal geomorphological map was beyond the scope of the project.

Four broad categories of classes were selected; non-bedrock surfaces, bedrock surfaces, aeolian bedforms, and boulder fields.

The surface classes were subdivided in terms of their metre scale relief and apparent roughness upon visual inspection. Bedrock classes exhibited clearly defined texture and relief, suggestive of outcrops, while non-bedrock was interpreted to consist of regolith or loose materials. Both were further subdivided according to the degree to texture present.

The bedform classes were distinguished from the non bedrock surfaces by the presence of clear aeolian ripple forms (Balme et al. 2008; Balme et al. 2017). They were subdivided based upon both the scale of the bedforms, and whether they were continuous or discontinuous. Large isolated ripples were labelled individually, but this was not practical for large fields of smaller discontinuous
ripples. These were thus classified based on whether they overlaid bedrock, or non bedrock surfaces. A very small number of sites within the study area also exhibit rectilinear ripples. These were only found on a large scale.

Finally boulder patches consist of block fields, and regions of boulder strewn ground. Individual blocks were too small to label, so patches of boulder covered ground were classified.

These classes were used to manually label a set of ~1500 training images, each being a small 128-128m “framelet” extracted from the larger HiRISE image. From these examples, the DNN learned to classify the entire site according to the prescribed classification scheme. The model output consisted of a classified raster image, of the same dimensions as the original HiRISE image. This was colour coded and overlain on the HiRISE images for further analysis.

NOAH-H performed very well when identifying the very distinct classes such as bedforms, boulder fields and areas of fractured ground. Distinguishing between surface classes proved less reliable. This is likely due to the fact that many of these classes form a continuous variation, and so cannot be divided into discrete types with 100% reliability. When similar classes are grouped, and all bedrock or non bedrock terrains are considered together, the reliability of the model increased dramatically.

The majority of confusion occurred within these broader groups, rather than between them. The final run of the model produced a mean Intersection over Union (IoU) of 74.15% for the full class list and 92.33% for the grouped classes.

A set of sample locations were also studied to determine how the geomorphology was represented in the output data. This analysis broadly supported the results of the IoU analysis. The pixel scale results were not always found to be a perfect match, due to subtle variations within and between classes. The model sometimes struggled with “fuzzy” boundaries between regions of contrasting terrains.

However, it was found that even in cases where some individual pixels were misclassified, the classification of the area as a whole was frequently still both useful and reliable. While only 53% of sampled locations were found to be correctly classified at the pixel scale, 72% were correct when the landscape of the area as a whole was considered. When classes were combined into groups, this increased to 88% of the sampled locations.

The model results are thus most useful when considered at the “landscape scale”. It provides a very reliable guide to the distribution of terrains within an area and will provide a valuable tool for geomorphological study. It is already being applied to detection of aeolian hazards at the Oxia Planum site (See: EPSC2020-572).

References


