CLASSIFYING PLANETARY SURFACES USING MACHINE LEARNING. A. M. Barrett[1], M. R. Balme[1], J. Wright[1,2], M.J. Woods[3], S. Karachalios, M.T. Malinowski, [1] School of Physical Sciences, The Open University, Walton Hall, Milton Keynes, MK7 6AA UK alexander.barrett@open.ac.uk; [2] European Space Agency (ESA), European Space Astronomy Centre (ESAC), Spain; [3] SCISYS Ltd.

Introduction: Deep learning (e.g. [1]) convolutional neural networks were trained to discriminate meter scale variations in surface texture in satellite images of planetary surfaces [2]. Two versions of the model were trained, one classifying surface textures and aeolian bedforms in HiRISE images of Mars [3]. The second to discriminate between blocky ejecta and textures indicative of impact melt in LROC-NAC images of the Moon [4].

An ever increasing volume of remote sensing data is being returned from the Moon and Mars, and while this has the potential to allow large scale mapping efforts at a greater spatial resolution than ever before, it is increasingly challenging for all relevant data to be surveyed in a reasonable amount of time.

The networks classify surface textures based on morphological criteria rather than making determinations of perceived geological origin. Automating the production of a geological map is not yet possible and would be counter-productive, since the purpose of a mapping effort is as much to build understanding of the geological history of a site through discussion and exploration as it is to classify the surface. Rather our aim is to use machine learning to augment the human mapping workflow, and speed up the initial surveying needed for such an effort. A network performs “triage” on unmanageably large datasets, indicating to a human operator areas where predefined surface texture assemblages indicate that features of interest could be present.

NOAH-H: The Novelty or Anomaly Hunter – HiRiSE (NOAH-H) [5] was developed as part of the ExoMars *Rosalind Franklin* [6] landing site selection process. It was trained on ~1500 example framelets, selected from across Arabia Terra. It performs semantic segmentation at a pixel scale, identifying surface textures as one of 14 ontological classes (fig 1). Seven surface classes define roughness types, six classes describe aeolian bedforms and the final class describes patches of boulders.

The classification scheme is hierarchical. Care was taken to ensure that the class definitions were purely morphological, providing a robust descriptive level. These descriptive classes were then grouped into thematic categories such as bedrock, non-bedrock etc. to form the interpretive layer of the system. This ensures that the classification does not require contextual evidence which would be unknown to the network. While still providing the scientist the information needed for interpretation.

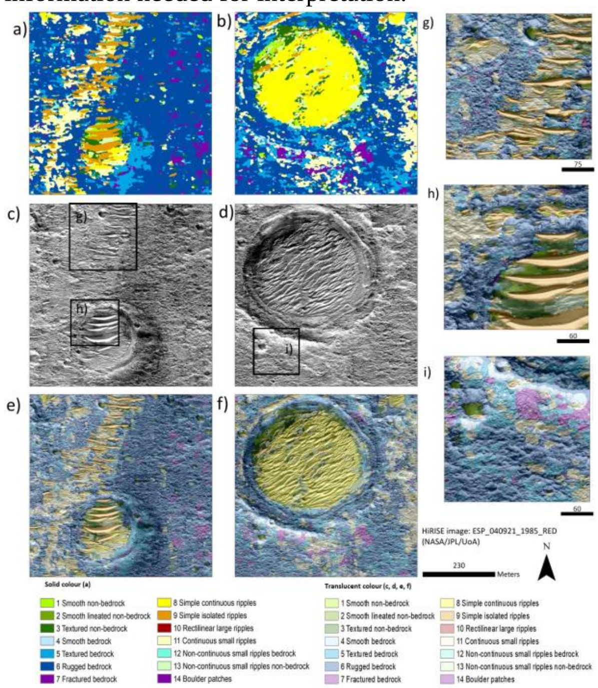

Fig 1: HiRISE image classified using NOAH-H. a) and b) Classified raster, c) and d) original HiRISE image, e) and f) translucent classification overlain on HiRISE image, g) aeolian bedforms over rugged bedrock, h) large and small bedforms, i) identification of fractures. Areas of different surface texture are indicated, though the boundaries might need further refinement by a human mapper. In the case of dispersed features such as the large aeolian bedforms, the results were found to reliably segment the features.

Fig. 1 shows an example where aeolian bedforms are reliably segmented from surfaces of various classes. The model is also good at identifying regions of fractured ground. Different roughness classes are used to distinguish smoother (most likely non-bedrock) terrains, from rougher ones (interpreted as bedrock).

Despite being trained on images from Arabia Terra, the NOAH-H system proved transferable to other areas of Mars where a similar suite of surface textures predominate. Images of the NASA Mars 2020 [7] landing site at Jezero Crater were successfully classified. This allowed for comparison with in-situ images from the Perseverance Rover and Ingenuity Helicopter. This analysis corroborated the results of the machine learning classification, suggesting that the texture classes observable in orbital HiRISE images correspond well with features on the ground [8]. Once the ExoMars rover lands at Oxia Planum, similar comparisons will be made for the training site.

Classification of images from Jezero also allowed for a direct comparison between our classification and human made geomorphological maps [9], since the area was well studied prior to landing. When our classes were grouped appropriately, a good correspondence was found between the two products.

Our results suggest that the NOAH-H system is a very useful tool for highlighting the textural changes which planetary geologists look for when creating a geomorphological map. NOAH-H can be used to identify regions of different classes at the pixel-scale, which can then be more generally characterized or summarized by a human mapper.

For some landforms, in particular the larger aeolian bedforms, NOAH-H was found to be capable of very precisely segmenting individual features, producing data which could form part of a manual mapping effort with little further human refinement. These bedform maps are now being used to constrain the aeolian history of Oxia Planum [10].

NOAH-L: NOAH for LROC-NAC is presently being developed. This extension of the project aims to test a more complex test case, by training the network to distinguish between impact melt and dry blocky ejecta around young lunar craters. This involves considerations not present for the martian case, such as the effect of solar incidence angle, which is much more variable in LROC-NAC images than in HiRISE.

Eleven ontological classes were used, covering six melt-related textures and a further five counter examples. Two classes represented areas covered by small boulders, and larger, isolated "megablocks", with dry granular flows forming the final class.

The aim was for each class to be morphologically distinct. However the more focused research question means that more contextual information may be required to distinguish melt related landforms from morphologically similar features produced through other processes. Each "melt" class was defined as a component of a broader impact melt assemblages. Thus a fractured terrain in proximity to flow features can be identified as being a potential indicator of melt, while fractures elsewhere in the image can potentially be disregarded. Work is underway to test this approach.

Discussion: In addition to the scientific outputs of this project, these case studies have provided insight around framing geomorphological questions in a way which will be approachable with a semantic segmentation methodology. Care must be taken when conceptualizing a research question, with the capabilities of the system in mind.

All geomorphology is a mixture of description and interpretation. The former can be approached as a semantic segmentation task, whereas the latter requires additional situational and contextual evidence which cannot be incorporated into a network via training on labelled frames. This limits the extent to which a machine learning system can (currently) replicate the human workflow. Its ability to recognize and identify features is impressive, however any discrimination which relies on broader context, or an understanding of specific geological processes remains challenging.

For example, when trained to identify the lobate margins of impact melt flows, the networks recognize the features, and perform well at locating more. However, they also detect other "arcuate edges", which occur at many scales and through many processes. The system does not appear to spontaneously assimilate the contextual evidence that a human would use to pick out a lobate flow front.

This can be solved in post processing by looking for assemblages of landforms, and only selecting those arcuate edges which are found in proximity to other melt indicative features. However, equifinality remains a major challenge, as "false friends" for these other features exist as well.

Where the system works best is when classifying structures like boulder fields, and aeolian bedforms. These are clear discrete features, and the boundaries between these classes and other are not "fuzzy" since few comparable features emerge by other processes.

Acknowledgments: We acknowledge funding from the European Space Agency, and the UK Science and Technologies Facility Council (Grant ST/T000228/1). We thank the HiRISE and LROC teams for providing us with so many great images to classify.

References: [1] LeCun et al. (2015) Nature 521 436-444 [2] Woods et al. 2020 I-SAIRAS Virtual Conference [3] McEwen et al. (2010) Icarus 205, 2–37. [4] Robinson et al. (2010) *Space Science Reviews, 150*, 81-124. [5] Barrett et al. (2022) Icarus 371, 114701. [6] Vago et al. (2017) Astrobiology 17, 471–510, [7] Farley et al. (2020) Space Sci. Rev. 216 [8] Wright et al., (in preparation) [9] Stack et al. (2020) Space Sci. Rev. 216 [10] Favaro et al. 2021 J. Geophys. Res. Planets 126.