AUTOMATIC CRATER DETECTION OVER THE JEZERO CRATER AREA FROM HiRISE IMAGERY:

A. Lagain[1], K. Servis[1,2], G.K. Benedix[1], D. Flannery[3], C. Norman[1], M. Towner[1]. J. Paxman[4], [1]Space Science and Technology Centre, School of Earth and Planetary Sciences, Curtin University, Perth, WA, Australia (anthony.lagain@curtin.edu.au), [2]CSIRO Resources Research Centre (ARRC) Kensington WA 6151, Australia. [3]School of Earth, Environmental and Biological Sciences, Queensland University of Technology, GPO Box 2434, Brisbane, QLD,Australia, [4]School of Civil and Mechanical Engineering, Curtin University, Perth, WA, Australia.

Introduction: Impact craters are used to determine the ages of planetary surfaces. Absolute dating of meteorites or *in situ* geochronology provide a few essential reference points, but these techniques are rare and not yet applicable at the planetary scale. Therefore, impact crater counting techniques will remain the major tool to decipher planetary surface history. This approach requires a tedious mapping and morphological inspection of a large number of circular features to distinguish true and primary impact craters from other surface features and secondary impact craters; in particular on Mars whose the surface exhibits a large variety of pseudo-circular features (e.g mounds, collapse pits of lava tubes, circular grabens, glacial cirques, calderas, etc…). The most complete database of Martian craters [1] includes a catalog of more than 384,000 impact structures larger than 1 km in diameter. This database is considered to be complete for this diameter range. A requirement to determine young surface ages on Mars must include smaller impact craters, typically a hundred meters in diameter, found on the area of interest [2]. Because crater number scales as a power law, the number of impact craters <1km across the entire surface of Mars could reach the millions, making manual analysis of local variations, over the entire surface, impossible without some kind of automation.

Automatic crater detection technique: Previously, we described our crater detection algorithm (CDA) [3-7] with subsequent improvements with ongoing advances in machine learning. Our CDA was initially trained by selecting 889 tiled THEMIS Day IR images where 1,762 impact craters from the Robbins database [1] have been identified and manually cleaned to select the most identifiable impact craters. Applied to the THEMIS mosaic between 45 degrees of North and South, our algorithm produce a true positive detection rate of 86% for craters larger than 1km in diameter [7]. We applied our algorithm to the CTX global mosaic [8] in order to access to the population of craters between 100m and 1 km in diameter [9]. Using the results, we demonstrated the ability of our CDA to reproduce model ages previously derived from manual counting with a precision of up to 5% on the age [6,7]. The ultimate goal of our work is now to automatically compile smaller impact craters (5m<D<100m) visible on HiRISE imagery dataset offering a resolution of 25cm/px, as well as create a user friendly software which can be used by anyone in the future. The high degree of detail found in this dataset makes our CDA inefficient due to the presence of decametre pseudo-circular structures on the surface as erosion or other particular features within field dunes. A retraining of the algorithm is therefore essential for an accurate detection of m-sized impact craters. There have been two significant tasks for retraining, namely iteration speed of the training-evaluation cycle for the model and the lack of a labelled dataset (training library) for HiRISE.

Improvement to the CDA: To address the first challenge, we decided to update the CDA in order to achieve a higher modularity of the components, but at the same time make the software more user friendly and able to be run both on a local machine for development but also on local HPC clusters at the Pawsey suprercomputer centre in Western Australia. To that end, we broke down all the processing steps to individual tasks, all with their own container that can be run individually and in parallel for different images, each of which goes through all the individual steps producing crater locations in geographical coordinates of the original GeoTIFF image (using the spatial reference system specified therein). When all the individual images have been processed, non-max suppression is performed on the union of the results in order to remove potential doublets and to obtain the final data product. In addition by using Ultralytics we have addressed both the issue of slow iterations and the shortage of labelled data. As the code has been written in Python it is much easier to make changes and diagnose issues that it is with Darknet which was used in the previous version of the CDA, which more than makes up for any slowdown. The software includes an augmentation on read which performs a number of affine transformations which we use (rotate, shear, scale, translate) in order to artificially augment out manually labelled dataset.

Training dataset: We trained our algorithm on a part of the HiRISE mosaic built by [8] covering a part of the Jezero crater (E77-5_N18_0) where 1650 craters have been manually identified. A portion of this population of craters has then be selected in order to be sure to include the most confident impact features in the training dataset, finally resulting to 1624 craters over this entire image. This labelled dataset is also augment-

ed by applying the range of transformations as described in the previous section.

Detection results over the Jezero crater: Our model has been applied over the entire HiRISE mosaic covering the Jezero crater where more than 27,298 craters have been detected [8]. Figure 1 illustrates the crater identification at different spatial scales (see caption for details). The bottom figure show the results of our model on the Jezero delta area. More than 1000 craters larger than 3m in diameter have been identified by the algorithm over an area of 12km x 9km where only 29 larger than 100m have been identified over the same area by using the previous version of the CDA on the CTX mosaic [7,8]. In order to validate our results, we compared the detection obtained on 30 tiles of 960px x 960px randomly chosen on a part of the mosaic (E77-25_N18-25) which have not been included into the training dataset with a manual identification, thus constituting the ground truth. For this purpose, we decided to categorize each tile according to the type of terrain mostly represented on each of them: rocky terrain, smooth terrain and dunes fields. We have also specified when the image exhibited some vertical stripes leading to the fourth category. On rocky and smooth terrains, the CDA produce very good results : only 5% of detection on the average are false detection and 16% of craters on average have not been detected by the CDA. However, the CDA is less efficient on dune fields since 35% of detection are false detection and 15% of craters have not been identified. Finally images exhibiting some vertical stripes significantly decrease the detection rate of the CDA since 56% of detection are false negative and 20% of craters have not been detected.

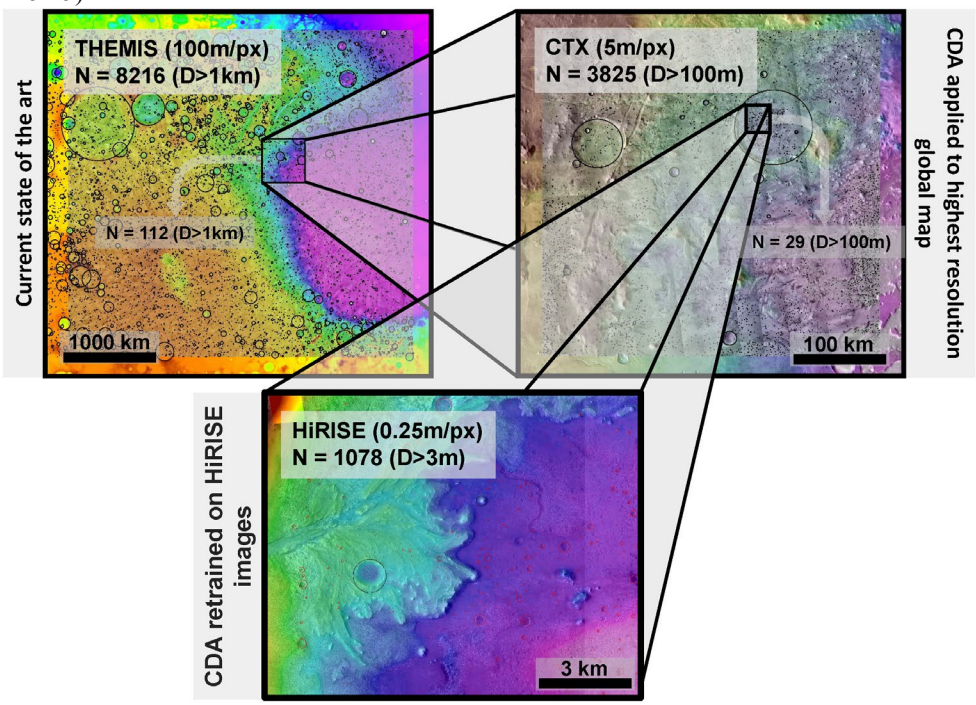

Figure 1: *Top left: Impact craters >1km identified by [1] on THEMIS imagery over the Syrtis Major region, Top right: Results of the CDA trained on THEMIS imagery using the manual crater database developed by [1] and applied on the CTX global mosaic [8] around the Jezero crater, Bottom: Results of the CDA retrained on the HiRISE mosaic [8] covering the Jezero crater. A part of the detection visible on the area shown here has been used in the validation dataset.*

Conclusion and future works: This version of the CDA produces an acceptable detection rate on HiRISE images where dunes fields are not present on the surface. The next step of our work will consist in using these detection over the Jezero crater in order to date the variation of its surface and validate the results against manual counting.

References: [1] Robbins, S. and Hynek, B. (2012) *JGR, 117*, E05004. [2] Quantin, C. et al. (2004) *Icarus Vol. 172*. [3] Norman, C.J. et al. (2018) *Planetary Science Information and Data Analytics Conference*, St. Louis, MO, Abstr #6002. [4] Benedix, G.K., et al. (2018) *49th LPSC*, #2202. [5] Benedix, G.K., et al. (2019) *50th LPSC*, #2140. [6] Lagain, A. et al., (2019) *9th International Conference on Mars*, #6017. [7] Benedix, G.K. et al. submitted to *Earth and Space Science* (19 Nov. 2019). [8] NASA/JPL/ MSSS/The Murray Lab. [9] Benedix et al., (2020) 51st LPSC.

Acknowledgement: This work is supported by the resources provided by the Pawsey Supercomputing Centre with funding from the Australian Government, Curtin University, and the Government of Western Australia.