USING MACHINE LEARNING TO COMPLEMENT NEW MARTIAN CRATER INVENTORIES.

Annabelle Gao[1], Ingrid J. Daubar[1], Daniel Wexler[1], Kiri L. Wagstaff[2], Valentin Bickel[2, 3, 4], Gary Doran[2].
[1]Brown University, Providence, RI (Annabelle_Gao@brown.edu). [2]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA. [3]MPS Goettingen, Goettingen, Germany. [4]ETH Zurich, Zurich, Switzerland

Introduction: The discovery and categorization of fresh impact craters on Mars in the last few decades [1] has granted insight into surface processes (e.g. [2, 3]) and crater retention ages (e.g. [4-6]), as well as subsurface phenomena (e.g. [7, 8]).

The investigation of crater clusters, or a multitude of craters that formed simultaneously and are spatially close, provide insight into atmospheric processes and fracturing dynamics [9]. Observing the frequency and nature of fresh impacts can contribute towards dating Martian surfaces. Although there are not yet dateable samples from Mars, a crater chronology model can still be extrapolated from the Moon and calibrated with current impact distributions [4-6]. Occasionally, a new impact reveals subsurface water ice. These impacts can provide information about ice-table depths and the nature of clean ice on Mars [7, 10].

All of the characterized new impacts recorded previously were found by manually sorting through and comparing Context Camera (CTX) [11] images with other datasets. This process involves finding two consecutive images of the same area, identifying a new impact in the later image, and ensuring that the same impact was not present in the earlier image. This is time-consuming and tedious, and it is possible that many new impacts were missed as a result of human oversight.

A machine learning algorithm has the potential to streamline the process of finding new impacts, filtering image pairs to those most likely to contain new impacts. Using this new method of locating new craters, we can explore biases in the impact dataset by comparing impacts found by the machine learning algorithm to those obtained by the manual method and give a more holistic and accurate picture of the overall current Martian bombardment. To date, we have applied the algorithm to PDS-released CTX images within +/- 60° latitude and requested follow-up images from HiRISE [12], which are needed to measure crater diameters, characterize ice-exposing craters [10], and observe changing albedos of blast zones [2].

Methods: The machine learning model to detect fresh impact craters in individual CTX observations was trained by adapting an existing deep convolutional neural network (inception-v3) [13]. A new training set was constructed using 1858 manually identified examples of fresh impact craters, as well as 4973 examples of non-impact regions of Mars' surface.

After training the fresh impact model, it was deployed across the entire CTX archive, containing approximately 112,000 observations at the time. Each CTX observation was broken into square tiles and the model produced a confidence value between 0 and 1 indicating the likelihood that a fresh impact is present within each tile. On average, the full system running in parallel on a high-performance computing cluster took about 5 seconds to process each observation. Even with the manual review described below, this is a vast improvement on the 30-40 minutes that would be required for manual examination of each image.

We manually reviewed the top 1000 highest confidence candidates identified by the machine learning classifier for which at least one possible "before" (probability of fresh impact < 0.5) and a possible "after" frame (probability >= 0.5) exists in the CTX archive. For each candidate, we inspected all CTX observations of the same site location within the archive. The manual examination of the ML results was enabled by significant experience in studying fresh craters. We also checked for HiRISE images of the same location; HiRISE observations were already available for 228 of the 1000 candidates.

If a candidate was deemed likely to be a fresh impact, we then checked the existing database for the site. For candidates that were not found in the database, we submitted requests via HiWish, the HiRISE public suggestion tool [14], to obtain follow-up high-resolution observations of each of these new sites to enable detailed measurements and analysis.

Once HiRISE images were acquired, the crater diameters were measured using the three-point tool in the JMARS program [15]. For crater clusters, all craters with a diameter ≥ 0.75 meters (3 bin-1 HiRISE pixels) were recorded. Any impacts with a smaller diameter were not included due to low resolution and possible inaccuracies in measuring. HiRISE images not yet available in JMARS were examined using the distance tool in HiView (www.uahirise.org/hiview/).

Preliminary Results: We found that 708 (71%) of the top 1000 candidates were fresh impacts. Of these, 76 were newly discovered fresh impacts that had not been previously identified manually. Another 161 candidates matched sites that were already known and included in the fresh impacts catalog [16]. The largest candidate group contained 461 features that visually appear to be fresh impacts but lack a "before" image in the CTX catalog that would allow us to constrain the date of formation. However, since blast zones fade on the order of decades [2], these are almost certainly geologically recent impacts. Cross-matching these impacts with earlier observations by Mars Global Surveyor, Mars Odyssey, or the Viking orbiters could enable the inclusion of some of these sites as dateable fresh impacts as well. This is an area of future work.

In addition to fresh impacts, the classifier also identified 165 impacts that did not appear to be "fresh" (e.g., absence of fresh ejecta, size, and presence of aeolian bedforms) and 89 features that were not impacts at all. An additional 48 candidates were determined to be duplicates of other candidates in the top 1000, due to minor errors in registration between images (Fig. 1).

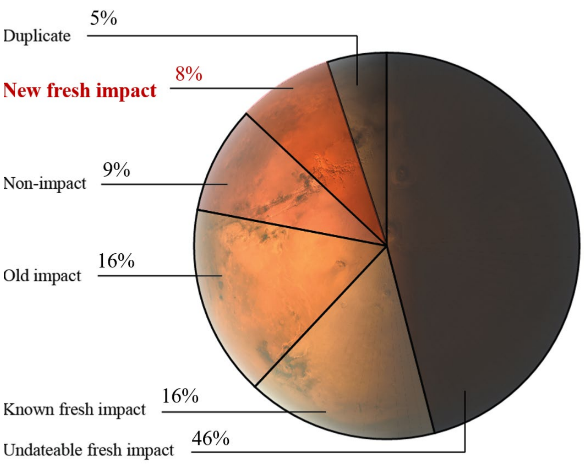

Fig. 1. *Results of manual evaluation of 1000 fresh impact candidates determined by the ML algorithm.*

Of the newly discovered 76 fresh impacts, 45 have been imaged with HiRISE thus far. When compared to the broader database of 1111 measured new impacts, the average diameter of craters found by machine learning was smaller, 3.9 ± 0.5 (N = 11) versus 6.9 ± 0.3 (N = 472) meters for singles and 6.0 ± 0.4 (N = 34) versus 6.7 ± 0.3 (N = 242) meters for clusters. The average diameter for clusters was calculated by taking the average of the effective diameters (($D_{eff} = (\sum_i D_i^3)^{1/3}$; [1-4]). Further, 75.6% of the 45 newly discovered impacts were clusters whereas only 57.5% of the original database fell into this category. Generally, the impacts found by machine learning displayed fewer rays (55.6% versus 62.5%) and more diffuse halos (93.3% versus 86.6%) (Fig. 2). The locations of impacts in the two datasets do not seem to differ much spatially, although more analysis on this has yet to be done. The 45 newly discovered impacts populate similar regions on Mars as the 1111 previously known impacts, areas which correlate well with thermal inertia and dust cover maps.

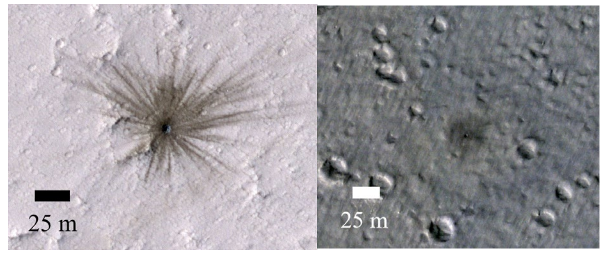

Fig. 2. *Comparison of a known new single impact currently in the database (left) and a machine learning new single impact (right). The known new crater is larger in diameter and has rays while the machine learning new crater has a diffuse halo.*

Preliminary Conclusions: These initial findings with a limited sample size (N = 45) suggest that the machine learning algorithm might be more adept at detecting clusters and smaller-diameter new impacts. This algorithm or a similar one may be useful in reducing the bias towards larger fresh impacts in the existing database and contribute to a more accurate size frequency distribution.

The differences that appear to exist between the populations can be attributed to a multitude of reasons, including biases towards impacts with more visible ejecta cover during the manual detection process. However, these findings are based on a very small sample size and contain uncertainties. As more HiRISE images of the 76 newly discovered impacts are obtained and released, the discrepancies, if there are any, between the original database and the machine learning dataset can be explored in more detail.

Acknowledgments: HiRISE data credit NASA/JPL/University of Arizona. AG, IJD, and DW supported by NASA SSW grant 80NSSC20K0789.

References: [1] Malin et al. (2006) *Science, 314,* 1573-1577. [2] Daubar et al. (2016) *Icarus. 267,* 86-105. [3] Bart et al. (2019) *Icarus 328,* 45-57. [4] Daubar et al. (2013) *Icarus. 225,* 506-516. [5] Hartmann et al. (2017) *Meteorit. Planet. Sci. 52*, 493–510. [6] Williams et al. (2014) *Icarus 235*, 23–36. [7] Bryne et al. (2009) *Science, 325,* 1674-1676. [8] Viviano et al. (2019) *Icarus 328,* 274–286. [9] Daubar et al. (2019) *JGR* 124 (4): 958-969. [10] Dundas et al. (2014) *JGR* 119 (1): 109-127. [11] Malin et al. (2007) *JGR* 112, E05S04. [12] McEwen et al. (2007) *JGR* 112, E05S02. [13] Munje et al. (2020) PCC 11, Abs. 2065. [14] https://www.uahirise.org/hiwish/; Chojnacki et al. (2020) *LPSC 51,* Abs. 2095. [15] Christensen et al. (2009) *Am. Geophys. Union Fall Meet,* IN22A-06. [16] Daubar et al. (2020) *PCC 11*, Abs. 2069.