COSMIC: Content-based Onboard Summarization to Monitor Infrequent Change

Gary Doran
Jet Propulsion Laboratory,
California Institute of Technology
4800 Oak Grove Drive
Pasadena, CA 91109
Gary.B.Doran.Jr@jpl.nasa.gov

Steven Lu
Jet Propulsion Laboratory,
California Institute of Technology
4800 Oak Grove Drive
Pasadena, CA 91109
You.Lu@jpl.nasa.gov

Maria Liukis
Jet Propulsion Laboratory,
California Institute of Technology
4800 Oak Grove Drive
Pasadena, CA 91109
Maria.Liukis@jpl.nasa.gov

Lukas Mandrake
Jet Propulsion Laboratory,
California Institute of Technology
4800 Oak Grove Drive
Pasadena, CA 91109
Lukas.Mandrake@jpl.nasa.gov

Umaa Rebbapragada
Jet Propulsion Laboratory,
California Institute of Technology
4800 Oak Grove Drive
Pasadena, CA 91109
Umaa.Rebbapragada@jpl.nasa.gov

Kiri L. Wagstaff
Jet Propulsion Laboratory,
California Institute of Technology
4800 Oak Grove Drive
Pasadena, CA 91109
Kiri.L.Wagstaff@jpl.nasa.gov

Jimmie Young
Jet Propulsion Laboratory,
California Institute of Technology
4800 Oak Grove Drive
Pasadena, CA 91109
Jimmie.D.Young@jpl.nasa.gov

Erik Langert
Palantir Technologies
15 Little W 12th St
New York, NY 10014
elangert@palantir.com

Anneliese Braunegg
Case Western Reserve University
10900 Euclid Ave
Cleveland, OH 44106
Anneliese.Braunegg@case.edu

Paul Horton
Arizona State University
781 S Terrace Rd
Tempe, AZ 85287
pahorton@asu.edu

Daniel Jeong
Columbia University
116th St & Broadway
New York, NY 10027
daniel.jeong@columbia.edu

Asher Trockman
Carnegie Mellon University
5000 Forbes Ave
Pittsburgh, PA 15213
ashert@cs.cmu.edu

Abstract— **Interplanetary exploration occurs at vast distances that severely limit communication bandwidth to spacecraft exploring other planets. It is possible to collect much more scientific data than can ever be downlinked given current communication capabilities. Therefore, we are developing a system called COSMIC (Content-based Onboard Summarization to Monitor Infrequent Change) that will opportunistically analyze data onboard a Mars orbiter to alert scientists when meaningful changes have occurred. COSMIC will allow future spacecraft to continuously collect data to search for rare, transient phenomena such as fresh impacts or seasonally changing polar landforms under a constrained downlink budget. In this paper, we describe the overall goals and architecture of COSMIC, plans to enable specific scientific studies, label acquisition to enable supervised approaches to surface landform classification, a new machine learning evaluation framework for analyzing the trade-offs between classifier accuracy and computational requirements, and lessons learned about constraints that COSMIC will face operating onboard a spacecraft. In particular, we discuss design considerations surrounding computational and storage constraints, change detection strategies, and localizing detected landforms of interest within a global coordinate frame. Finally, we describe challenges and open research questions that must be addressed prior to deploying COSMIC.**

TABLE OF CONTENTS

978-1-7281-2734-7/20/$31.00 ©2020 IEEE

1. INTRODUCTION

The modern era of space exploration has witnessed an explosion in science instrument capability. Higher spatial, temporal, and spectral resolutions now capture observations that have revolutionized planetary science, permitting unprecedented maps of landforms, mineralogy, and thermal properties [1], [2]. Unfortunately, the resulting rise in instrument data collection has outstripped the much more slowly growing communications bandwidth back to Earth. This has created a data crisis at the point of collection: we can trivially gather more data than can be downlinked for analysis. This is especially true for outer planets and other distant targets of exploration, but is even the case for neighboring planets like Mars. Mission scientists have responded by carefully targeting observation requests: we obtain low-resolution maps of large regions to inform hand-selection of areas of interest for high-resolution observation using complementary instruments. For example, on the Mars Reconnaissance Orbiter (MRO) spacecraft, landforms identified with the 6 m/pixel resolution Context Camera (CTX) instrument [3] might be used to target observations with the 0.25 m/pixel resolution High Resolution Imaging Science Experiment (HiRISE) instrument [1]. The role of the mission planning scientist has thus become a fundamental part of the flight operations process. Expert scientists able to receive new data, deduce its scientific content, and optimize follow-on behavior ensure the highest science yield from these expensive, rare missions. For example, the MRO science planning cycle is performed on a fortnightly schedule [4].

Scientific inquiry using this human-in-the-loop model has provided an excellent understanding of the surface morphology of Mars, seasonal behavior, and the dynamic past forces resulting in the impressive diversity of the red planet. Recent evidence has shown, however, that the surface of Mars is not nearly as static as once thought. Frozen carbon dioxide and water ice are seasonally redistributed from pole to pole with sublimation forming pitted terrain resembling Swiss cheese and violent eruptions of dusty vapor that are dispersed over the surface by wind [5], [6]. Dust devils seasonally scour the surface leaving vast, interconnected trails [7], [8]. Superimposed upon this dynamic landscape, meteors create blast-like marks radiating from sharp, fresh craters, some of which can reveal ice trapped beneath the surface [9].

Mission instruments have captured dynamic events in scientist-targeted images chiefly through serendipity. As the onset of these events is unknowable, it is only by luck that scientists happened to request images of the same location both before and after, or during, the change, often in an image taken for an entirely different reason. Thus, surface change detection is under-served by the human-in-the-loop model due to the need to downlink repeat images of any locations of interest. For example, there are nearly 7000 HiRISE observations in the Planetary Data System (PDS) archives whose purpose is described as some form of "monitoring." Therefore, the detection, capture, and summarization of change and dynamic events has become one of the primary goals of onboard autonomy, in Mars orbit and beyond [10]. To serve this need, a novel operational paradigm and associated autonomy capability is required.

In this work, we address the specific challenges of detecting surface change onboard future Mars orbiters using imagers modeled on the current HiRISE and CTX instruments. We assume that future imagers at Mars will be capable of observing as frequently as Earth-observing spacecraft, whose power and thermal systems are designed to operate for longer periods to take advantage of the higher downlink bandwidth in low-Earth orbit. We propose to couple the more frequent imaging cadence with upcoming advances in High-Performance Spaceflight Computing (HPSC) platforms [11] to develop a new autonomous system, COSMIC, capable of detecting and summarizing active change to avoid unnecessarily downlinking repeat images for the purpose of ground-based monitoring. The faster cadence of change detections at a global scale will allow scientists to better characterize transient Martian events and sample change frequency in a more unbiased manner. COSMIC will supplement regular scientist-selected observations.

This paper does not report on a completed, validated system but rather exposes the foundational concepts within COSMIC and reports on lessons learned, which are applicable more generally to deployed machine learning systems under computational constraints. We describe how the system architecture of COSMIC allows change detection of small landforms globally, including challenges related to onboard storage and localization of landforms. Further, we describe plans to address trade-offs between accuracy and computational complexity of classifiers for landform detection given onboard processor constraints. Finally, we describe our experience collecting crowd-sourced labels to train supervised classification algorithms for detecting specific landforms or regions of interest. Our preliminary results demonstrate the feasibility of a system like COSMIC, highlight its benefits to future Mars missions, and establish areas of focus for future work.

2. System Architecture Overview

The core goal of COSMIC is global-scale change monitoring to detect rare, transient events on the surface of Mars. One approach to change detection uses at least two images of the same location that are co-registered and compared at the pixel level to determine areas of alteration. While COSMIC can support this direct form of change detection, we show in Section 3 that such an approach is only feasible for small regions or at low resolution globally. To solve the global detection problem, COSMIC will build a large onboard database of summarized findings within high-resolution images. Summaries include both identified landforms and regions identified by visual salience then classified into one of several categories [12]. Change is detected by comparing these landforms and regions rather than raw pixels. Figure 1 shows the modules of the COSMIC architecture used to implement this change detection strategy, with more details provided below.

Investigation Coordination - The Arbiter

The Arbiter will drive investigations, coordinate the execution of modules, and manage data storage. It will serve as the primary interface through which the system is accessed and commanded by scientists. Note that COSMIC does not interfere with any scientist-targeted observations. This is a crucial issue for mission infusion, as onboard autonomy has only begun to build trust in the planetary science community. Instead, COSMIC will enable scientists to pose new questions about previously unattainable observations, receive alerts and reports on potential change events, and monitor Mars in the global change detection context.

Process Support

Process Support modules will provide services that are utilized across the Science Modules. Traditionally, the functionality provided by these modules would be part of a ground-based data processing system. However, COSMIC's success requires these functions to be performed onboard.

Alignment—This module will enable the alignment or localization of one image with respect to another image or a global coordinate frame. The alignment module will work by identifying *tie-points* within a new image, which are matched with tie-points extracted from a reference image. The requirements of this localization subsystem are described further in Section 3.

Vault—The Vault will be an onboard database storing data products from prior and on-going investigations, supporting metadata, reference images of key regions of interest, and various configurations. As discussed in Section 3, the format of data products will emphasize the conservation of space, relying on terse descriptive representations as opposed to complete image data. Highly salient regions would be summarized here, providing an encyclopedic, global map of overall potential interest. Due to its size, the onboard database is not intended to be downlinked completely. Instead, scientists may issue analytical SQL-like queries to the spacecraft where they are evaluated to provide summaries returned to Earth. For example, scientists might ask for all

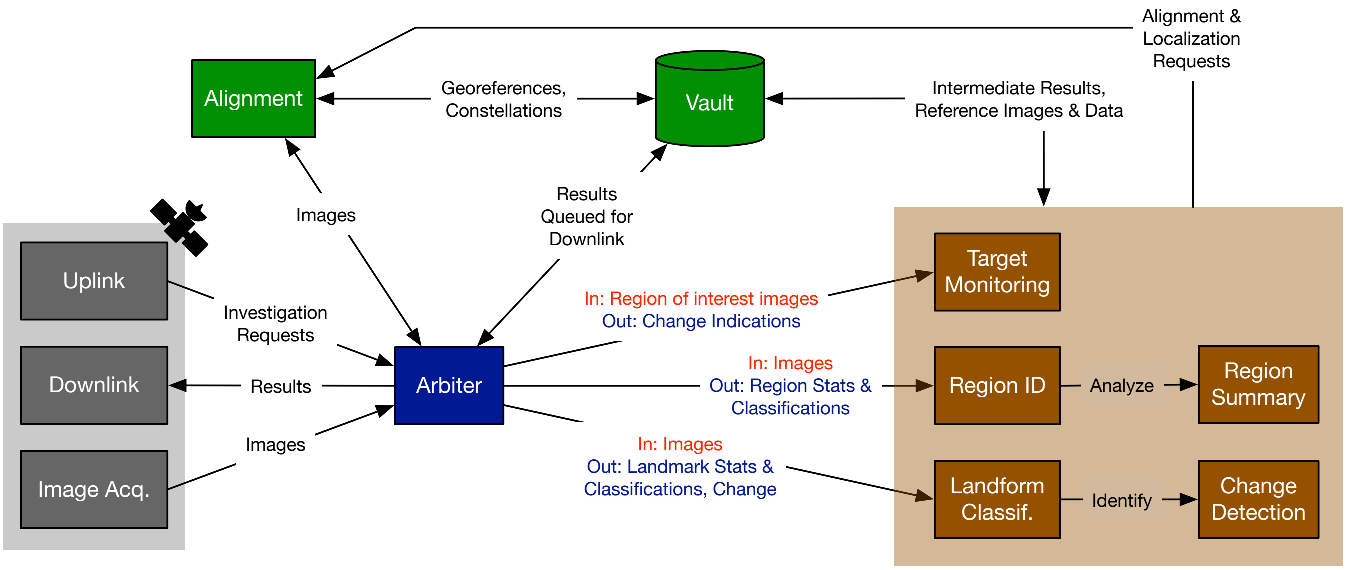


Figure 1: The architecture of the COSMIC system. The gray boxes at the left represent functions that are accessed through the spacecraft's operating system: *uplink* of commands from Earth, *downlink* of data back to the ground, and *image acquisition* through the orbiter's imaging instruments. The remaining modules are color-coded according to their functionality: blue for *Process Coordination*, green for *Process Support*, and brown for *Science Modules*.

fresh impacts between given date and latitude ranges.

Science Modules

Science modules will identify, classify, and detect change in surface landforms and regions of interest. The modules will support different modes of investigation, described below.

Target Monitoring—Scientists may specify small regions with change potential that are imaged at full resolution at each opportunity. Each image is aligned with previous images for pixel-based change detection. When changes are detected, alerts will be sent to the ground to notify scientists, who can then choose to downlink the detections or save them for later access. This saves bandwidth over the approach of downlinking all images and analyzing them on the ground.

Region ID—Global planetary change detection includes a large regional context that can span much or all of an image rather than individual landforms. The detection and identification of large regions of interest will be another COSMIC capability, by discovering both regions similar to a known catalog of scientifically interesting terrain types and regions that are different from their surroundings without a guiding library of known targets.

Region Summary—Each identified region will be characterized by boundaries and summaries of shape and visual properties. Image segmentation will be used to summarize visually distinct regions within images that do not have explicitly defined types identifiable using Region ID.

Landform Classification—Supervised classification algorithms will be used to find landforms of interest (e.g., craters, dunes, dust devil tracks). As classification is a computationally expensive process, a relatively quick visual salience estimate [12] will be used first to focus attention on candidate regions. The salience preprocessor is necessary for real-time processing of streaming, high-resolution imagery. Classifications with high confidence will be stored in the onboard database to provide additional context for alerts and change reports. While the classification process does not explicitly perform change detection, some surface landforms are direct indicators of recent change, such as the distinct patterns surrounding fresh meteorite impacts.

Change Detection—The Change Detection module will compare classified landforms with previous detections to determine if any new landforms have appeared or any prior landforms are no longer visible [12]. This form of change detection, based on differencing objects rather than raw pixels, addresses the data volume and computational constraints of an onboard planetary monitoring system. Maps of planetary change can be continually updated, with alerts and reports sent to Earth for areas of change.

Together, these capabilities will autonomously produce a rich and thorough map of the surface of Mars in terms of visual interest, identified landforms, regions of interest, statistically distinct regions, and high-resolution pixel captures surrounding ground-specified target regions. The utilization of this database to provide alerts and reports will better inform ground scientists, who will remain in full control without being required to regularly maintain the operation of the autonomous system.

3. Design Considerations

When designing a system such as COSMIC to operate onboard a spacecraft, there are unique design considerations that must be made to accommodate the computational, memory, power, and other requirements. Here, we describe a subset of those considerations for the science, storage, and localization subsystems.

Science Modules

The capability of classifying scientifically interesting regions and landforms is critical to COSMIC. Onboard image classification is a complex problem that requires considering many factors, including CPU speed and memory constraints. To estimate the computational requirements, we are exploring three different image classification strategies: standard feature-based classification, pixel-based classification, and state-of-the-art deep learning classifiers.

Classical machine learning requires manual feature extraction before classification. Features must be engineered carefully to differentiate between classes of interest. General texture-related characteristics can be captured by extracting the minimum, maximum, median, and other statistics from images, while Scale Invariant Feature Transform (SIFT), Local Binary Pattern (LBP), and Gray-Level Co-Occurance Matrix (GLCM) features can be used to summarize the local structures of images [13]. Although feature extraction can be computationally expensive, carefully engineered features can provide an efficient and effective summary of image contents that enables computationally inexpensive classification.

As an alternative to classical machine learning algorithms operating on pre-extracted image features, the TextureCam algorithm uses a random forest to perform pixel-wise classification for recognizing geological features [14]. Because TextureCam only computes features as they are needed at each branch in each decision tree, it can feasibly be evaluated using onboard processor architectures. For example, TextureCam was deployed on the Intelligent Payload EXperiment (IPEX) CubeSat [15] which carries a 400 MHz Atmel ARM9 CPU without hardware floating point unit and 128 MB of RAM, and the even more constrained Earth Observing-1 (EO-1) spacecraft, which has only a 12 MHz Mongoose M5 processor [16].

Since the breakthrough of AlexNet [17] in the 2012 ImageNet competition, deep convolutional neural networks (CNNs) have become increasingly popular for image classification problems. One of the advantages of deep CNNs is the performance measured in classification accuracy, which comes at the cost of relatively high computational complexity. The number of multiply-add operations required for inference can be measured as a function of the size of the input image and the structures of layers in the network. For example, 15.5 billion operations are required in order for GoogLeNet to perform a single forward pass for one image. As a point of comparison to spacecraft computing rates, the radiation-hardened RAD750 processor onboard the MRO spacecraft can process approximately 200 million instructions per second (MIPS). With the current limited spacecraft computational resources, it is not feasible to deploy deep CNN models even with the recent advances in reducing precision of operations, number of operations, and model size [18]. NASA's HPSC program aims to improve the current state-of-art spacecraft processors by two orders-of-magnitude by 2021 [11], which will bring the computational rate to approximately 50,000 MIPS. The possibility of deploying deep CNN models using an HPSC system remains to be examined in the future.

Science Storage

A key component of COSMIC is the ability to store a database onboard the spacecraft of known landforms and regions to serve as a basis of comparison for change detection. One straightforward approach to change detection is to store an onboard global map of the surface with high enough resolution to detect the appearance of new landforms of interest. However, with small landforms of interest, the storage requirement for a map to resolve such landforms quickly becomes infeasible. On the other hand, a database that just stores landform locations and some additional metadata has a size that depends only on the density of the landform, not its size.

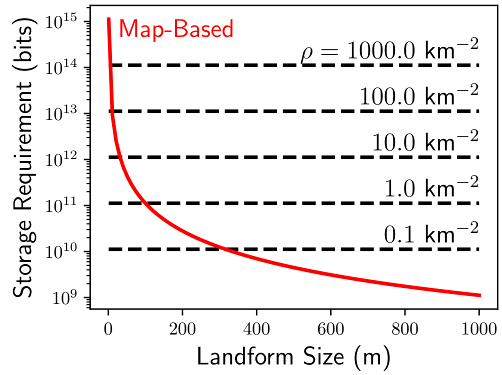


Figure 2: A comparison of the storage required by landforms of given sizes and densities, using a map-based approach (solid) versus a database-based approach (dashed).

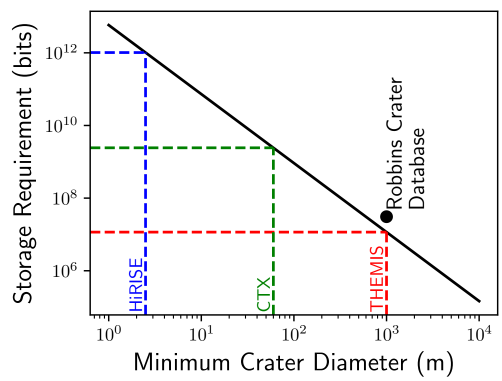


Figure 3: The estimated storage requirement for craters exceeding a given diameter. For comparison, the size of the existing Robbins Crater Database derived from THEMIS images and estimated sizes for databases derived from imagers with other resolutions are indicated with dashed lines.

As a concrete comparison, suppose a landform has a constant size (diameter) s and average surface density ρ. Let $A_{♂}$ be the surface area of Mars. Then there will be $\left(\rho A_{♂}\right)$ entries in a database, versus $\left({}^{A_{♂}}/{}_{s^2}\right)$ pixels in a map with pixel size on the order of the landform size. If a pixel is represented with 1 byte, and a database entry with 100 bytes, then Figure 2 shows the storage requirement as a function of landform size for the map-based approach in contrast with the requirement for the database-based approach with various landform densities. The figure shows that a map-based approach only requires on the order of 1 gigabit to resolve landforms on the order of 1 km in size, but 1 petabit to resolve landforms on the order of 1 m in size. In contrast, landforms of any size can be stored in less than 100 gigabits as long as their density remains below $1\,\mathrm{km}^{-2}$.

Above, we made a simplifying assumption that each landform is roughly constant in size. However, there exist classes of landforms that manifest at various scales. In particular, impact craters of all sizes exist on the surface of Mars.

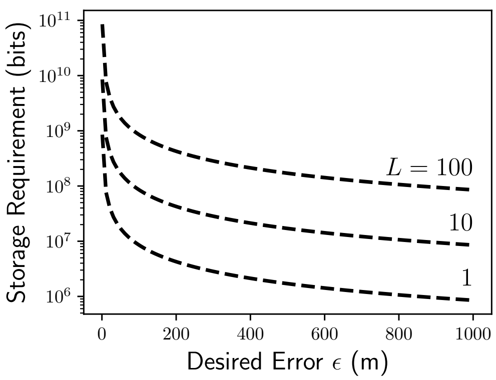


Figure 4: The required storage size for a single polygon defining a polar cap boundary with errors at most ϵ. The requirement is shown for different values of Lipschitz constant L parametrizing the smoothness of the boundary.

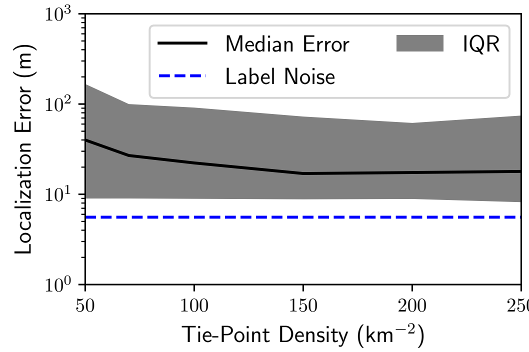


Figure 5: Registration error of CTX images using the VPT algorithm with various tie-point densities. Both the median and IQR of registration errors are shown, which the dashed line indicating the theoretical floor due to label noise.

Furthermore, the density of craters on Mars is not constant with respect to their size. Instead, cumulative crater counts on Mars follow a power law given by $N_{\geq d} \propto d^{-1.9}$, where $N_{\geq d}$ is the number of craters exceeding some diameter d [19]. Using this power law estimate for the size–frequency distribution of impact craters on Mars, we can derive storage requirements for all craters exceeding some diameter. Assuming roughly 10 bytes per crater to store each crater's location and diameter, Figure 3 shows the storage requirement for all craters exceeding a given diameter. The figure compares the estimate with the actual size of the Robbins Crater Database, derived from Thermal Emission Imaging System (THEMIS) images [20], [2]. Figure 3 also shows estimates of how much space would be required to store all craters identifiable in images with resolutions comparable to that of either the HiRISE or CTX instruments. The corresponding crater diameters are computed by multiplying the resolutions of CTX (6 m) and HiRISE (0.25 m) by 10, since the Robbins Crater Database was only able to robustly identify craters that were 10 pixels across. In this case, we see that all craters identifiable at CTX resolution can be stored within 10 gigabits of memory.

In addition to landforms, COSMIC will also need to store regions of interest over time to detect change. The same argument as above can be used to show why storing polygonal region boundaries is more efficient than storing map-like inclusion masks for regions, assuming region boundaries are relatively smooth so that their perimeter–to–area ratios remain less than the ratio of storage costs of mask pixels to vertices. To see how large an individual region might become, consider a concrete example region type: an entire polar ice cap that can vary in extent seasonally. For simplicity, assume that the smoothness of the cap boundary satisfies a Lipschitz condition, meaning that for every change in longitudinal extent by Δx, the change in latitudinal extent is bounded by $L\,(\Delta x)$, for constant L. Then to trace out the polar cap extent to within ϵ everywhere along the perimeter, it suffices to sample points with a spacing in the longitudinal direction of $(2\epsilon/L)$. As a conservative upper bound, suppose a polar cap can grow to consume an entire hemisphere of Mars. In this case, the perimeter in the longitudinal direction is the circumference of Mars at its equator, so the number of vertex points required given the spacing above is $n \geq \frac{2\pi R_{♂} L}{2\epsilon} = \frac{\pi R_{♂} L}{\epsilon}$, where $R_{♂}$ is the radius of Mars. Assuming each vertex requires on the order of 10 bytes to represent, Figure 4 shows the overall storage requirement for the full polygon given the desired error ϵ in the boundary for a range of values of the Lipschitz constant L. Required storage ranges from 1 megabit to 1 gigabit for errors down to 100 m, even for large constants. These sizes are small enough to compute and store on a regular basis.

What amount of onboard storage can we expect for a future Mars orbiter? The MRO spacecraft, launched in 2005, has over 100 gigabits of onboard storage [4]. It is not unreasonable to assume that an order of magnitude increase in this capacity is possible for future Mars orbiters. Therefore, the examples described above—all landforms with an average density less than 1 km^{-2}, all craters visible with CTX resolution, polar cap extents with 100 m accuracy—can all feasibly fit within an onboard database.

Localization

Given the strategy for change detection via comparing detected landforms with a database of known landforms, COSMIC will be required to accurately map locations in image coordinates to real-world coordinates on the surface. For landforms that are locally dense (such as impact craters, dunes, and polar features), the level of accuracy required is roughly on the order of the size of those landforms on the surface. Very accurate estimates of a spacecraft's ephemeris, or position and velocity with respect to Mars, can be made periodically to within several meters, but the spacecraft's state begins to drift slowly between these updates. For the MRO spacecraft, the drift rate translated to ground pointing error can be as high as 8.5 m/h [4]. The ephemeris of MRO is updated frequently enough so that pointing error never exceeds several kilometers. Therefore, COSMIC must reduce pointing errors from potentially several kilometers to a level suitable for identifying small landforms on the surface.

The process of localizing a spacecraft with respect to landforms is also performed during Terrain Relative Navigation (TRN), which is typically used during Entry, Descent, and Landing (EDL) or navigation around small bodies and relies on storing onboard maps [21]. However, as for science purposes, storing an onboard map for the entirety of Mars at a resolution sufficient for precise localization is infeasible. Instead, a set of image features such as SIFT can be pre-extracted from a global map, then matching features can be found within orbital imagery during flight [22].

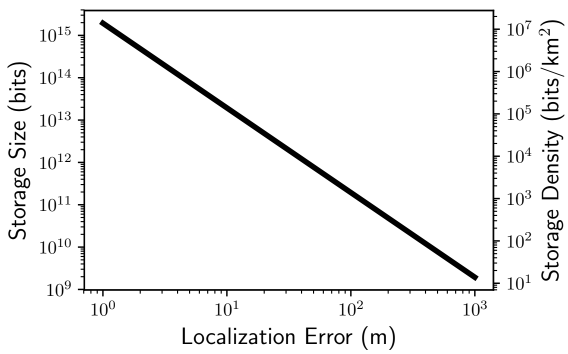


Figure 6: Required onboard storage capacity to achieve the corresponding localization error. Storage requirement is expressed both in terms of total storage for global coverage as well as storage per unit area.

There are many algorithms that perform tie-point based image alignment, but one implementation designed for use on spacecraft is the Visual Precision Targeting (VPT) algorithm developed for Mars rover instruments [23]. We evaluated the ability of VPT to accurately align two orbital images using 10 pairs of manually co-registered CTX images. From the known alignment, the images were offset randomly by a Gaussian with standard deviation corresponding to the offsets typically experienced by MRO. Figure 5 shows the resulting registration error when various densities of tie-points are extracted from each of the images. Both the median and interquartile range (IQR) across the random samples are shown. Due to errors in the human-labeled ground truth tie-points, error below the dashed line is not expected for images of this resolution.

While the experiments above evaluate registration performance using CTX images, we could also in principle use higher resolution images for localization, assuming we had already refined the spacecraft's ephemeris to estimate the pointing location on the surface to within a comparable factor as MRO's pointing error is to CTX resolution. Thus, we can rescale both the tie-point density and registration error in Figure 5 to correspond to images of other resolutions. The resulting achievable trade-off between localization error and tie-point density is illustrated in Figure 6. For a desired localization error along the x-axis, the required storage for VPT tie-points (40 bytes each) is shown along the y-axis. Required storage is expressed both in terms of total storage for global coverage as well as storage per unit area if high-precision localization is desired only within a relatively small region of interest. Given around 1 terabit of onboard storage, it is feasible to localize anywhere on Mars to within 10-100 m, or more accurately within select regions.

Finally, we consider how frequently onboard localization must be performed. Suppose we desire to estimate the spacecraft's ground-track pointing location to within e at all times. We could localize to within $e/2$, then wait for the spacecraft's estimate to drift (at roughly 8.5 m/h) back to e. Even if the desired error $e = 1\,\mathrm{m}$, VPT would only need to be invoked once every several minutes. In fact, it might be possible to extract tie-points from only certain latitude bands across the surface of Mars in such a way that the spacecraft can still be re-localized at the required cadence. This strategy could maintain the desired localization error while reducing the total storage required for tie-points.

4. EXAMPLE SCIENCE APPLICATIONS

There are several ongoing and dynamic events and processes on Mars that are valuable to identify and catalog. In this section, we describe three examples of scientific phenomena that create surface changes of the kind that the COSMIC system is designed to detect. We use data from the HiRISE instrument on the MRO spacecraft to provide examples of each landform. The greatest surface coverage comes from HiRISE red-band observations (550–850 nm), which are displayed as grayscale images. We are collecting labels for each of these landforms to determine how capable the supervised classification strategies described in Section 3 are at detecting these landforms under onboard computational limitations.

Fresh Meteorite Impacts

The surface of Mars is heavily cratered, and it continues to accumulate new meteorite impacts each year. Typically these manifest as blast zones that modify the surface, e.g., by removing light-colored dust and exposing darker bedrock [9]; see Figure 7a. The radial shape of the blast varies as a function of the angle of the impact, current wind conditions, and local terrain that may interfere with the ejecta or dust clearing processes. Some meteors break up into multiple impactors and create clusters of impacts, or remove darker dust from a lighter underlying bedrock. In some cases, the impact exposes underlying subsurface ice, which sublimates and disappears within a few months [24]. By studying these impacts to make cratering models, the duration of surface processes and the age of landscapes can be better estimated, making it easier to interpret Martian geological history. However, only a few hundred fresh impact sites have been imaged by HiRISE, primarily in homogeneous dusty areas [9]. Thus, the ability of COSMIC to better detect and image fresh impacts would be valuable for planetary science.

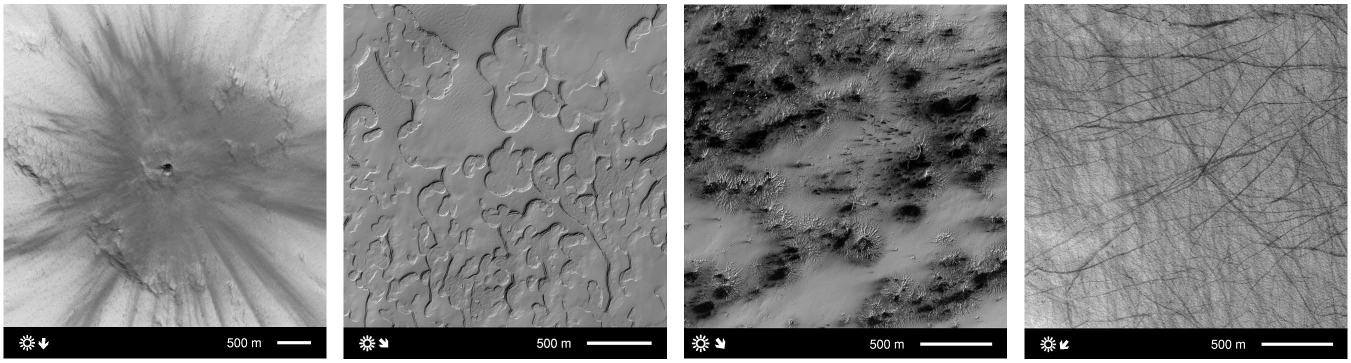
To date these impacts have been discovered by careful inspection of lower-resolution (6 m/pixel) images obtained by the CTX instrument on MRO, followed by the collection of a high-resolution image from HiRISE for confirmation. No one has yet attempted to catalog all impacts in the more than $50{,}000$ HiRISE images (each one up to gigapixels in size) collected since 2006.

The Mars InSight mission, which landed in November 2018, carries a sensitive seismometer that is able to detect any nearby meteorite impacts that occur. InSight is expected to observe 1 to 3 impacts per year that strike within 1200 km of its landing site [25]. In addition to capturing the first in-situ recording of a Martian impact, InSight will use the seismic signal and its reflections to build a map of the subsurface structure of the crust and upper mantle. Correlating any detected signals with visual images of their associated impact sites will be of key scientific value for the mission.

Polar Landforms

On the southern pole of Mars, scientists have observed two particularly interesting types of dynamic terrain called "Swiss cheese" and "araneiforms." Both are consequences of seasonal carbon dioxide (CO_2) ice accumulations on or below the surface. During the Martian winters, CO_2 will cool and form ice both above and below the surface. As springtime approaches, the sun will warm the CO_2 ice and cause it to sublimate. The rapid transformation of the CO_2 results in the formation of different landforms.

Swiss Cheese Terrain was first discovered by the Mars Global Surveyor in 2000 and can be characterized as cavities



(a) A fresh impact (ESP_041953_1885)
(b) Swiss cheese terrain (ESP_022054_0930)
(c) Araneiforms with deposits (ESP_047922_0990)
(d) Dust devil tracks (ESP_036324_2325)

Figure 7: Transient or dynamic landforms observed from the HiRISE instrument as examples of COSMIC applications.

in the surface of Mars due to the sublimation of CO_2. The size of these cavities can be hundreds of meters wide and less than ten meters deep [5]. Their name comes from their similar appearance to the holes in Swiss cheese; examples can be seen in Figure 7b.

Araneiforms, sometimes called "spiders," are another result of subterranean CO_2 ice sublimation. As the surface warms, the sublimated CO_2 will form cracks in the surface to create vents for the gas to escape [6]. The spider-like appearance of these landforms can be attributed to darker dust accumulating around the surface vents formed by the escaping subterranean gas, as seen in Figure 7c. The formation mechanisms for these landforms are not entirely understood, so they are objects of continued study that will provide insight into the past and current Martian climate. Because polar landforms undergo significant seasonal change, COSMIC will be well-suited to studying them and alerting scientists as the most interesting changes take place.

Dust Devils

Dust Devils are tornado-like vortices that form due to differences in local air temperature and are rendered visible by collected dust and other debris [26]. While naturally occurring on Earth, we are interested in those found on Mars. Earth-based dust devils are 2-3 orders of magnitude less powerful than normal tornadoes, but their dust-lifting effects can be rather significant on Mars [26]. Due to the reduced atmosphere of Mars, tracks of dust devils can be found on the Martian surface and are visible from current Mars orbiters, as shown in Figure 7d. Locating and identifying dust devils tracks form an important part of COSMIC's science capabilities.

Dust devils are particularly relevant to spacecraft operations and power management as they provide a means for critical dust removal from the solar panels of rover and landers [27]. Dust devils also contribute to noteworthy changes in albedo on the Martian surface and provide vital insight for geologists [7]. Geologists utilize the presence of dust devil activity as a way to understand the granularity of the surface particles. The tracks also provide information about the winds in the lower atmosphere which can be used to model erosion and deposition on the surface [8].

Label Gathering

To enable COSMIC to automatically identify the scientific landforms of interest described above, labeled examples of such phenomena are used to train supervised learning algorithms. Since label collection can be an expensive and labor-intensive process, we have created a citizen-science project on the Zooniverse web platform to crowd-source labels.[2] Zooniverse provides customizable user interfaces for labeling data in various ways and establishes a communication channel between researchers and citizen scientists. We created four Zooniverse workflows to collect labeled data for fresh impacts, Swiss cheese, araneiforms, and dense dust devil fields. Citizen scientists are provided with detailed tutorials on how to perform the labeling tasks, and talk boards are made available to both citizen scientists and planetary domain experts to discuss ambiguous examples or interesting observations.

In order to utilize different classification strategies, citizen scientists are asked to provide both pixel-level and image-level labels. Pixel-level labels in the form of arbitrary polygons would enable a machine learning system to mark the exact locations of identified landforms and regions, whereas image-level labels would enable a machine learning system to classify landforms already identified by bounding boxes. This labeling strategy is somewhat different from that employed by similar, previous Zooniverse projects such as Planet Four, which had users explicitly label certain shape properties of araneiforms to evaluate hypothesized formation processes [6]. In contrast, our task is intended only to learn more generally how to localize these landforms within orbital imagery rather than to extract specific properties for scientific analysis.

There are a total of $14{,}512$ images across the four labeling workflows, and we specified that each image should be labeled by 11 different users to enable robust consensus labeling. Label accuracy can depend on a user's background, such as previous experience analyzing orbital imagery, so requiring that each image be labeled multiple times helps improve label quality. For image-level labels, if a strong majority of users provide the same label, then this is taken as ground truth. Otherwise, a domain expert will take a closer look to break ties. For pixel-level labels, a probabilistic label can be derived for every pixel by counting the fraction of users who included that pixel in their labeled region.

[2] https://www.zooniverse.org/projects/wkiri/cosmic

5. SYSTEM EVALUATION STRATEGY

The labels collected for the science use cases described in Section 4 will be used to determine how effective machine learning algorithms are at identifying targets of interest. More importantly for COSMIC, we are interested in the trade-off between classification performance in terms of accuracy, precision, recall, etc., and computation requirements such as runtime and memory consumption. We have developed a new framework, Distributed Optimization of ML Incorporating Nested Evaluation (DOMINE), to robustly evaluate these trade-offs.

Manually constructing a quality machine learning algorithm is a multi-step process which includes data understanding, model selection, hyperparameter optimization, and evaluation. Our goal is to create an automated machine learning system that supports both traditional and deep learning based algorithms with hyperparameter optimization in a fully cross-validated manner. In addition, we also aim at building the system with reasonable interpretability and traceability such that it can help machine learning researchers and engineers determine the model that is most suitable for the problem.

DOMINE provides a distributed solution to this need. Its architectural diagram is shown in Figure 8. DOMINE consists of one configuration file, one server instance, as many client instances as needed, one evaluation instance, and one database instance. The configuration file is the entry point that defines the parameters of a DOMINE experiment, including the location of the data set on each client (processing) machine, the cross-validation method, and the learning algorithms with associated hyperparameter ranges. The server instance is responsible for generating tasks based on the information provided in the configuration file. A task in DOMINE is defined as the combination of a learning algorithm with the associated hyperparameters to be explored with respect to training and testing data. The client instance is responsible for processing the tasks and returning the results back to the server instance so that these results can be stored in the database instance.

The results from all folds and hyperparameter options are saved in the database to ensure traceability, interpretability of results, and assessment of how sensitive an algorithm is to changes in those parameter values. The distributed nature of DOMINE enables many different parameter options to be explored (using random or grid search) in parallel across each fold-based split of the data into train and test sets.

DOMINE by default supports both traditional and deep learning classification algorithms, and it can also be expanded to support custom classification algorithms. For traditional algorithms, we utilize scikit-learn (`sklearn`) as the classification framework. We enable `sklearn`'s methods such as `sklearn.linear_model.LogisticRegression` by building a wrapper layer. Feature vectors are required to use these traditional algorithms. DOMINE does not provide a feature engineering capability, but DOMINE can be configured to enable `sklearn`'s recursive feature elimination (RFE) estimator for feature ranking. For deep learning algorithms, we support two CNN structures: AlexNet [17] and Inception v3 [28]. We exploit transfer learning techniques to fine-tune the CNN classifiers. The CNN classifiers were pretrained on 1.2 million ImageNet [29] images, and DOMINE adapts them to recognize new classes.

An example of results generated by DOMINE is shown in Figure 9. This plot shows the trade-off in accuracy and classifier evaluation time across three classifiers for a dataset derived from the Zooniverse fresh impact labels. The x-axis shows the classifier accuracy at distinguishing fresh impacts from other images of the Martian surface, and the y-axis shows the amount of time required to evaluate the trained classifier on a new test image. The logistic regression and random forest classifier are classical machine learning (ML) approaches that use hand-engineered features (such as properties of the image histogram), whereas AlexNet uses a CNN-based approach with feature derived from the ImageNet dataset. This shows that while AlexNet can achieve marginally better accuracy, it does so at an increased computational expense. Evaluating these trade-offs will be a key aspect of COSMIC going forward.

6. LESSONS LEARNED

The process of developing a preliminary design for COSMIC has revealed several important lessons about performing change detection onboard a spacecraft. Foremost are the design choices imposed by the relatively constrained computational environment onboard most spacecraft computers. Unlike many machine learning problem domains for which virtually unlimited computing resources can be brought to bear on the problem, here a careful trade-off between model performance and computational cost must be made.

In addition to being compute-limited, the onboard setting is also storage-limited. Accordingly, the strategy of direct image comparison for change detection is not feasible for onboard global change detection. Instead, we have identified an alternative strategy, which is to extract landforms of interest from orbital images, then perform change detection using a database of known landforms. This strategy enables change detection at a much smaller scale than would be feasible if a high-resolution reference map were required.

We also identified localization of landforms in images as a major challenge to the strategy of storing landforms with their locations in a database. From Section 3, we see that although tie-point based registration appears to be feasible to localize landforms to within 10-100 m of their true locations, the number of tie-points required to do this at a global scale requires storage comparable to that for many of the science targets themselves. If there are reliably identifiable types of landforms such as impact craters that have sufficient global density, then these landforms might serve as tie-points for navigation while also serving a scientific purpose.

Unlike many ML problems in which classifiers and other image processing algorithms are expected to encounter (almost surely) unique test images drawn from some hypothetical distribution, the COSMIC use case is different in that each location on the surface will be classified repeatedly. Thus, for COSMIC, we are making a special effort to evaluate the *stability* of our algorithms under realistic variations in lighting conditions, viewing perspective, and seasonal changes. Our initial plans to evaluate algorithms under the common assumption of independent and identically distributed test examples required modification to accommodate this stability analysis, which explicitly compares outputs across pairs or tuples of correlated examples.

Finally, we have learned about the costs and limitations of labeling data to train systems that will be used to recognize landforms of interest on the surface. Table 1 shows the number of examples for each of the Zooniverse workflows,

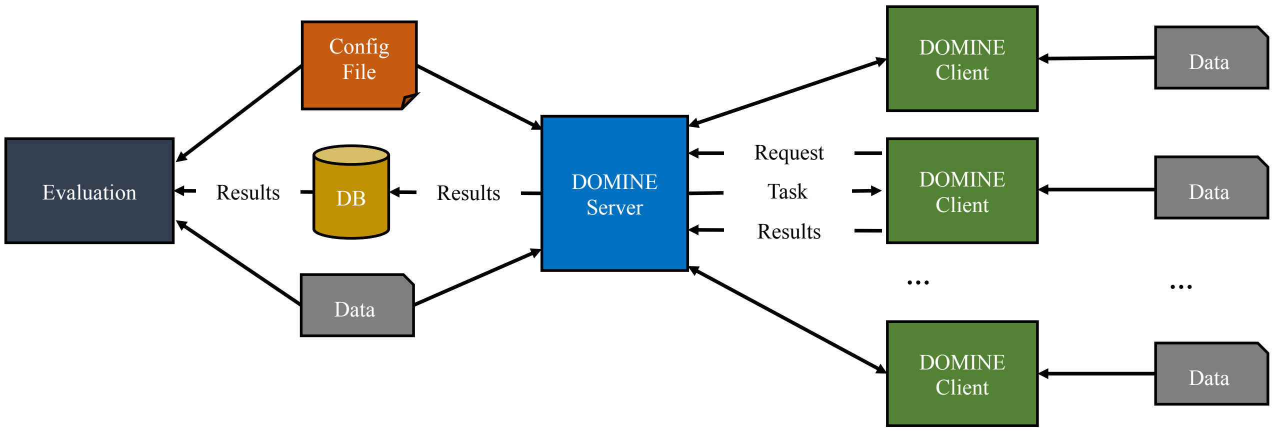


Figure 8: The architecture of the DOMINE system.

Table 1: Labeling costs and rates for Zooniverse workflows

Workflow	Examples	Time/Label	Total Effort	Label Rate	Label Efficiency	Total Time
Fresh Impacts	308	76 s	71 h	$1.9\,h^{-1}$	4.1 %	0.2 yr
Dust Devils	1138	8 s	29 h	$8.2\,h^{-1}$	1.9 %	0.2 yr
Swiss Cheese	5974	6 s	117 h	$6.4\,h^{-1}$	1.1 %	1.2 yr
Araneiforms	7092	37 s	793 h	$5.4\,h^{-1}$	5.5 %	1.7 yr

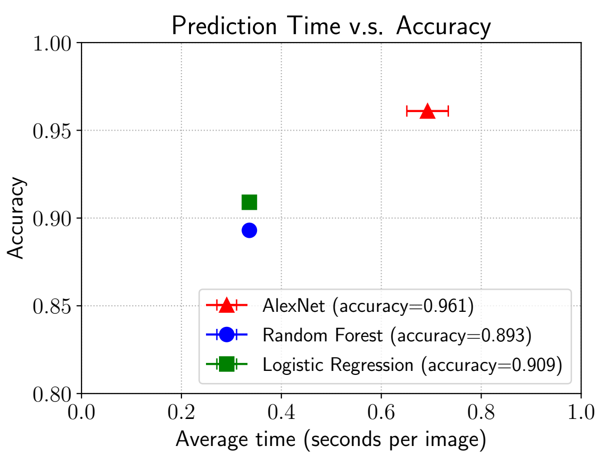


Figure 9: An example output from DOMINE showing the trade-off in accuracy and evaluation time requirements for classical and deep learning algorithms.

as well as median time taken to label each example to date. The total effort is our estimate for how many person-days it will take to acquire 11 labels for each example (the retirement criterion). Varying labeling times correspond to varying difficulty of the labeling tasks; some tasks require users to draw detailed polygons for each example, whereas others only require a simple "yes" or "no" response. The "Total Effort" assumes that there is a single person labeling continually, whereas in reality, examples are labeled at a particular "Label Rate" depending on how many people are online and actively labeling concurrently. We also show the "Label Efficiency," which is the observed utilization of labelers relative to a single person labeling continually. This shows that while we are drawing from a large pool of volunteers, we effectively have the equivalent of one person labeling only 1–5% of the time. The final column show the estimate total time it will take to acquire all labels. Since we deployed the Zooniverse project in December 2018, two of the workflows (fresh impacts and dust devil tracks) have been completed, but it is estimated to 1–2 years total for the araneiforms and Swiss cheese workflows, which are still ongoing. Thus, there is a substantial effort required to acquire the labels necessary to train classifiers for COSMIC, in addition to the several months that were required to set up the Zooniverse project itself. An open question for future work is to determine how many of these labels are necessary to achieve good classifier performance; we likely can achieve reasonable performance even with a subset of the labels.

While crowd-sourcing can help reduce the total time required to acquire labels, there are limitations to the distinctions that citizen scientists can reliably make while labeling, compared with domain experts. For example, rather than focus on distinguishing specific scientific classes of phenomena such as araneiforms, we instead asked users to make higher-level visual distinctions. We adapted our strategy to focus on labeling according to classes that both amateur humans and machine learning algorithms are more likely to discriminate rather than attempt to make subtle distinctions that only an expert might reliably discern.

7. Conclusions and Future Work

The development of COSMIC to its current state both demonstrates the potential benefits of monitoring for changes and transient events onboard a Mars orbiter and highlights directions of future work required to deploy such a system successfully. Given a better understanding of the requirements and behavior of each component, we must build a framework that integrates the various components together. Key to this integrated system will be the Arbiter module that performs simple planning and scheduling of analyses to be performed as data is collected. Because the Arbiter will be the primary interface that scientists use to specify their preferences to COSMIC, there are interesting open questions about how to effectively command the Arbiter in a way that is both flexible and intuitive, but leads to predictable and reliable behavior by the system.

Characterizing the reliability of COSMIC's performance on identifying scientifically relevant landforms and regions will consume a large portion of the effort going forward. Given labels acquired from Zooniverse, we will perform a rigorous analysis of the trade-offs in performance and classification accuracy across the approaches described in Section 3. Performance across a variety of computer architectures such as the RAD750 and HPSC will inform spacecraft design for future missions that wish to use a system such as COSMIC to increase the amount of science value that can be returned given a constrained downlink budget.

In addition to describing the computational constraints on deploying machine learning systems onboard spacecraft, we have discussed alternative designs for onboard storage systems that enable planetary-scale change detection without requiring that a high-resolution map be saved as a baseline. Similarly, we discussed strategies for localization of targets identified by COSMIC using a tie-point based approach. Finally, we described an overall system architecture that allows scientists to selectively monitor certain landforms and regions of interest such as fresh impacts, polar landforms, and dust devils. Our preliminary analysis indicates that although there remain many areas of future research, COSMIC's approach to onboard science analysis is promising for future spacecraft both at Mars and elsewhere in the solar system.

ACKNOWLEDGMENTS

We thank the Jet Propulsion Laboratory Research and Technology Development program for funding support. This publication uses data generated via the `Zooniverse.org` platform, development of which is funded by generous support, including a Global Impact Award from Google, and by a grant from the Alfred P. Sloan Foundation. Part of this research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

REFERENCES

[1] A. S. McEwen, E. M. Eliason, J. W. Bergstrom, N. T. Bridges, C. J. Hansen, W. A. Delamere, J. A. Grant, V. C. Gulick, K. E. Herkenhoff, L. Keszthelyi, R. L. Kirk, M. T. Mellon, S. W. Squyres, N. Thomas, and C. M. Weitz, "Mars reconnaissance orbiter's high resolution imaging science experiment (hirise)," *Journal of Geophysical Research: Planets*, vol. 112, no. E5, 2007.

[2] P. R. Christensen, B. M. Jakosky, H. H. Kieffer, M. C. Malin, H. Y. McSween, K. Nealson, G. L. Mehall, S. H. Silverman, S. Ferry, M. Caplinger *et al.*, "The thermal emission imaging system (THEMIS) for the Mars 2001 Odyssey mission," *Space Science Reviews*, vol. 110, no. 1-2, pp. 85–130, 2004.

[3] M. C. Malin, J. F. Bell III, B. A. Cantor, M. A. Caplinger, W. M. Calvin, R. T. Clancy, K. S. Edgett, L. Edwards, R. M. Haberle, P. B. James, S. W. Lee, M. A. Ravine, P. C. Thomas, and M. J. Wolff, "Context camera investigation on board the Mars reconnaissance orbiter," *Journal of Geophysical Research: Planets*, vol. 112, no. E5, 2007.

[4] D. Wenkert, N. Bridges, W. Eggemeyer, A. Hale, D. Kass, T. Martin, S. Noland, A. Safaeinili, and S. Smrekar, "Mro's evolving process for science planning," in *AIAA Space 2007 Conference & Exposition*, 2007.

[5] P. C. Thomas, M. C. Malin, P. B. James, B. A. Cantor, R. M. Williams, and P. Gierasch, "South polar residual cap of Mars: Features, stratigraphy, and changes," *Icarus*, vol. 174, no. 2, pp. 535–559, 2005.

[6] M. E. Schwamb, K.-M. Aye, G. Portyankina, C. J. Hansen, C. Allen, S. Allen, F. J. C. III, S. Duca, A. McMaster, and G. R. Miller, "Planet four: Terrains – discovery of araneiforms outside of the south polar layered deposits," *Icarus*, vol. 308, no. 1, pp. 148–187, 2018.

[7] M. Balme and R. Greeley, "Dust devils on earth and mars," *Reviews of Geophysics*, vol. 44, no. 3, 2006.

[8] R. Greeley, P. L. Whelley, and L. D. V. Neakrase, "Martian dust devils: Directions of movement inferred from their tracks," *Geophysical Research Letters*, vol. 31, no. 24, 12 2004. [Online]. Available: https://doi.org/10.1029/2004GL021599

[9] I. J. Daubar, A. S. McEwen, S. Byrne, M. R. Kennedy, and B. Ivanov, "The current martian cratering rate," *Icarus*, vol. 225, no. 1, pp. 506–516, 2013.

[10] K. Di, Y. Liu, W. Hu, Z. Yue, and Z. Liu, "Mars surface change detection from multi-temporal orbital images," *IOP Conference Series: Earth and Environmental Science*, vol. 17, 2014. [Online]. Available: https://iopscience.iop.org/article/10.1088/1755-1315/17/1/012015/meta

[11] W. Powell, "High-performance spaceflight computing (HPSC) program overview," in *Space Computing and Connected Enterprise Resiliency Conference*, 2018.

[12] K. L. Wagstaff, J. Panetta, A. Ansar, R. Greeley, M. Pendleton Hoffer, M. Bunte, and N. Schörghofer, "Dynamic landmarking for surface feature identification and change detection," *ACM Transactions on Intelligent Systems and Technology*, vol. 3, no. 3, 2012. [Online]. Available: http://doi.acm.org/10.1145/2168752.2168763

[13] U. Bayram, G. Can, S. Duzgun, and N. Yalabik, "Evaluation of textural features for multispectral images," in *Image and Signal Processing for Remote Sensing*, 2011.

[14] D. R. Thompson, A. Allwood, D. Bekker, N. A. Cabrol, T. Estlin, T. Fuchs, and K. L. Wagstaff, "Texturecam: Autonomous image analysis for astrobiology survey," in *43rd Lunar and Planetary Science Conference*, 2012.

[15] S. Chien, J. Doubleday, D. R. Thompson, K. L. Wagstaff, J. Bellardo, C. Francis, E. Baumgarten, A. Williams, E. Yee, E. Stanton, and J. Piug-Suari, "Onboard autonomy on the Intelligent Payload EXperiment CubeSat mission," *Journal of Aerospace Information Systems*, 2016.

[16] K. L. Wagstaff, A. Altinok, S. Chien, U. Rebbapragada, S. Schaffer, D. R. Thompson, and D. Tran, "Cloud filtering and novelty detection using onboard machine learning for the eo-1 spacecraft," in *IJCAI 2017 Workshop on AI in the Oceans and Space*, 2017.

[17] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Ima-

genet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, 2012.

[18] V. Sze, Y.-H. Chen, T.-J. Yang, and J. S. Emer, "Efficient processing of deep neural networks: A tutorial and survey," in *Proceedings of the IEEE*, vol. 105, 2017, pp. 2295–2329.

[19] E. S. Kite and D. P. Mayer, "Mars sedimentary rock erosion rates constrained using crater counts, with applications to organic-matter preservation and to the global dust cycle," *Icarus*, vol. 286, pp. 212–222, 2017.

[20] S. J. Robbins and B. M. Hynek, "A new global database of Mars impact craters $\geq$1 km: 1. database creation, properties, and parameters," *Journal of Geophysical Research: Planets*, vol. 117, no. E5, 2012.

[21] A. E. Johnson and J. F. Montgomery, "Overview of terrain relative navigation approaches for precise lunar landing," in *2008 IEEE Aerospace Conference*. IEEE, 2008, pp. 1–10.

[22] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, pp. 91–110, 2004.

[23] G. Doran, D. R. Thompson, and T. Estlin, "Precision instrument targeting via image registration for the Mars 2020 rover," in *Proceedings of the International Joint Conference on Artificial Intelligence*, 2016.

[24] S. Byrne, C. M. Dundas, M. R. Kennedy, M. T. Mellon, A. S. McEwen, S. C. Cull, I. J. Daubar, D. E. Shean, K. D. Seelos, S. L. Murchie, B. A. Cantor, R. E. Arvidson, K. S. Edgett, A. Reufer, N. Thomas, T. N. Harrison, L. V. Posiolova, and F. P. Seelos, "Distribution of mid-latitude ground ice on mars from new impact craters," *Science*, vol. 325, no. 5948, pp. 1674–1676, 2009.

[25] N. A. Teanby, "Predicted detection rates of regional-scale meteorite impacts on mars with the insight short-period seismometer," *Icarus*, vol. 265, pp. 49–62, 2015.

[26] R. D. Lorenz, M. R. Balme, Z. Gu, H. Kahanpää, M. Klose, M. V. Kurgansky, M. R. Patel, D. Reiss, A. P. Rossi, A. Spiga, T. Takemi, and W. Wei, "History and applications of dust devil studies," *Space Science Reviews*, vol. 203, no. 1, pp. 5–37, 2016.

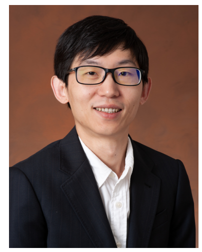
[27] R. Greeley, P. L. Whelley, R. E. Arvidson, N. A. Cabrol, D. J. Foley, B. J. Franklin, P. G. Geissler, M. P. Golombek, R. O. Kuzmin, G. A. Landis, M. T. Lemmon, L. D. V. Neakrase, S. W. Squyres, and S. D. Thompson, "Active dust devils in gusev crater, mars: Observations from the mars exploration rover spirit," *Journal of Geophysical Research: Planets*, vol. 111, no. E12, 2006.

[28] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2016.

[29] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, "Imagenet large scale visual recognition challenge," *International Journal of Computer Vision*, 2015.

BIOGRAPHY

***Gary Doran** received his Ph.D. in Computer and Information Sciences from Case Western Reserve University in 2013. He is currently a member of the Machine Learning and Instrument Autonomy Group at the Jet Propulsion Laboratory, California Institute of Technology, where his research focuses on processing data onboard spacecraft to identify science targets for autonomous instrument targeting or downlink prioritization.*



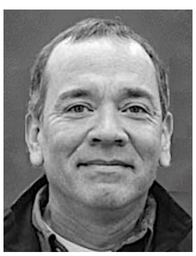
***Steven Lu** received his B.S. and M.S. degrees in Computer Science from Northeastern University in 2009 and California State Polytechnic University Pomona in 2015, respectively. He is a Data Scientist in the Machine Learning and Instrument Autonomy Group at the Jet Propulsion Laboratory, California Institute of Technology. Steven's current work focuses on applying machine learning and image processing algorithms to solve various Earth and Planetary Science problems.*

***Maria Liukis** received her Masters in Applied Mathematics from St. Petersburg State University, Russia, in 1992. She is currently a member of the Machine Learning and Instrument Autonomy Group at the Jet Propulsion Laboratory, California Institute of Technology, where she applies her software development skills to support research projects.*

***Dr. Lukas Mandrake** is a senior researcher and group supervisor of the Machine Learning (ML) and Instrument Autonomy group at JPL since 2002. His group brings ML techniques to bear against a wide range of applied problems in space, science, technology, and medicine. Our products have empowered flight missions (MSL Curiosity rover, MER Opportunity/Spirit rovers, EO-1, OCO-2/3, ...), advanced the state of the art for commercial customers like Kaiser, Statoil, and Chevron, were infused into numerous university settings, and addressed governmental agency needs (USDA, DOE, and DOD). Specifically, Mandrake specializes on scoping a question to make it answerable given available data, followed by creating ML-based solutions that are explainable and scientifically defensible (e.g. no "black box"). He is also the ML liaison to the Information and Data Science program office at JPL, representing rigorous ML to scientists, engineers, policymakers, and upper management. While working on the Orbiting Carbon Observatory (OCO-2) mission, he invented a unique system for estimating the uncertainty of observations called the Data Ordering Genetic Optimization (DOGO), 2^{nd} winner for the prestigious, NASA-wide Software of the Year award. Mandrake is also an enthusiastic and award-winning speaker and educator, bringing the promise of ML to new fields and applications as well as science education in* *general. He passionately works to educate children in science literacy and critical thinking, and has hosted hundreds of school children on interactive tours of JPL. He has received a BSE in Engineering Physics (U of AZ), a MS in Theoretical Plasma Physics (UCLA), and a PhD in Computational Plasma Physics (UCLA).*

Dr. Umaa Rebbapragada *is a Data Scientist in the Machine Learning and Instrument Autonomy group at the Jet Propulsion Laboratory, California Institute of Technology. Her work is focused on the infusion of state-of-the-art techniques from the machine learning community into large-scale data science systems, with a current focus on astronomical surveys. Dr. Rebbapragada received her Ph.D. in Computer Science from Tufts University.*

Kiri L. Wagstaff *received a Ph.D. in Computer Science from Cornell University in 2002, followed by an M.S. in Geological Sciences from the University of Southern California in 2008 and an MLIS in Library and Information Science from San Jose State University in 2017. She currently works at the Jet Propulsion Laboratory, California Institute of Technology, where she applies machine learning methods to problems from planetary science and astronomy and investigates new methods for spacecraft onboard data analysis and autonomy.*



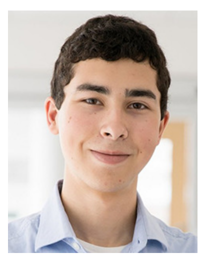
Jimmie Young *received his B.S. in Computer Science from California State Polytechnic University, Pomona in 1990. He is currently a member of the Computer Science for Data Intensive Applications Group at the Jet Propulsion Laboratory, California Institute of Technology where his work focuses on Science Data Systems (SDS) for Earth orbiting satellites.*

Erik Langert *received his M.S. in computer science from Binghamton University in 2017. He is currently employed at Palantir Technologies since September 2019. He currently works as a deployment strategist on the Business Development team.*

Anneliese Braunegg *is a software systems engineer specializing in machine learning. She interned at the Jet Propulsion Laboratory in the summer of 2018 and is now working full-time at The MITRE Corporation. In her spare time, Anneliese enjoys rowing, singing in ensembles, and playing board games.*

Paul Horton *is pursuing his Ph.D. in Exploration Systems Design (Systems Engineering) at Arizona State University and plans to graduate in 2023. His research is primarily focused on data science for planetary science and astronomy. He works closely with the Machine Learning and Instrument Autonomy Group at the Jet Propulsion Laboratory, California Institute of Technology, on multi-spectral anomaly detection with Curiosity and on board autonomy for balloon operations.*

Daniel Jeong *is a B.S. candidate in Computer Science at Columbia University. He is currently in his final year and plans to pursue a Ph.D. degree upon graduation. His research interests are in developing novel machine learning approaches for solving key problems in computational neuroscience/biology and space exploration.*



Asher Trockman *is pursuing his Ph.D. in Computer Science at Carnegie Mellon University. His research interests are in machine learning and optimization. He has previously worked on empirical software engineering in the Institute for Software Research at Carnegie Mellon University and deep learning in the Machine Learning and Instrument Autonomy Group at the Jet Propulsion Laboratory.*