ENABLING AUTONOMY IN COMMERCIAL-CLASS LUNAR MISSIONS

Virtual Conference 19–23 October 2020

Kaizad Raimalwala[1], Michele Faragalli[1], Evan Smal[1], Melissa Battler[1], Ewan Reid[1], Braden Stefanuk[2], Krzysztof Skonieczny[2]

[1]Mission Control Space Services, 162 Elm St. West, Ottawa, ON K1R 6N5, Canada
E-mail: kaizad@missioncontrolspaceservices.com

[2]Concordia University, 1455 De Maisonneuve Blvd. W., Montreal, QC, Canada, Email: bstefanuk@gmail.com

ABSTRACT

Early Lunar micro-rover missions will be short in duration and have constrained downlink capacity. To maximize the scientific return of these missions, Mission Control is developing technologies to autonomously classify geological features and detect novel features in rover camera imagery, which can be used to support intelligent decision-making for prioritizing data for downlink and instrument targeting. In a recently completed concept study, a trade-off analysis and performance evaluation were conducted for the terrain classifier and novelty detector algorithms across multiple datasets. The terrain classifier developed achieved accuracies of 77%-86% and Intersection over Union (IoU) scores of 0.667-0.680 across 10 different terrain type, on 3 distinct data sets (totalling 928 images), demonstrating the robustness of the approach to varying illumination conditions. In ongoing work, a comprehensive Lunar analogue dataset is being developed to continue prototyping, and the algorithms are being developed on an embedded processor for a flight demonstration opportunity.

1 INTRODUCTION

In this section, we begin with highlighting the challenges facing Lunar rover operations and the motivations for including technologies that enable autonomous perception and decision-making. In Section 2, we explain intended onboard and offboard use cases and User Experience (UX) considerations for how image annotations that are autonomously created by the terrain classifier can be used in a Lunar rover operations tactical cycle. In Section 3, we provide an overview of technologies with recent results from trade-off analysis of the terrain classifier on datasets. In Section 4, we highlight open challenges and next steps.

1.1 Challenges in Commercial Lunar Missions

Between the harsh Lunar environment and economic pressures on companies, early Lunar surface missions will be limited to a Lunar day (14 Earth days). and nominal operations at mid/low latitudes will likely be 10-12 Earth days. Payloads must also share a constrained downlink capacity. Payload operators generating high volumes of data face the problem of not receiving data in a timely fashion to influence their operations or worse, leaving valuable data on the Moon. These constraints are motivating the need for innovative concept of operations and technologies to ensure customer satisfaction, and the viability of this new model of exploration.

1.2 Mobility and Science Operations

Several factors are driving the need for autonomy in mobile science operations. In traditional Mars rover operations, visual surface characterization and subsequent analysis and decision-making takes place in day-long tactical cycles [1]. Upcoming Lunar rover missions, however, will have reduced latency, short lifetimes, and constrained bandwidth. This will result in a need for rapid tactical decision-making processes with limited data, leaving little time to analyze data, identify features of interest, and make decisions.

The highly anticipated NASA VIPER rover that will fly to the south polar region is a large rover (~300kg) but will have a constrained direct-to-Earth communications channel of 230 kbps [2]. Small-scale commercial Lunar rovers will also be constrained; a 10 kg rover deployed in Astrobotic's Mission One will be allocated 200 kbps according to standard payload data rate allocation advertised in their Payload User Guide (PUG) [3]. As per their CubeRover PUG, a 6kg payload will be allocated 60 kbps [4].

Sensing capabilities are growing increasingly powerful but data transfer rates are not sufficiently high to downlink high volume data in short decision-making timescales. To maximize scientific return, it will be important to have methods to intelligently compress or select data to downlink in real-time or to select key geological features to measure.

The nature of scientific discovery makes onboard autonomy compelling. It increases the chances of detecting valuable novel/sparse features that may otherwise be missed in scenarios that prioritize driving and other mission needs. For example, NASA's Opportunity rover was driven 600 ft past the Block Island meteorite, one of its biggest discoveries, before the science team discovered it and decided to drive back to investigate it [5].

With tactical cycles a few minutes long and pressure to achieve science objectives, missions will benefit from autonomy in data processing and decision-making. The ASAS-CRATERS (Autonomous Soil Assessment System: Contextualizing Rocks, Anomalies and Terrains in Exploratory Robotic Science) system developed by Mission Control offers such capabilities, with the goal of maximizing scientific return in upcoming missions [6].

1.3 State of the Art in Autonomous Perception

In a previous paper [7], we offered a detailed survey of modern perception and modeling technologies for planetary surface robotics. The state-of-the-art in terrain classification leverages high performance Convolutional Neural Networks (CNNs) that find natural features and complex patterns in the image.

For example, Soil Property and Object Classification (SPOC) [8] has a terrain classifier that uses Fully CNNs (FCNNs). Gonzalez and Iagnemma [9] recently published a comparative analysis of CNNs, Deep Neural Networks, and classical algorithms such as Support Vector Machines. These and other works have focused on classifying Mars surface images to improve autonomy for Mars rovers.

For Lunar applications, terrain classification motivated by scientific research has focused on crater detection using orbital data. Stepinski et al. and Chung et al. offer a review of traditional machine learning techniques, including SVMs [10], [11]. More recently, Silburt et al. [12] explored the use of CNNs to detect craters using a DEM merged from LRO and Kaguya data.

While these studies have successfully demonstrated the use of deep learning to improve terrain classification of images from Mars rover datasets or from a laboratory setting, only recent work by Mission Control has holistically studied terrain classification in a real-time system for a science-driven rover mission and its implications on mission operations [13]. Additional work, as presented in this paper, has demonstrated the use of this technology on Lunar datasets.

The Mission Control terrain classifier was first developed under the CSA-funded Autonomous Soil Assessment System [14]. In 2019, it was used onboard a rover to classify eight Mars-relevant terrain types in real-time at ~15 FPS as the rover drove at 20cm/s at a high-fidelity analogue site in Iceland (see *Figure 1*). This was a part of SAND-E (Semi-Autonomous Navigation for Detrital Environments), a NASA PSTAR funded project to inform Mars2020 operations [13].

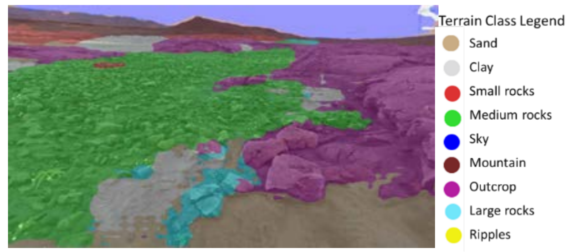

Figure 1: Result from field-testing the deep-learning based terrain classifier in Iceland.

Recent work by Kerner et al. [15], [16] have demonstrated the capability to detect novel geological features in multispectral images of the Martian surface. They show that a spatial-spectral error map can enable both accurate classification of novelty in multispectral images as well as human-comprehensible explanations of the detection.

2 USE CASES AND USER EXPERIENCE IN LUNAR ROVER OPERATIONS

In this section, we focus on the use cases of the perception algorithms (terrain classification and novelty detection) in onboard applications to enable autonomous behaviour to benefit science and navigation alike, and offboard applications to support science backroom tasks.

2.1 Onboard Applications

In Lunar rover operations, we expect two types of data collection modes for onboard science instruments, in particular those that are scanning the surface:

Continuous Mode: The instrument is set to capture data at a set time or distance-based rate. This mode is relevant during scouting operations or driving between destinations. Instrument measurements are desired at whatever rate is feasible to downlink given other payload and telemetry allocation.

We assume a stronger constraint on downlink capacity than on available power for data capture, i.e. more data can be captured and stored onboard than is possible to downlink in real-time. In this case, we must maximize the scientific value of this data.

Targeted Mode: The instrument is triggered to capture data at specific instances due to higher power constraint that does not make continuous measurements feasible. It is desired to target the instrument to scan specific and high-priority features of interest.

In both modes, autonomous perception can support onboard decision-making, i.e. intelligent downlink of continuous data in the first mode, or intelligent targeting of the instrument in the second mode.

With a semantic understanding of the nearby terrain, and using science operator defined rules, the onboard system can decide whether the instrument data is suitable for immediate downlink. Similar to the AEGIS system pioneered for Mars rovers [17], we intend to use strategies such as key target signatures, novelty detection, and representative sampling.

Additionally, as a semantically useful terrain representation, the terrain classifier outputs can be used by path planning algorithms to enable autonomous and intelligent navigation to scientific targets of interest.

2.2 Offboard Use Cases

Consider a driving scenario where the rover is driven to a prioritized destination without planned science stops. The science backroom team is tasked to continuously monitor camera feed to assess the terrain for any high-priority feature that are worth a short stop and quick inspection, e.g. a high-resolution image or instrument scan. NavCam images are compressed and downlinked at about 1Hz, and so the image quality is not optimal for scientific terrain assessment.

Consider the following mission parameters. The rover is driving at 8 cm/s with navigation stops every 5m. From the time a particular image is taken, the call to stop the rover to analyze a feature must be done before the rover drives 3m ahead such that that feature is still within a reasonable range; this translates to 37.5 seconds later if the rover is driving continuously during those 3m. Accounting for system latency and data processing time, the science team has approximately 30s to look at the image and: i) decide to annotate a particular feature, ii) annotate it, and iii) communicate it to the rover operations team.

While numbers may change, this mission scenario is expected to be common for early commercial Lunar micro-rover operations, and as such, autonomous and intelligent image annotation can greatly assist in the process of scientific terrain assessment in these rapid tactical cycles. The following is a summary of benefits with autonomously annotating features in real-time:

- Efficient evaluation and selection of a feature to request a navigation stop or instrument targeting, followed by communication of the request internally within the operations team.
- Features can be catalogued in a database, enabling feature-based query in real-time which can be highly beneficial in short-duration missions. E.g., an operator can quickly retrieve all fresh craters of size 3-4m in a specific geographic area.
- Features can be projected on a map frame, and map-based data products can be easily integrated into GIS tools for rapid analysis with the context of scale and other information layers derived *in situ* or from orbit.

The rapid classification and cataloguing of Lunar surface features supports analyses, e.g. crater counting and size-frequency distribution estimation. This can help scientists to inform their models and hypotheses that might guide decisions within the 10-12 Earth day mission timeframe.

To test these hypotheses on UX design, we will work with Lunar scientists with an interest in operations and conduct field tests, followed by working with operations personnel on targeted flight demonstrations.

3 OVERVIEW OF TECHNOLOGIES AND EARLY RESULTS

3.1 Terrain Classification

The terrain classifier uses a CNN in an encoder-decoder architecture to perform semantic segmentation. In the recently completed ASAS-CRATERS concept study, a detailed trade-off analysis and performance evaluation was conducted for the terrain classifier, using three datasets (see the figure below for example images): two pre-existing labeled datasets, one from the CSA-MET (Canadian Space Agency Mars Emulation Terrain), and one from Iceland as part of the 2019 SAND-E field test campaign. The third dataset was compiled and labeled during the ASAS-CRATERS concept study; it consists of images from the Chang'E-3 Yutu-1 PCAM (Panoramic Camera) and the Chang'E-4 lander-based TCAM (Terrain Camera). Table 1 highlights the simple classification scheme used for early prototyping.

Figure 2: Left to right: Images from the Chang'E, CSA-MET, and SAND-E datasets.

Table 1: Simple Lunar terrain classification scheme defined for early trade-off analysis and prototyping

Sky	Black region on top of the image
Crater	Any distinct crater
Boulder	Any distinct rock
Crater Field	Groups of craters
Rock Field	Cluster of rocks, ejecta, etc.
Regolith	Regolith interspaced between rocks
Hills	Visible in the background

To choose an architecture for the terrain classifier, a trade-off analysis was completed, comparing the performance of five different encoder-decoder architectures. Two are variations of DeepLabV3+ [18], one of which uses a truncated pre-trained MobileNetV2 model as the feature extractor in the encoder module and the other uses a truncated pre-trained Xception model. The remaining three architectures, which we refer to as SmallNet, TinyNet and MiniNet, are novel encoder-decoder architectures of three sizes. All three are relatively small compared to most state-of-the-art CNNs for semantic segmentation; they use a truncated pre-trained MobileNetV2 as the encoder module and a novel decoder module. All five architectures use ImageNet as the pre-training dataset.

The five architectures were tested on the CSA-MET and SAND-E datasets at three different input sizes: 96x96, 128x128 and 224x224, all downsized from the original images. The input size affects classifier accuracy, resolution, and memory and downlink requirements. The tests were done with GPU implementations of the CNNs. The top two architectures on the CSA-MET and SAND-E datasets (DeepLabV3+-MobileNetV2 and DeepLabV3+-Xception) were selected for evaluation on the Chang'E dataset (see figures below for results). With the 224x224 input, they achieved accuracies of 75.5% and 76.9%, and mean IOU of 66.3% and 66.5%.

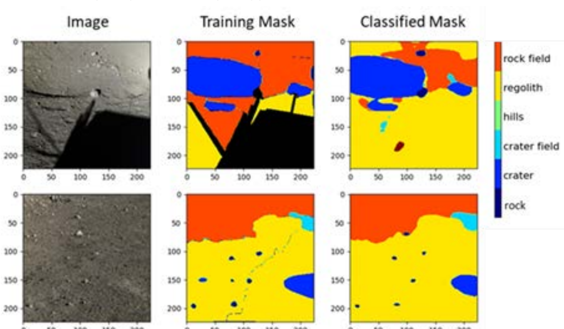

Figure 3: Two Chang'E image examples classified using DeepLabV3+-MobileNetV2. The black mask indicating the rover shadow was not used in classification.

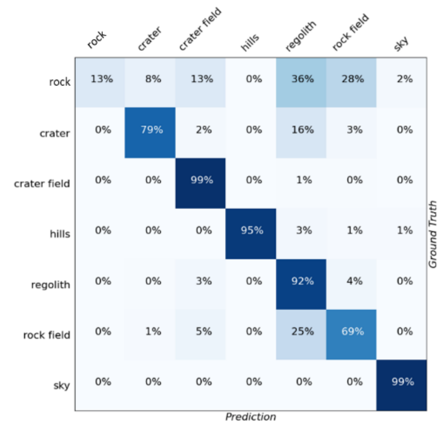

Figure 4: Confusion matrix for the Chang'E dataset, using DeepLabV3+-MobileNetV2 with an 224x224 input size image.

The loss function to minimize during training has a significant impact on the network's performance. We trained the CNNs on two datasets using the standard crossentropy (CE), weighted CE, focal CE [19], dice loss, focal dice loss, a combination of dice loss and CE (referred to as combo loss), tversky loss (also called generalized dice loss), focal tversky loss and Lovasz-softmax loss. We then selected the top four functions from testing on the CSA-MET and SAND-E datasets: CE, weighted CE, combo loss (CE+Dice) and focal CE. In our study the best performing loss, with both the highest accuracy and mean IoU, was the focal CE. The following table and figure show some results of the loss function trade-off analysis.

Table 2: Loss Function analysis on the SAND-E dataset using DeepLabV3+-MobileNetV2

Loss Function	Accuracy	Mean IOU
CE	75.3%	62.6%
Weighted CE	67.5%	50.9%
CE+Dice	68.4%	55%
Focal CE	**76.1%**	**65.2%**

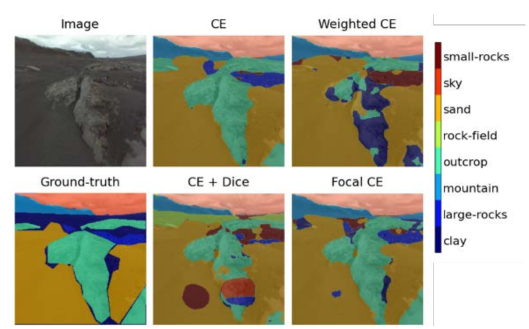

Figure 5: An image from SAND-E classified by DeepLabV3+-MobileNetV2 with four loss functions.

3.2 Novelty Detection

The purpose of the novelty detection algorithm is two-fold: 1) to detect known geological features that are rare to find and have limited data for training, and 2) to detect anomalous features that are not consistent with typical surface features as defined in a curated training dataset.

For unknown novelties, the novelty detection algorithm is implemented using semi-supervised deep learning methods consisting of a convolutional autoencoder for input reconstruction and a binary classification network for identifying novel data. A convolutional autoencoder is used to generate error maps that are then fed to a classification network determines the input to novel or typical.

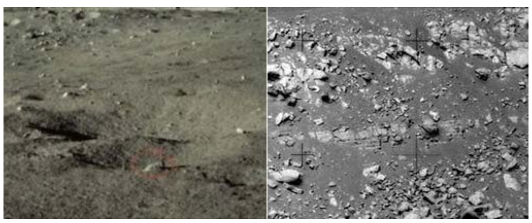

Figure 6: Left: A 'low-erosion' rock flagged as anomalous by the Yutu-2 operations team. Right: Outcrop of bedrock seen by Apollo 15 astronauts.

For known novelties, i.e. those that are well-documented or well-described but occur rarely or are outliers in the given setting, we may also consider a supervised learning approach with limited training data, typically referred to as one-shot (or few-shot) learning approaches. Examples of novelties that could be detected using this approach include mantle xenoliths in basalt, outcrops of bedrock, meteorites, secondary craters, and pyroclastic material with colouration.

To evaluate the semi-supervised novelty detection approaches using convolutional autoencoder architectures (developed by Kerner et al.), two datasets were studied: the Chang'E dataset labeled with 'novel' features during the ASAS-CRATERS concept study (see *Figure 7*), and an additional dataset of Mars images from MSL made publicly available by Kerner et al [16]. A parallel paper at this conference by Stefanuk et al. includes detailed results on the work done to evaluate the novelty detector approaches on both datasets.

Due to the limitations of the Lunar dataset and the definitions of novel features, we discovered several challenges with evaluating the semi-supervised approach on the Lunar dataset. Challenges and next steps are discussed in Section 4.

Figure 7: Example novel features labeled in the Chang'E dataset.

3.3 Feature Mapping

A feature mapping system can aggregate classifier results on map tiles. This is useable by onboard algorithms for instrument targeting and navigation, and offers easy integration into GIS tools for analysis with geospatial context and other mission data. *Figure 8* shows a concept diagram of this mapping.

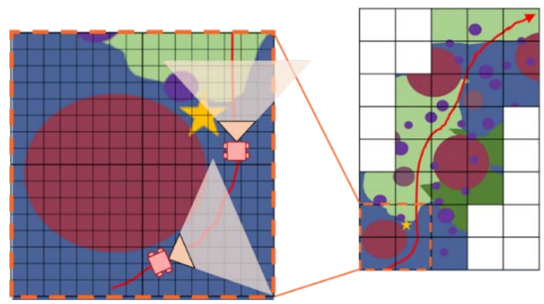

Figure 8: Illustration of aggregating outputs from the terrain classifier onto map tiles.

The terrain classifier and novelty detector provide outputs in the image coordinate frame. A set of algorithms can then project these results onto a map frame and aggregate them onto a uniform map-tile data product, similar to an occupancy grid used in rover navigation. This has several benefits.

First, it acts as an ensemble method to improve classifier accuracy and eliminates contradictory feature outputs in overlapping classified images. Multiple classification instances of a specific feature at a spatial point can 'smooth out' the single-instance outputs of the terrain classifier and improve classification accuracy.

Second, it better supports scientific analysis. Map tiles more readily offer geospatial context of the classified features and can be more easily integrated into GIS tools and overlaid on data from multiple sources as there is conformity in the same reference frame.

Third, the map tiles can be used for science and navigation autonomy, e.g. autonomous instrument targeting and path planning.

To enable the mapping, the system would require a source of relative localization and a depth image that is co-registered with the classified frame. A stereo-processing pipeline can generate this information but requires additional onboard software which adds complexity and increases computation. While it is desired to implement this as part of the onboard flight suite for a future mission, a contingency option is to implement this as part of the ground segment data processing suite. On the ground, the map tiles can be geo-tagged with available global localization data, along with other derived data products.

3.4 Feature Prioritization for Intelligent Downlink or Targeting

Once an onboard system can infer some knowledge of the surrounding terrain's geological features, it can also be programmed with some decision-making capabilities. Inspired from methods used in Mars rover missions [17], we define some methods to leverage the outputs of the terrain classifier and novelty detector to inform intelligent prioritization for data downlink or instrument targeting:

1) Novel features as identified by a semi-supervised detector, or a supervised detector for known but rare features.

2) Representative sampling: Features which contribute to build a representative sample of specific features and characteristics, identified using the classes present in the terrain classifier output and the expected distribution of classes in the region. For example, the science team may want measurements of instrument data taken of an equal number of 'small', 'medium', and 'large' craters (relative size), or an equal number of 'fresh', 'semi-degraded', or 'ghost' craters.

3) High-priority classes, such as a fresh crater or rocks of a specific type. Depending on which classes are defined for model training, this functionality could include prioritizing images containing rocks of high/low albedo or angularity.

Scientific data itself (e.g. from a neutron spectrometer or multi-spectral imager) can also be analyzed for novel or high-priority feature detection.

3.5 High-Performance COTS Processors

We have completed a preliminary feasibility assessment for and are prototyping algorithms for development on high-performance processors by Xiphos Technologies. Primarily, we consider the Q8S, which is a compact, low-cost, high-performance, COTS, and flight-qualified processor that is suitable for computationally intensive algorithms for planetary science and robotics missions. The Q8S is 90g, 4-25W, and 85.8x80x22.6mm. At the core of the Q8S is a high-performance, low power Xilinx Zynq Ultrascale+ Multiprocessor System-on-Chip (MPSoC) FPGA, which enables high-performance parallel computing.

Using off-the-shelf components will lower costs for development and enable commercialization. For more physically and power constrained platforms, we also consider the Xiphos Q7S (24g, 1W, and 78x43x9mm).

In the ASAS-CRATERS concept study, the algorithm performance was evaluated on the Q7S and Q8S. We tested the DeepLabV3-MobileNetV2 architecture at the 224x224 input size directly on the ZCU102 board with the Xilinx Vitis-AI toolkit which achieves a DPU processing time of 0.004s (250 Hz) with additional post processing time of 0.024s. In total, the end to end throughput for this mode was ~35 Hz. This could be improved dramatically by moving the final softmax layer from the CPU to the FPGA. The high throughput and low latency of this profiling analysis demonstrates feasibility on a flight-qualified processor.

A preliminary radiation assessment has also been completed for the Q8S. Requirements were identified through radiation modelling; a maximum Total Ionizing Dose (TID) of 6 krad was determined to be the worst-case dosage of a Lunar mission, assuming nominal shielding thickness provided by the lander and rover during transit and surface operations. Low Dose Rate (LDR) Radiation testing of the Q8S platform has been conducted under biased and unbiased EQM to 30 krad total dose. The Q8S remained fully operational until 20 krad with a Low Dropout Regulator failure at 20krad. Single Event Effects (SEE) Proton Testing performed at TRIUMF 105 MeV where 4 different tests were completed to 10^{10} total fluence with flux ranging from 10^6 to 10^7 protons/sec. No destructive latch-up events were detected during proton testing

This initial testing indicates that the Q8S is suitable for a Lunar mission.

4 DISCUSSION

4.1 Open Challenges

Autonomous Lunar terrain classification is a challenging endeavor. A balance must be achieved between performance, fidelity of the classification scheme to detect increasingly complexity, and the constraints of deep space computing systems and architectures.

The quality of the training dataset is a primary factor for the performance of any deep learning model. In particular, the training dataset quality enforces an upper bound on both the quantitative and qualitative performance. Through the lessons learned during early prototyping, we have compiled several recommendations for constructing datasets for terrain classification using ASAS-CRATERS, and identified challenges to address in future development

The minimally acceptable accuracy and/or mean IoU metrics depend on the types of features defined for classification. Certain scenarios such as identifying macro homogeneous terrain like regolith and rock fields may be able to tolerate coarse semantic estimates which could be refined using a simple image erosion operation. Other scenarios where fine segmentation of small interspersed craters and rocks may require more accuracy and higher IoU models to be effective during the mission operations. We can also improve classification results by removing labels such as ‘crater field’, which can be assigned as a meta-label in post-processing a group of craters that meets a certain density threshold is detected.

For novelty detection, challenges remain in characterizing the performance of convolutional autoencoders in Lunar applications. While examples of features have been defined, the ‘novelty’ of a feature is also dependent on the context of the landing site, e.g., a volcanic feature found at a highland site. Finding known novel features such as bedrock outcrop may also be better suited for supervised approaches such as one-shot learning. Lastly, in many cases, the definition of novel features is subjective to the interests and objectives of specific mission scenarios.

4.2 Next Steps

To advance state of the art of terrain classification and novelty detection for Lunar surface applications, a high quality and comprehensive dataset that represents the views from a Lunar rover is required. This should cover varying environmental conditions in terms of lighting angles and resulting shadows on the terrain, the varying types and layouts of terrain and geological features, the presence of novel or anomalous terrain features, and artificial elements such as the rover’s own self, shadows, and wheel tracks. To build such a comprehensive dataset, we are currently building a controlled Lunar analogue environment where these variables can be introduced.

Once such a dataset is acquired, we will train and validate deep learning models to complete the prototyping phase of ASAS-CRATERS. In future development, we plan to increase the complexity of classifier features to improve the classifier capability. This includes adding attributes to features such as the degradation state of a crater or the tone of the regolith.

Following the prototyping phase, these technologies will be integrated into analogue missions to validate their operational hypotheses and use cases. This includes the SAND-E 2021 field season, which offers an opportunity to test the technologies at a high-fidelity Moon and Mars natural analogue site in Iceland.

The terrain classifier is currently at TRL 5 and is undergoing prototyping on a flight-qualified embedded processor. A flight demonstration is targeted for missions in 2022 and 2023. In the first flight demonstration, the terrain classifier will be embedded onto a Xiphos processor integrated on a Lunar lander, while the other technologies will be integrated into the ground segment to demonstrate the feasibility and usability in a real Lunar mission.

6 CONCLUSION

In this paper, we presented results from early prototyping of the terrain classifier algorithm developed by Mission Control for Lunar applications. Based on these results, we choose the DeepLabV3+ architecture with a truncated pre-trained MobileNetV2 model as the feature extractor in the encoder module. The MobileNetV2 encoder is selected due to its highly efficient design for mobile processors. We introduced the targeted flight processors and highlighted key results from algorithm profiling and radiation assessment for a mission one Lunar day long.

We also provided a discussion on the problem of novelty detection for Lunar applications. For both perception technologies, we provided an overview of the benefits and use cases to augment the autonomy of micro-rover operations. In the flight segment, autonomous perception enables intelligent decision-making, such as prioritizing data for real-time downlink when science instruments are continuously collecting data, and prioritizing features for the science instrument to target when continuous data collection is not feasible. In the ground segment, autonomous classification, mapping, and cataloguing of features can enable real-time analysis and feature-based query. This can help the science operations team update models and hypotheses to adapt decision-making processes during the mission itself, which may be 10-12 Earth days long.

Acknowledgement

The authors thank the Canadian Space Agency for funding this technology, and the ASAS-CRATERS partners for science and technology support: Dr. Ed Cloutis at the University of Winnipeg, Dr. Gordon Osinski and Dr. Kenneth McIsaac at Western University, Dr. Ryan Ewing at Texas A&M University, and Xiphos Technologies.

References

[1] D. Gaines *et al.*, “Productivity challenges for mars rover operations,” in *Proceedings of 4th Workshop on Planning and Robotics (PlanRob)*, 2016, pp. 115–125.

[2] A. Colaprete, “An Overview of the Volatiles Investigating Polar Exploration Rover (VIPER),” presented at the Missions and Instruments Virtual Workshop, SSERVI, Feb. 07, 2020, [Online]. Available: https://youtu.be/mUnkY_Rwo44.

[3] “Astrobotic - Payload User Guide.pdf.” Aug. 2019.

[4] “Astrobotic CubeRover - Payload User Guide.” Apr. 2020.

[5] “NASA - Meteorite Found on Mars Yields Clues About Planet’s Past.” https://www.nasa.gov/mission_pages/mer/news/mer-20090810.html (accessed Sep. 18, 2020).

[6] K. Raimalwala, M. Faragalli, E. Reid, E. Smal, and M. Battler, “Autonomous Soil Assessment System: Contextualizing Rocks, Anomalies and Terrains in Exploratory Robotic Science (ASAS-CRATERS),” presented at the Lunar Surface Science Workshop 2020, Denver, CO, Apr. 2020, Accessed: Mar. 03, 2020. [Online].

[7] K. Raimalwala, M. Faragalli, E. Reid, and E. Smal, “Machine Learning Applications for Safe and Efficient Rover Mobility Operations and Planning,” presented at the International Astronautical Congress (IAC), Bremen, Germany, Oct. 2018.

[8] B. Rothrock, R. Kennedy, C. Cunningham, J. Papon, M. Heverly, and M. Ono, “SPOC: Deep Learning-based Terrain Classification for Mars Rover Missions,” presented at the AIAA SPACE 2016, Sep. 2016, doi: 10.2514/6.2016-5539.

[9] R. Gonzalez and K. Iagnemma, “DeepTerramechanics: Terrain Classification and Slip Estimation for Ground Robots via Deep Learning,” *arXiv:1806.07379 [cs]*, Jun. 2018, Accessed: Aug. 17, 2018. [Online]. Available: http://arxiv.org/abs/1806.07379.

[10] T. F. Stepinski, W. Ding, and R. Vilalta, “Detecting impact craters in planetary images using machine learning,” in *Intelligent data analysis for real-life applications: theory and practice*, IGI Global, 2012, pp. 146–159.

[11] C. P. Chung, C. O. Chung, and C. H. Yam, “Lunar Crater Identification from Machine Learning Perspective,” *Acta Futura*, p. 8, 2014.

[12] A. Silburt *et al.*, “Lunar Crater Identification via Deep Learning,” *Icarus*, vol. 317, pp. 27–38, Jan. 2019, doi: 10.1016/j.icarus.2018.06.022.

[13] R. C. Ewing *et al.*, “Overview and Initial Results of SAND-E: Semi-Autonomous Navigation for Detrital Environments,” presented at the Lunar and Planetary Science Conference, Woodlands, Texas, 2020.

[14] M. Faragalli, K. Skonieczny, and K. Raimalwala, “Predicting Non-geometric Hazards Using Machine Learning for Safer Rover Mobility Operations,” presented at the International Symposium on Artificial Intelligence, Robotics and Automation in Space (i-SAIRAS) 2014, Jun. 2018.

[15] H. R. Kerner, D. F. Wellington, K. L. Wagstaff, J. F. Bell, C. Kwan, and H. Ben Amor, “Novelty Detection for Multispectral Images with Application to Planetary Exploration,” *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, pp. 9484–9491, Jul. 2019, doi: 10.1609/aaai.v33i01.33019484.

[16] H. R. Kerner *et al.*, “Comparison of novelty detection methods for multispectral images in rover-based planetary exploration missions,” *Data Mining and Knowledge Discovery*, Jun. 2020, doi: 10.1007/s10618-020-00697-6.

[17] R. Castano *et al.*, “Oasis: Onboard autonomous science investigation system for opportunistic rover science,” *Journal of Field Robotics*, vol. 24, no. 5, pp. 379–397, May 2007, doi: 10.1002/rob.20192.

[18] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, “Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation,” 2018, pp. 801–818, Accessed: Jun. 29, 2020. [Online]. Available: https://openaccess.thecvf.com/content_ECCV_2018/html/Liang-Chieh_Chen_Encoder-Decoder_with_Atrous_ECCV_2018_paper.html.

[19] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollar, “Focal Loss for Dense Object Detection,” 2017, pp. 2980–2988, Accessed: Jun. 26, 2020. [Online]. Available: http://openaccess.thecvf.com/content_iccv_2017/html/Lin_Focal_Loss_for_ICCV_2017_paper.html.