Check for updates

RESEARCH ARTICLE

Design of robotic traverses on the Archytas Dome on the Moon

[version 1; peer review: 1 approved, 1 approved with reservations]

Lewis Lovell[1], Isabella C. Adriani[2], Giacomo Nodjoumi[3], Javier E. Suarez-Valencia[3], Daniel Le Corre[4,5], Anita Heward[6], Angelo Pio Rossi[3], Nick L. J. Cox[5]

[1]School of Environmental Sciences, University of East Anglia, Norwich, England, NR4 7TJ, UK
[2]University of Rome La Sapienza, Rome, Lazio, 00185, Italy
[3]Constructor University, Bremen, 28759, Germany
[4]Centre for Astrophysics and Planetary Science, University of Kent, Canterbury, England, CT2 7NH, UK
[5]Centre d'Etudes et de Recherche de Grasse, Grasse, 06130, France
[6]Dill Faulkes Educational Trust, London, England, EC3M 6BN, UK

v1 **First published:** 13 Jun 2024, **4**:116
https://doi.org/10.12688/openreseurope.17424.1
Latest published: 13 Jun 2024, **4**:116
https://doi.org/10.12688/openreseurope.17424.1

Abstract

Background

In recent years, we have seen renewed efforts to study and explore the Moon; modern techniques like machine learning can be important in this context, especially in recognising and classifying the lunar surface. The EXPLORE Machine Learning Lunar Data Challenge was a public initiative during the last quarter of 2022. Its objective was to encourage participants to apply machine learning techniques to identify potential hazards for a planetary mission and to design a robotic traverse for exploring the lunar surface.

Methods

The lunar region targeted by the challenge was the Archytas Dome in Mare Frigoris, a location with a varied geology and a potential zone for future exploration. We provided training datasets of craters and boulders to the participants, who used them to complete the three steps of the challenge: creating a model that detects these landforms, applying these models to the Archytas Dome region, and computing a traverse for optimal exploration of the zone. In this paper, we showcase the results and considerations of the team that won the challenge. The first step was to enhance the training data by

Open Peer Review

Approval Status ? ✓

	1	2
version 1 13 Jun 2024	? view	✓ view

1. **Shreya Santra**, Tohoku University, Sendai, Japan

2. **Mohammad Mehdi Taghavi Nasrabadi**, University of Houston, Houston, USA
Heidar Malki, University of Houston, Houston, USA

Any reports and responses or comments on the article can be found at the end of the article.

generating new labels and resizing the existing ones. The original and the improved dataset were then used to train four iterations of a neural network model.

Results

The model with the enhanced dataset yielded the best scores when applied to the Archytas Domes zone (75.46\%). Finally, the traverse was calculated using proximity analysis while avoiding steep slopes and dangerous landforms.

Conclusions

We found that the variations between tasks and the different approaches necessary to solve them turned out to be the major difficulty of the challenge, as it required backgrounds in both remote sensing and computer sciences. This was reflected in the low participation and the multidisciplinary of the members of the winning team.

Plain language summary
The EXPLORE Machine Learning Lunar Data Challenge was a public activity launched by the EXPLORE consortium during the last quarter of 2022. Its objective was to challenge the scientific community to use machine learning techniques to face problems expected in lunar exploration, including creating an optimal path that a rover or human can follow. The exercise was focused on the Archytas Dome on Mare Frigoris, an interesting location for lunar exploration. The challenge consisted of three steps: creating a machine learning model to detect craters and boulders, applying that model to detect these landforms in the Archytas Dome region, and computing an optimal path for exploring the area. We present the results of the team that won the challenge and discuss the feasibility of the project and the problems faced in its development. We found that the challenge was excessively long and complex, explaining the low participation and the composition of the winning team, which was comprised of members with a remote sensing and computer sciences background.

Keywords
Data Methods, Machine Learning, Lunar Exploration, Data Challenge, Lunar Domes, Lunar Traverse

H2020 This article is included in the Horizon 2020 gateway.



This article is included in the Machine Learning collection.

Corresponding author: Lewis Lovell (l.lovell@uea.ac.uk)

Author roles: Lovell L: Formal Analysis, Investigation, Methodology, Writing – Original Draft Preparation, Writing – Review & Editing; **Adriani IC**: Formal Analysis, Investigation, Methodology, Writing – Original Draft Preparation, Writing – Review & Editing; **Nodjoumi G**: Conceptualization, Data Curation, Writing – Original Draft Preparation, Writing – Review & Editing; **Suarez-Valencia JE**: Conceptualization, Data Curation, Writing – Original Draft Preparation, Writing – Review & Editing; **Le Corre D**: Conceptualization, Data Curation, Writing – Original Draft Preparation, Writing – Review & Editing; **Heward A**: Writing – Review & Editing; **Rossi AP**: Writing – Review & Editing; **Cox NLJ**: Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004214 (Innovative Scientific Data Exploration and Exploitation Applications for Space Sciences [EXPLORE]).
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2024 Lovell L *et al*. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Lovell L, Adriani IC, Nodjoumi G *et al.* **Design of robotic traverses on the Archytas Dome on the Moon [version 1; peer review: 1 approved, 1 approved with reservations]** Open Research Europe 2024, **4**:116 https://doi.org/10.12688/openreseurope.17424.1

First published: 13 Jun 2024, **4**:116 https://doi.org/10.12688/openreseurope.17424.1

1 Introduction

In recent years, new efforts to explore and study the Moon have been ramping up, exemplified by ongoing missions from several space agencies, like the Artemis program developed by NASA and ESA[1]. Although human missions are considered at later stages of these programs, using rovers to scout the surface remains the preferred method of exploration. Artificial intelligence (AI) is also expected to play a major role in these missions, as it can improve decision-making and traverse computations for a robotic mission on site[2].

Within this framework, The EXPLORE Machine Learning Lunar Data Challenge was developed by the team of the EXPLORE consortium during the last quarter of 2022. The challenge required its participants to generate robotic traverses across a hypothetical lunar landing site, using AI and other computational techniques to survey the landscape and identify potential hazards. Here, we will briefly discuss the technical aspects of the challenge, to later describe and discuss the methodologies and results of the winning team. This work was developed as part of the EXPLORE project, whose main goal is to develop platforms that promote exploring space science data using AI and visual analytic tools.

1.1 Lunar exploration

Robotic instruments have been a cornerstone of lunar exploration since the initial stages of space flight. The first instrument that achieved a soft landing on the Moon was Surveyor 1 in 1967[3]. This was simply a landing test, but even then, the instrument had a set of instruments that fed a central computer to control the landing, showing the importance of feedback and machine decision-making processes. During the era of the Apollo missions, the main focus was human exploration, and the vehicles were designed to be used by astronauts and not for autonomous control[4]. Nevertheless, during this period, in 1970, the Soviet Union developed the Lunokhod 1, the first moving instrument on the surface of the Moon[5]. The following years saw a rapid stagnation of lunar exploration, and when interest and activities were renewed at the end of the last century, the means of exploration shifted heavily from landers to orbiters. Several missions have been placed around the Moon in the last two decades, such as the Lunar Reconnaissance Orbiter (LRO)[6], with the main advantages of having bigger coverage and time spans than rovers, greatly enhancing our knowledge of the lunar surface and subsurface.

In the last few years, interest in rovers has revived. The renewed intentions of taking humans to the Moon will be preceded by numerous rovers scouting the surface. A recent example is the rover Yutu-2, part of the Chang'e 4 Chinese mission[7]. This rover is the first to explore the far side of the Moon and is part of a more ambitious exploration program[8]. Upcoming vehicles are also being prepared, like the VIPER mission to the lunar south pole[9]. Finally, the private sector is starting to play an increasing role in exploration with missions such as ispace's Hakuto-R Mission 1 carrying a rover from the Emirates Lunar Mission[10]. Overall, developing new ideas and technologies will be extremely important for the future of space exploration.

1.2 Machine Learning based planetary mapping

The application of Artificial Intelligence (AI), Machine Learning (ML), Deep Learning (DL), and Neural Networks (NN) in object detection extends significantly into areas of analysing satellite images, offering capabilities for monitoring and understanding planetary surfaces. By leveraging these technologies, it becomes possible to automatically identify surface features, both natural and man-made, from space with impressive accuracy and efficiency[11–14]. Deep Learning, in particular, has shown remarkable success in satellite imagery analysis due to its ability to handle the vast complexity and variability of data captured from orbit. Neural networks, especially Convolutional Neural Networks (CNNs)[15–18], are adept at processing images often present at different spatial and spectral resolutions. These networks can distinguish subtle features across different terrains and conditions, making them highly effective for tasks ranging from mission planning to hazard detection and surface characterization. Integrating these technologies in satellite image analysis enables the automation of previously labour-intensive and time-consuming tasks, such as surface classification, change detection over time, and resource mapping. This capability is crucial for addressing global challenges in planetary sciences like crater counting[19,20], boulder detection[21,22], and more general surface segmentation[23–25]. By training neural networks on large datasets of labelled satellite images, these systems learn to detect objects and patterns of interest with high precision, opening up new avenues for research, policy-making, and commercial applications. The ability to process and analyse satellite imagery with AI and ML thus represents a significant leap forward in our capacity to observe and understand the planetary surfaces from a global perspective.

1.3 The lunar data challenge

Data challenges are useful to evaluate current and future space exploration problems as they seek to harness the scientific community to solve practical questions. For that reason, the EXPLORE Machine Learning Lunar Data Challenge was launched in 2022, in which the participants were challenged to use ML techniques to analyse the geological hazards of a potential lunar landing site and create a viable traverse for a robotic mission.

Since the challenge aimed to go beyond just processing data, three steps were established for the participants to follow in order to complete it successfully. The first step consisted of creating an object detection model for lunar landforms. In order to do this, the participant selected, created and trained a deep-learning model that could recognize the target landforms. For this challenge, we decided to focus on detecting craters and boulders since they would be the most immediate hazards for a rover. The second step was to apply the trained and verified model to the target location on the Moon, detecting the desired landforms. Finally, in the third stage, the participants integrated the prior results to generate a realistic traverse across some previously defined Regions of Interest (ROI).

To evaluate the results of the first steps, we reviewed the trained models, confusion matrices, and training scores. For the other two steps, we checked manually mapped landforms

and traverses with the resulting labels, maps, and traverses developed by the participants.

The data challenge was open from September to December of 2022. In this work, we will showcase the methodology, results and discussion arising from the winning team of the challenge, as well as some considerations about the challenge itself.

2 Data selection and preparation

To accomplish the objectives of this ML data challenge, we decided to prepare two sets of data with different purposes. The first dataset is related to the data necessary to train the model in the first stage, with images collected from existing databases from scattered locations on the Moon. The second set of data corresponds to the zone selected to apply the model during the second and third steps of the model, the Archytas Dome in Mare Frigorifis (0.71° E, 55.71° N).

2.1 Selection of the region

As this lunar data challenge aimed to highlight these systems' usefulness in exploring the Moon, it was essential to select a region that could be interesting to host a real robotic mission. Hence, the location should have a reasonable extension, good geological diversity, and scientifically interesting features. Furthermore, good data coverage and spatial resolution were needed to fully use the capacities of the ML techniques.

In the search for geological diversity, we decided to explore some landforms not usually studied in detail, such as intrusive igneous domes. They are the result of magma that cooled below the surface, eventually creating a noticeable landform due to its ascent through the crust or via faulting[26]. A list of intrusive domes was compiled by Lena *et al.*[27], from which we selected the Archytas Dome as our study case since it fulfils the parameters described above.

The Archytas Dome is located south of Mare Frigoris in the vicinity of the rim of the Imbrium Basin (see Figure 1 for a simplified geological map of the zone). Similar to other intrusive domes on the Moon, it has a gentle slope of 0.25°, a low altitude of 70 m, and a large diameter of 33 km[27]. The structure may be the result of the flexural deformation made by an ascendant intrusion, which stagnated as a laccolith near the surface[28]. The geologic diversity near the dome is high; close to the dome lies an elongated rille and a concentric crater, the latter being an uncommon type of crater with a poorly delineated inner rim and a flat floor. Wöhler and Lena[28] suggest that this may be the result of an impact event during the emplacement phase of the laccolith. Furthermore, south of the dome, there is a sinuous rille that travels downhill from the Imbrium Basin rim, adding a more highlands-like set of rocks to the zone.

The selected location is 68 km large and 12 km wide, which is a reasonable coverage for a modern long-term robotic mission, e.g. Farley *et al.*[29]. All the geologic units mapped are dated as Imbrian (3.85 Gyr – 3.2 Gyr)[30]; hence, the region is heavily cratered and has experienced few resurfacing events.

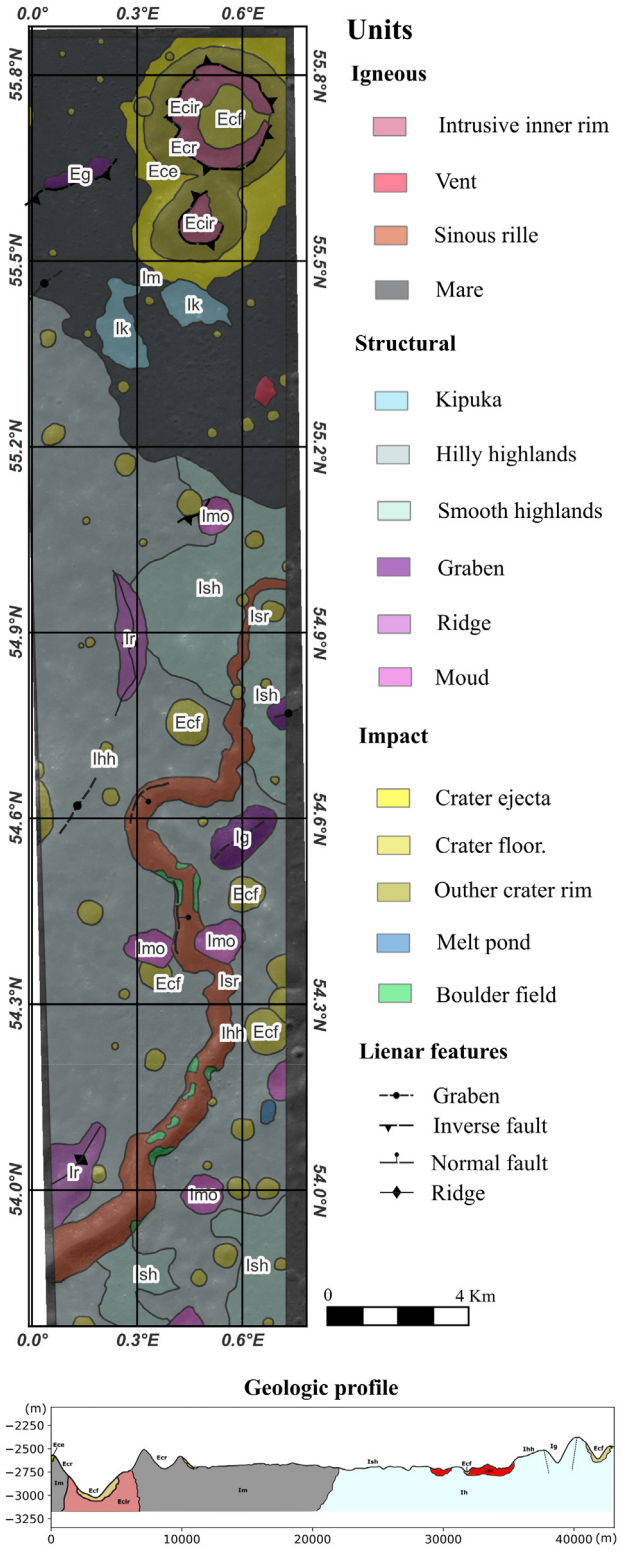

Figure 1. Geologic map of Archytas dome done in this work.

There is a clear accumulation of boulders near the rims of the larger craters and down-slope of the rille. The prior

configuration provides plenty of boulder fields and craters to be recognised by the models to be trained by the participants.

2.2 Data preparation

A large amount of data is necessary to train ML models for image recognition[31]. This data should be representative of the target feature and its diversity, and the feature itself has to be labelled in every image. Craters and boulders are the most studied landforms on the Moon; therefore, a large amount of labelled data is already available (Figure 2 a, b).

The dataset of craters was collected from the work of Fairweather *et al.*[32], who ran their model on images of the Narrow Angle Camera (NAC) onboard the Lunar Reconnaissance Orbiter (LROC)[6], and detected craters with diameters between 20 and 500 m. As the Archytas dome is covered by a single NAC image, the scale of these labels suits our needs.

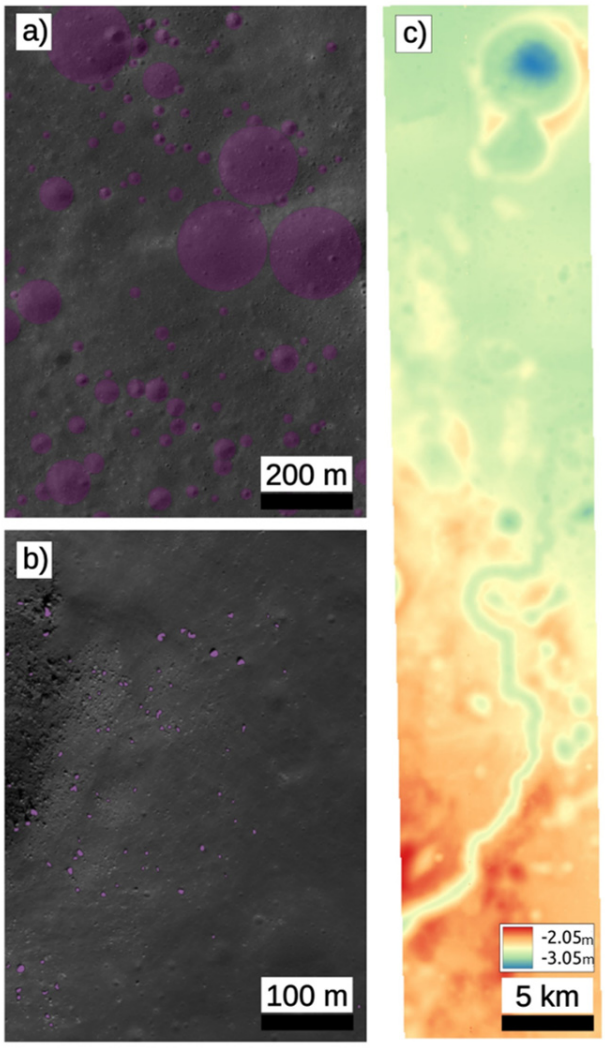

Figure 2. a) Labels of crater from Fairweather *et al.*[32]; **b**) Labels of boulders from Watkins *et al.*[21]; **c**) Digital Elevation model of the Archytas Dome, elevation in meters.

We compiled 3257 labelled images with craters of different sizes and shapes, although all of them were simple craters. The dataset of boulders was collected from Watkins *et al.*[21], which also detected these landforms in NAC images. The size of the boulders is naturally smaller than craters, ranging from 1 to 5 m. We compiled 9555 labelled images with boulders and boulders fields. Later, these datasets were randomly divided into test, valid, and training groups, containing 10 %, 20 %, and 70 % of the data, respectively. A list of the LROC images used for this purpose can be seen in Table 1.

For the data regarding the Archytas domes, we also used NAC images. The M1103544450LE and M1103544450RE takes were downloaded from the Lunar Orbital Data Explorer and processed up to map-projected images using the software ISIS[33]. Finally, a digital elevation model with a spatial resolution of 3 m was computed from the prior images using the Ames Stereo Pipeline (ASP)[34] (Figure 2c).

2.3 Ground truth

To evaluate the fidelity of the ML model after its application to the Archytas Dome zone, it was necessary to compare it with a ground truth. As the existing databases do not cover the location, we created our own set of labels for craters and boulders. This step was made manually in QGIS; we created a squared grid to cover the area better and labelled features with sizes and shapes similar to those of the training sets of data. In total, we collected 586 craters and 746 boulders.

3 Methods

The solution proposed by the winning team to solve the three steps of the challenge was divided into three phases respectively: a network is trained to segment and predict boulders and craters on a general lunar surface, then the network is used to predict these obstacles on the Archytas Dome area of the Moon, lastly a traverse is calculated to reach certain ROIs while avoiding obstacles and inaccessible areas computed in the prior steps. The full process of the methodology is represented in Figure 3 and discussed in detail below.

3.1 Segmentation and classification of boulders and craters

3.1.1 Data pre-processing. Neural networks aim to identify objects and their positions in an image, segmenting them into the desired distinctions. The ground truth of the segmentation is computed by unifying a few points of the object contours and then identifying the outside pixels as terrain and inside pixels as craters or boulders.

The images in the dataset constructed for the initial training contained boulders and craters of different sizes, which were resized to fifty-by-fifty pixels due to the neural network hardware requirements. The images are augmented with random translations and scales to reduce the number of boulders and craters of similar shape and size centred in the images. We applied additional shear values, random rotations, and vertical and lateral image flipping to increase the robustness of the network.

Table 1. Images used to obtain the labels for the training of the model.

Class	Label Type	Instrument	Image Name	Source
Crater	Bounding box + polygon	LROC NAC	M112963850lE, M1288076949lE, M1320016983lE, M1325197569lE, M1338833866lE, M1361391010lE	Fairweather *et al.* (2022)
Boulders	Bounding box + polygon	LROC NAC	M1108182629LE, M1108182629RE, M1259058367LE, M150633128LE, M150633128RE, M152770233LE, M152770233RE, M165645700LE, M165645700RE, M165998991RE, M181065865LE	Watkins, Ryan (2019)

We supplemented the dataset with augmented images from Mt. Marilyn[35–37] to improve the robustness of the model by adding topographic variations to the boulder and crater images.

3.1.2 Neural Network Architecture. The main task of the challenge is to classify boulders and craters on the lunar surface. Two different models, a Convolutional Neural Network (CNN) and a Vision Transformer (ViT), have been proposed and compared. A CNN learns the importance of the different elements of an image by assigning them weights and biases[38]. A ViT uses images divided into patches and considers the order in which these are given[39]. For an image representing a boulder or a crater, the network predicts its segmentation, whereby each image pixel is classified as either boulder, crater or terrain. The model uses an Adam optimiser and a cross-entropy loss, where the overall loss is minimised during training. The model training and testing, as well as the data handling, is executed by using Python, Pytorch and Jupyter Notebooks[40].

3.2 Prediction of obstacles in the Archytas Dome

The second step of the challenge required the implementation of the trained model to the NAC image of the Archytas Dome zone. Due to limitations in the computational resources, especially the memory, we segmented the image of the Archytas Dome into smaller fifty-by-fifty pixel sub-images to pass through the network. The network classified each pixel as boulder, crater or terrain, and then the classified images were joined back together to the dimension of the original NAC image.

3.3 Traverse calculations

For the third step, the calculation of a lunar traverse, our solution aimed to develop an efficient methodology for generating paths that avoid obstacles and reach pre-selected regions of interest (ROI) by combining various algorithms within QGIS. A non-exhaustive list of existing path-generating methods and algorithms include: 1D Convolutional neural networks[41], A*, D*, rapidly exploring random trees (RRT), potential fields[42], Dubins' paths, and Hamilton- Jacobi-Bellman methods[43]. Our simplified approach to this problem involved weighted ROI path selection and optimal traverse calculation requiring multiple image processing steps.

The QGIS[44] raster terrain analysis was initially applied to the DEM NAC Archytas image to compute slopes above fifteen degrees. This image was then converted to a binary raster image and combined with the boulder and crater images to form an obstacle map. In our reduced complexity approach, the minimum path was found by creating a map of cumulative ROI distance costs and obstacle proximity distances to act as pseudo-topography for the r.drain algorithm (available in the QGIS GRASS package), generating vectorised least-cost paths.

Repeating the above steps for all ROI pairs is computationally expensive and, as such, the ROI path order is automatically pre-selected based on a cost-minimised route based on ROI distance and ROI importance (low, medium, high). This calculation is achieved by forming a list of Cartesian distances between all pairs of ROIs via the QGIS 'Shortest line between

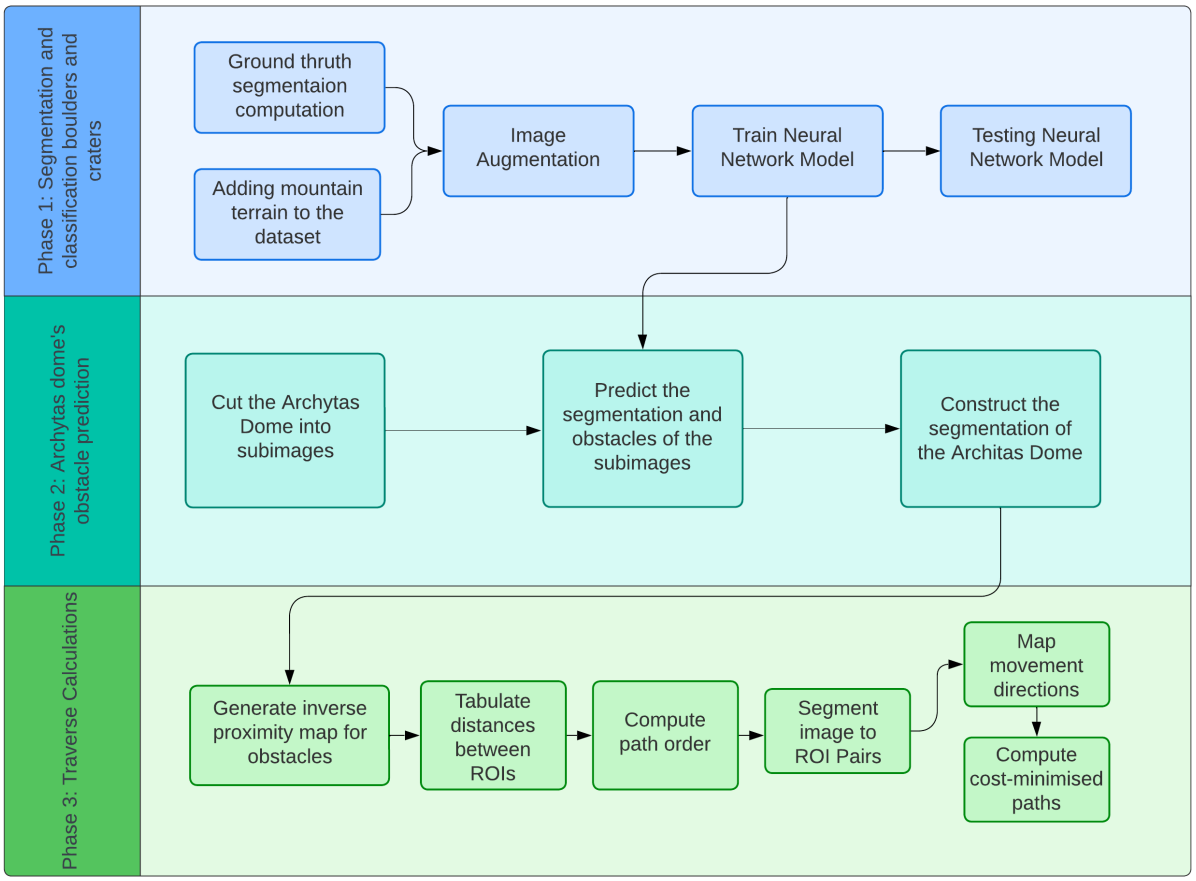

Figure 3. Flow diagram of each phase of the project; 1) segmentation and classification, 2) obstacle prediction, and 3) traverse calculations.

features' algorithm, adding ROI importance, and then applying our developed Dijkstra-based algorithm[45] to compute the cost-minimised-path from the 'Start' to 'End' ROI.

The GDAL raster distance proximity algorithm is applied to pre-compute obstacle distances for the whole map with geo-referenced coordinates. The obstacles are clipped by subtracting the binary image, and the resulting image is inverted to give the highest weighting closest to the obstacles.

The map is segmented for each ROI pair, and cumulative costs and movements between the ROIs are found by the QGIS r.cost algorithm (Figure 4) before generating the paths via the QGIS r.drain algorithm, which are then combined to form a single traverse.

4 Results

4.1 Boulders and craters segmentation and classification

We tested four models to predict boulders and craters across the Archytas Dome, and the results are summarised in Table 2. Via image augmentation (Subsection 3.1.1), the original dataset was doubled for the first three models and expanded eleven-fold for the fourth model. Figure 5 shows each model's heatmap, while Figure 6 shows the second phase prediction of boulders and craters across the Archytas Dome. Vision Transformers (ViTs) and Convolutional Neural Networks (CNNs) both achieved image segmentation. From the results, we can see that a convolutional neural network works best with a segmentation problem, likely due to its general purpose backbone[39]. Adding learnable weights to the classes does not seem to increase the model's overall score, but balances out the accuracy of the categories. Expanding the dataset improved the network's robustness as the network could observe a larger variety of obstacles. This is particularly noticeable in the prediction of obstacles of Archytas Dome in the second phase (Figure 6).

4.2 Traverse calculations

The applied methodology successfully computed a ~68.2 km long object avoidance traverse weighted between the start and end locations, reaching one low priority, four medium priority, and two high priority ROIs (Figure 7). The traverse could not be applied to the full set of initial ROIs due to two image artefacts in the DEM acting as slope barriers. The traverse does not cross any boundaries, and due to the stage two results, boulders and craters were not included in the final object avoidance image.

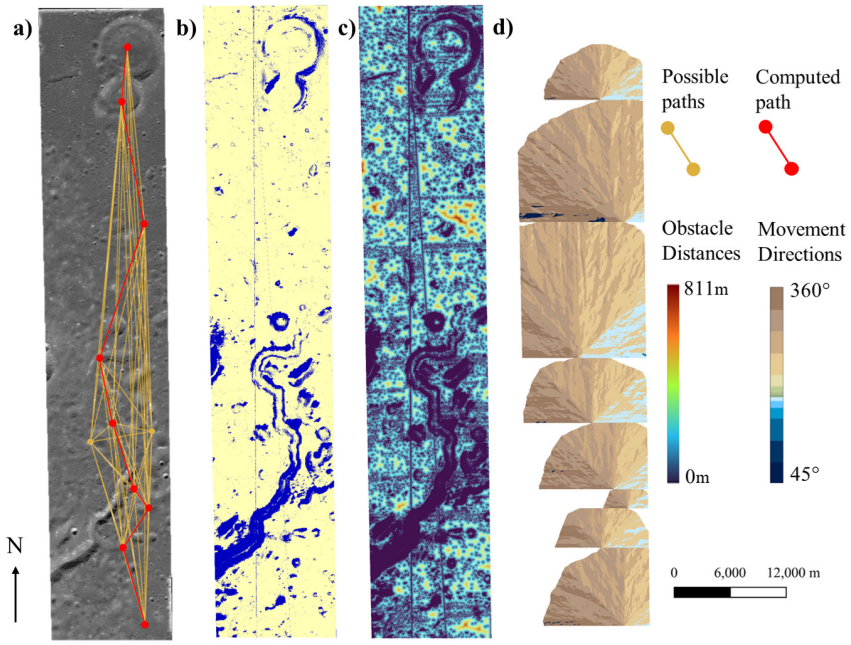

Figure 4. Steps to compute the optimal traverse. a) Dijkstra weighted ROI path selection. **b**) Binary obstacle image (blue obstacles, yellow non-obstacles). **c**) Object proximity distances. **d**) Least cost movement directions.

Table 2. Model results summary of F1 scores. Visual Transformer (ViT), Convolutional Neural Network (CNN), CNN with class weights, CNN with eleven times larger dataset.

Model	F1 (%)
ViT	62.31
CNN	71.44
CNN with class weights	70.69
CNN with larger dataset	**75.46**

5 Discussion

5.1 The model

The initial data has very similar features whereby the obstacles are centred in the image and have similar proportions with respect to the image size. This led us to augment the data with rotations, translations and scaling to acquire images of many types of boulders and craters. The visual transformer gives the least successful results due to its patching scheme. Meanwhile, a CNN learns how to give importance to the different areas of an image in no particular order with learnable weights and biases. Images of Mt. Marilyn have been added to the dataset to help the model distinguish obstacles on flat terrain from geologically mountainous terrain. Additionally, the model accuracies are also reduced due to the highly imbalanced dataset with more terrain than boulders and craters. The addition of weights balances the accuracy of the classes but not the overall F1 score. Enlarging part of the dataset images would greatly improve the F1 score in the first phase but the mountain chain south-west of the Archytas Dome would be fully identified as either a crater or a boulder. The addition of samples in the dataset improves the performance of the model due to the high number of possible craters and boulders. The Archytas Dome has been downsized and divided into smaller images due to its incompatibility with the square-shaped images of the first phase and its large size which does not follow the hardware constraints.

5.2 The traverse

The computed traverse methodology successfully computes weighted paths between ROIs automatically[46]. However, the cost-minimised path is based on maximum distances from objects (steep slopes), suggesting that there is scope to reduce the total path length with a traverse that comes closer to objects. The ROI order is pre-computed using a Dijkstra-based algorithm to reduce computational costs via a fully connected graph of distances between all ROIs. The complete path is not totally optimised as these distances assume straight-line paths, which do not account for the re-routing distances necessary for object avoidance. The cost-minimised traverse takes the ROI path as an input vector, where the path selection algorithm does not automatically allow paths to return to an ROI, preventing the automated addition of two medium-priority ROIs. The ROI weights are computed based on subtracting distances, which require arbitrary tuning, and are bounded by the minimum allowable distances of positive values, and do not give options to ensure a certain number of ROIs are reached. Our fully

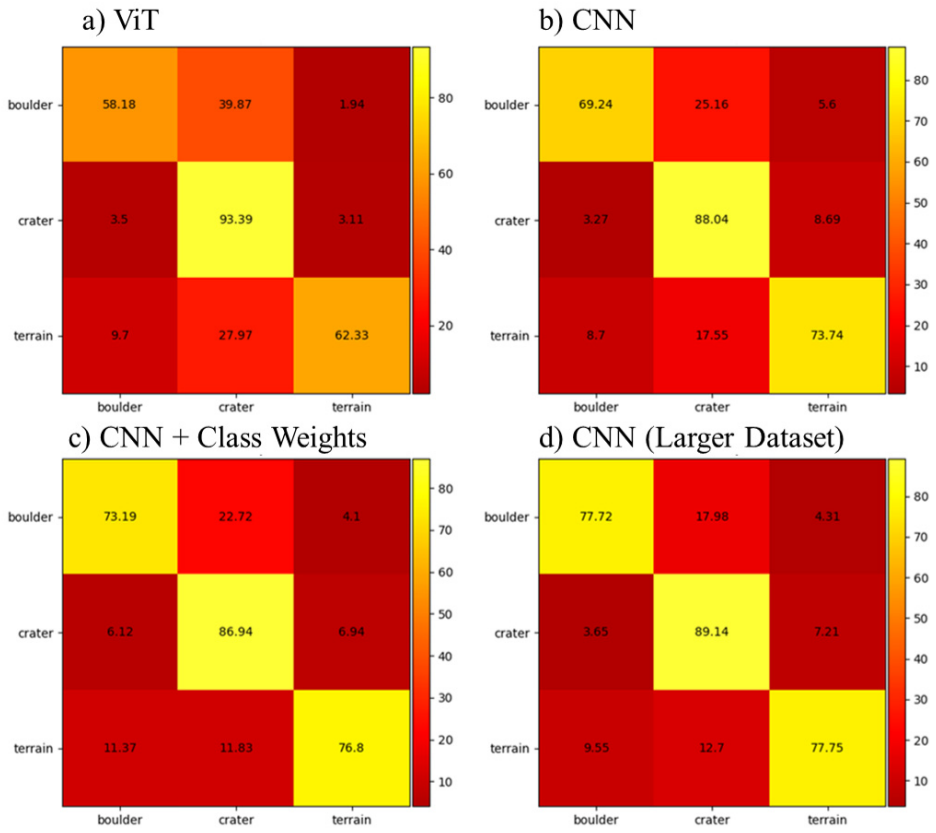

Figure 5. Model heatmaps. a) Vision Transformer (ViT), **b**) Standard CNN, **c**) CNN with class weights, **d**) CNN with a larger dataset.

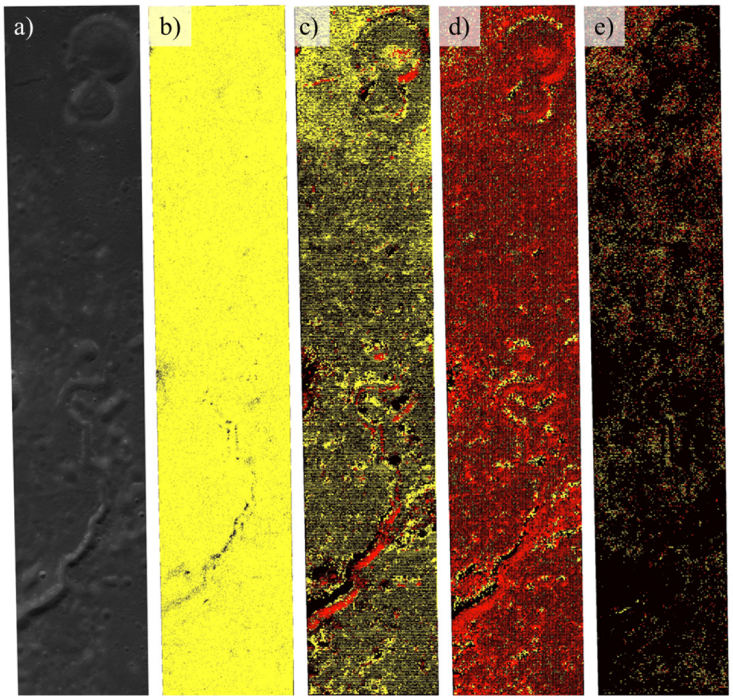

Figure 6. Results of the obstacle prediction phase for terrain (black), boulders (red), and craters (yellow). a) Input NAC image, **b**) ViT, **c**) CNN, **d**) CNN with class weights and **e**) CNN with a larger dataset.

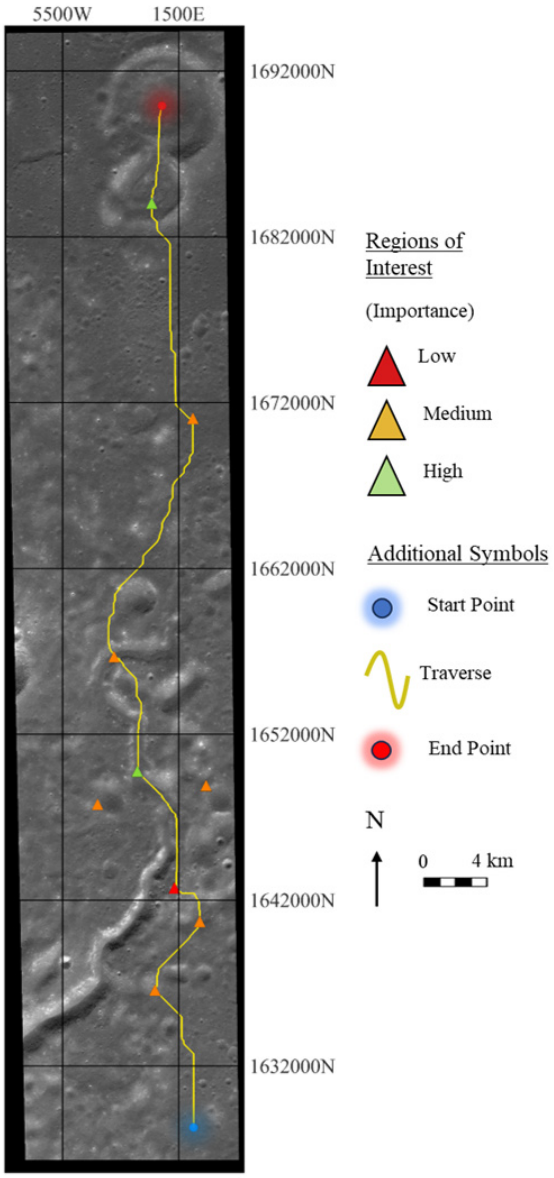

Figure 7. Automated minimised path length lunar traverse, avoiding slopes above fifteen degrees and contacting regions of interest around the Archytas Dome region.

automated approach accomplished the traverse computations using simple QGIS functions and a bespoke Dijkstra-based algorithm; future directions should seek to overcome our method limitations and explore additional realistic constraints such as accounting for terrain roughness, minimising total slopes travelled, maintaining sunlight exposure, and including potential distance limitations for communication relays.

5.3 The challenge framework

Apart from the results of the winning team, we can draw some lessons from the organization and development of the data challenge itself. The number of participants was low; around six teams registered for the activity, of which two presented results, and only the winning team provided a meaningful approach to solving the challenge. Although the results of the winning team were consistent and demonstrated a good understanding of the challenge, the limited participation indicates missions overlooked during the challenge conception.

One explanation for the lack of participants was the low diffusion time of the challenge. After launch, we presented the challenge in several conferences, forums, and academic groups focused on machine learning, but earlier dissemination could have given potential participants more preparation time. Another obstacle to achieving larger project participation may be the high challenge complexity. We found that the majority of the participants needed help with integrating the different steps of the challenge. Building a rover traverse from real environmental conditions is complex and multidisciplinary; hence, the teams must be equally diverse. The successful team comprised a computer engineer and a geophysicist, each focusing on the machine ML and traverse computation, respectively. This three-step challenge configuration likely discouraged individual participants as it required knowledge of multiple disciplinary backgrounds. Defining more constrained activities or goals may be more beneficial for future data challenges.

A technical problem that arose was the training data quality. The mixed results obtained by the winning team in the Archytas Dome could be partially related to the training data provided by the EXPLORE team. The combination of labels for both craters and boulders is likely suboptimal due to the great difference in size between both landforms, ranging from craters with hundreds of meters in diameter to boulders that only occupied a couple of pixels on the image. Another factor may be the difference between the images used for the training dataset and the one of the Archytas region. Several factors may vary from one image to the next, from the illumination to the textural properties of the rocks. As observed in highly illuminated slopes, a big variance from one data set to another could cause poor model behaviour. A more consistent training dataset and using one type of landform would likely allow participants to achieve better results.

6 Conclusions

The EXPLORE Machine Learning Lunar Data Challenge was an interesting activity to test the use of ML techniques to solve real problems faced in lunar exploration. In the search for replicating a semi-real scenario of exploration, the complexity of the challenge escalated beyond a point where non-specialised users could not participate in the challenge effectively, resulting in a low participation by the public. Although a simpler challenge would have attracted more participants, it is worth noting that the winning team successfully understood and delivered solutions for the three proposed steps. Even if those results could be improved by a better training dataset or simpler tasks, their approach proves the usefulness of ML techniques in the future of planetary exploration.

The winning team constructed a neural network that can locate the position of boulders and craters, and successfully created an automated system to compute a cost-minimised path

targeting weighted regions of interest that avoids obstacles such as slopes above fifteen degrees, boulders, and craters. This solution was based on connecting several QGIS functions and writing a bespoke Dijkstra-based script. The process relied on the cost-to-obstacle calculations and initial path selection from straight-line distances; the overall method was not fully optimised for minimising path lengths but may still be a useful tool for automatically generating traverse paths.

Ethics and consent

Ethical approval and consent were not required.

Data and software availability

Underlying data

Zenodo: EXPLORE Expert Data Challenges 2022 - Craters dataset. https://zenodo.org/doi/10.5281/zenodo.7096474[47].

Zenodo: EXPLORE Expert Data Challenges 2022 - Boulders dataset. https://zenodo.org/doi/10.5281/zenodo.7096496[48].

Zenodo: EXPLORE Machine Learning Lunar Data Challenges 2022 - QGIS project. https://zenodo.org/doi/10.5281/zenodo.7179841[49].

Data are available under the terms of the Creative Commons Attribution 4.0 International license (CC-BY 4.0).

Extended data

Zenodo: Lunar Traverse Algorithm - EXPLORE Machine Learning Lunar Data Challenge 2022. https://zenodo.org/doi/10.5281/zenodo.10722635[46].

Data are available under the terms of the Creative Commons Attribution 4.0 International license (CC-BY 4.0).

Zenodo: Design of robotic traverses on the Archytas Dome on the Moon. https://zenodo.org/doi/10.5281/zenodo.10972860[40]

Code is available under the terms of the Creative Commons Attribution 4.0 International license (CC-BY 4.0) and Apache license V2.0.

Other codes related to the EXPLORE project can be found at https://github.com/explore-platform.

Author contribution

Conceptualization: Giacomo Nodjoumi, Javier E. Suarez-Valencia, Daniel Le Corre

Data curation: Giacomo Nodjoumi, Javier E. Suarez-Valencia, Daniel Le Corre

Methodology: Lewis Lovell, Isabella C. Adriani

Formal analysis and Investigation: Lewis Lovell, Isabella C. Adriani

Writing - Original Draft: Lewis Lovell, Isabella C. Adriani, Javier E. Suarez-Valencia, Giacomo Nodjoumi, Daniel Le Corre

Writing - Review and Editing: Lewis Lovell, Isabella C. Adriani, Javier E. Suarez-Valencia, Daniel Le Corre, Giacomo Nodjoumi, Anita Heward, Angelo Pio Rossi, Nick L. J. Cox

References

1. Smith M, Craig D, Herrmann N, *et al.*: **The artemis program: an overview of NASA's activities to return humans to the Moon.** In: *2020 IEEE Aerospace Conference.* 2020; 1–10.
Publisher Full Text
2. Oche PA, Ewa GA, Ibekwe N: **Applications and challenges of artificial intelligence in space missions.** *IEEE Access.* 2021; **12**: 44481–44509.
Publisher Full Text
3. NASA: **Surveyor I flight performance.** Technical Report SSD 68222R, NASA, October, 1996.
4. Costes NC, Farmer JE, George EB: **Mobility performance of the lunar roving vehicle: terrestrial studies, apollo 15 results.** NASA, February, 1972.
5. DARPA: **Lunokhod-1 soviet lunar surface vehicle.** Technical report, DARPA, July, 1971.
6. Robinson MS, Brylow SM, Tschimmel M, *et al.*: **Lunar Reconnaissance Orbiter Camera (LROC) instrument overview.** *Space Sci Rev.* 2010; **150**(1–4): 81–124.
Publisher Full Text
7. Ding L, Zhou R, Yuan Y, *et al.*: **A 2-year locomotive exploration and scientific investigation of the lunar farside by the Yutu-2 rover.** *Sci Robot.* 2022; **7**(62): eabj6660.
PubMed Abstract | Publisher Full Text
8. Zheng Y, Ouyang Z, Li C, *et al.*: **China's lunar exploration program: present and future.** *Planet Space Sci.* 2008; **56**(7): 881–886.
Publisher Full Text
9. Colaprete A, Elphic R, Shirley M, *et al.*: **The Volatiles Investigating Polar Exploration Rover (VIPER) mission: measurement goals and traverse planning.** In: *AGU Fall Meeting Abstracts.* 2021; **2021**: P53B–08.
Reference Source
10. Els SG, Almaeeini S, Almarzooqi H: **The science instrumentation for the rashid rover of the emirates lunar mission.** In: *52nd Lunar and Planetary Science Conference.* Lunar and Planetary Science Conference, 2021; 1905.
11. Ahmad J, Farman H, Jan Z: **Deep learning methods and applications.** In: *SpringerBriefs in Computer Science.* Springer Singapore, 2019.
Publisher Full Text
12. Hassoun MH: **Fundamentals of artificial neural networks.** *Proceedings of the IEEE.* 2005.
Publisher Full Text
13. Sonka M, Hlavac V, Boyle R: **Image processing, analysis, and machine vision.** *Thomson Learning.* 2008; 812.
Publisher Full Text
14. Stepinski TF, Ghosh S, Vilalta R: **Machine learning for automatic mapping of planetary surfaces.** *Proceedings of the National Conference on Artificial Intelligence.* 2007; 1807–1812.
Publisher Full Text
15. Dineva R, Koichev A: **Benchmarking of CNNs for low-cost, low-power robotics applications.** *Intern Dis.* 2000.
16. Gu J, Wang Z, Kuen J, *et al.*: **Recent advances in convolutional neural networks.** *Pattern Recogn.* 2018; **77**: 354–377.
Publisher Full Text

17. He K, Gkioxari G, Dollár P, *et al.*: **Mask R-CNN.** *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 2020; **42**(2): 386–397.
Publisher Full Text

18. Krizhevsky A, Sutskever I, Hinton GE: **ImageNet classification with deep convolutional neural networks.** In: *ImageNet Classification with Deep Convolutional Neural Networks.* 2012.

19. La Grassa R, Cremonese G, Gallo I, *et al.*: **YOLOLens: a deep learning model based on super-resolution to enhance the crater detection of the planetary surfaces.** *Remote Sens.* 2023; **15**(5): 1171.
Publisher Full Text

20. Wang Y, Wu B: **Active machine learning approach for crater detection from planetary imagery and Digital Elevation Models.** *IEEE Transactions on Geoscience and Remote Sensing.* 2019; **57**(8): 5777–5789.
Publisher Full Text

21. Watkins RN, Jolliff BL, Mistick K, *et al.*: **Boulder distributions around young, small lunar impact craters and implications for regolith production rates and landing site safety.** *J Geophys Res (Planet).* 2019; **124**(11): 2754–2771.
Publisher Full Text

22. Prieur NC, Amaro B, Gonzalez E, *et al.*: **Automatic characterization of boulders on planetary surfaces from high-resolution satellite images.** *J Geophys Res: Planet.* 2023; **128**(11): e2023JE008013.
Publisher Full Text

23. Barrett AM, Balme MR, Woods M, *et al.*: **NOAH-H, a deep-learning, terrain classification system for mars: results for the ExoMars rover candidate landing sites.** *Icarus.* 2022; **371**: 114701.
Publisher Full Text

24. Wilhelm T, Geis M, Püttschneider J, *et al.*: **DoMars16k: a diverse dataset for weakly supervised geomorphologic analysis on mars.** *Remote Sens.* 2020; **12**(23): 1–38.
Publisher Full Text

25. Nodjoumi G, Pozzobon R, Sauro F, *et al.*: **DeepLandforms: a deep learning computer vision toolset applied to a prime use case for mapping planetary skylights.** *Earth Space Sci.* 2023; **10**(1): e2022EA002278.
Publisher Full Text

26. Schmiedel T, Galland O, Breitkreuz C: **Dynamics of sill and laccolith emplacement in the brittle crust: role of host rock strength and deformation mode.** *J Geophys Res Solid Earth.* 2017; **122**(11): 8860–8871.
Publisher Full Text

27. Lena R, Wöhler C, Phillips J, *et al.*: **Lunar domes.** Springer Milan, Milano, 2013.
Publisher Full Text

28. Wöhler C, Lena R: **Lunar intrusive domes: morphometric analysis and laccolith modelling.** *Icarus.* 2009; **204**(2): 381–398.
Publisher Full Text

29. Farley KA, Williford KH, Stack KM, *et al.*: **Mars 2020 mission overview.** *Space Sci Rev.* 2020; **216**(8): 142.
Publisher Full Text

30. M'Gonigle JW, Schleicher D: **Geologic map of the Plato quadrrangle of the Moon**. 1972.
Publisher Full Text

31. Pak M, Kim S: **A review of deep learning in image recognition.** In: *2017 4th International Conference on Computer Applications and Information Processing Technology (CAIPT).* Kuta Bali, IEEE. 2017; 1–3.
Publisher Full Text

32. Fairweather JH, Lagain A, Servis K, *et al.*: **Automatic mapping of small lunar impact craters using LRO-NAC images.** *Earth Space Sci.* 2022; **9**(7): e02177.
Publisher Full Text

33. Sucharski T, Mapel J, Jcwbacker, *et al.*: **USGS-Astrogeology/ISIS3: ISIS 4.2.0 public release.** *Zenodo.* 2020.
Publisher Full Text

34. Beyer RA, Alexandrov O, McMichael S: **The Ames Stereo Pipeline: NASA's open source software for deriving and processing terrain data.** *Earth Space Sci.* 2018; **5**(9): 537–548.
Publisher Full Text

35. Estes NM, Hanger CD, Licht AA, *et al.*: **Lunaserv web map service: history, implementation details, development, and uses.** [Abstract]. *44th Lunar and Planetary Science Conference.* 2013; 2609.
Reference Source

36. Robinson MS: **Lunar Reconnaissance Orbiter Camera experimental data record, LRO-L-LROC-2-EDR-V1.0.** *NASA Planetary Data System.* 2010.

37. Robinson MS: **Lunar Reconnaissance Orbiter Camera (LROC) instrument overview.** *Space Sci Rev.* 2010; **150**: 81–124.
Publisher Full Text

38. O'Shea K, Nash R: **An introduction to convolutional neural networks.** *CoRR.* 2015; abs/1511.08458.
Publisher Full Text

39. Thisanke H, Deshan C, Chamith K, *et al.*: **Semantic segmentation using vision transformers: a survey.** *Eng Appl Artif Intell.* 2023; **126**(Part A): 106669.
Publisher Full Text

40. Adriani IC: **Design of robotic traverses on the Archytas Dome on the Moon.** *Zenodo.* 2024.
http://www.doi.org/10.5281/zenodo.10972861

41. Visca M, Bouton A, Powell R, *et al.*: **Conv1d energy-aware path planner for mobile robots in unstructured environments.** In: *2021 IEEE International Conference on Robotics and Automation (ICRA).* IEEE, 2021; 2279–2285.
Publisher Full Text

42. Gopikrishnan S, Shravan BVS, Gole H, *et al.*: **Path planning algorithms: a comparative study.** *Space Transportation Systems.* 2011.
Reference Source

43. Parkinson C, Arnold D, Bertozzi AL, *et al.*: **Optimal human navigation in steep terrain: a Hamilton-Jacobi-Bellman approach.** *arXiv preprint arXiv: 1805.04973.* 2018.
Publisher Full Text

44. QGIS Development Team: **QGIS Geographic Information System**. QGIS Association, 2024.
Reference Source

45. Dijkstra EW: **A note on two problems in connexion with graphs.** In: *Edsger Wybe Dijkstra: His Life, Work, and Legacy.* Association for Computing Machinery, 2022; 287–290.
Publisher Full Text

46. Lovell L: **Lunar Traverse Algorithm - EXPLORE machine learning lunar data challenge 2022 (Version v1).** Zenodo. [Code]. 2024.
http://www.doi.org/10.5281/zenodo.10722636

47. Nodjoumi G, Valencia JS: **EXPLORE Expert Data Challenges 2022 - craters dataset (Version 1).** Zenodo. [Dataset]. 2022.
http://www.doi.org/10.5281/zenodo.7181924

48. Nodjoumi G, Valencia JS: **EXPLORE Expert Data Challenges 2022 - boulders dataset (Version 1).** Zenodo. [Dataset]. 2022.
http://www.doi.org/10.5281/zenodo.7181923

49. Nodjoumi G, Valencia JS: **EXPLORE machine learning lunar data challenges 2022 - QGIS project.** Zenodo. [Dataset]. 2022.
http://www.doi.org/10.5281/zenodo.7179842

Open Peer Review

Current Peer Review Status: ? ✓

Version 1

Reviewer Report 28 September 2024

https://doi.org/10.21956/openreseurope.18831.r44140

© **2024 Malki H et al.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

✓ **Mohammad Mehdi Taghavi Nasrabadi**
University of Houston, Houston, Texas, USA
Heidar Malki
University of Houston, Houston, Texas, USA

The paper provides a clear and detailed account of how machine learning models can be used to address challenges in lunar exploration. The authors' use of convolutional neural networks and vision transformers to classify lunar landforms is well-executed, and the methodology for calculating an optimal traverse is effective. The work is well-supported by relevant literature and available datasets. However, the training dataset could benefit from a broader variety of images to further improve model robustness, and alternative path planning algorithms such as A* or D* could be explored for efficiency. Overall, the paper demonstrates an important contribution to planetary exploration with machine learning techniques.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and does the work have academic merit?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: No competing interests were disclosed.

We confirm that we have read this submission and believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 11 September 2024

https://doi.org/10.21956/openreseurope.18831.r43375

© 2024 Santra S. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Shreya Santra
Tohoku University, Sendai, Japan

This article summarizes the approach taken by the winning team of the EXPLORE Machine Learning Lunar Data Challenge. While it doesn't present novel findings, it is commendable for its reproducibility and practical demonstration of the methodology.

The strengths of the work include a detailed description of the lunar target site and thorough dataset preparation. The team developed a neural network capable of identifying boulders and craters, and successfully built an automated system to compute a cost-optimized path through weighted regions of interest while avoiding obstacles like steep slopes, boulders, and craters. This solution integrated several tools such as processing using QGIS functions, with a custom Dijkstra-based script to automatically generate traverse path, leveraging cost-to-obstacle calculations and initial path selection based on straight-line distances.

All datasets used in the project are openly available, allowing for easy replication of the work.

The comparison between the performance of Vision Transformers (ViTs) and Convolutional Neural Networks (CNNs) for segmenting craters and boulders ensured that the most effective method was chosen to enhance the results. Expanding the dataset proved beneficial, as it increased the network's robustness by exposing it to a wider range of obstacles. However, this expansion may have also led to overfitting or generated false positives, a potential issue that was not fully explored in the study.

That said, the method and results could have been further refined with a more diverse training dataset that separately represented boulders and craters, given their significant variation in size. Incorporating a data augmentation strategy, such as using virtual lunar simulators, could have added a novel dimension to the approach. The path planning approach could have been further validated by testing it against alternative algorithms such as A* and D*, which are known to address some of the limitations of Dijkstra's algorithm. Incorporating these planners would allow for a more comprehensive evaluation, potentially improving the efficiency of the pathfinding and applicability in different lunar terrain scenarios.

Note: It is unclear whether the authors of this article were part of the winning team or the challenge organizers. If the authors are not members of the winning team, proper acknowledgment of the team's contribution should be included in the article.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and does the work have academic merit?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Aerospace engineering, space robotics perception, control and navigation

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.