LunarLoc: Segment-Based Global Localization on the Moon

Annika Thomas, Robaire Galliath, Aleksander Garbuz, Luke Anger, Cormac O'Neill, Trevor Johst, Dami Thomas, George Lordos, Jonathan P. How

***Abstract*—Global localization is necessary for autonomous operations on the lunar surface where traditional Earth-based navigation infrastructure, such as GPS, is unavailable. As NASA advances toward sustained lunar presence under the Artemis program, autonomous operations will be an essential component of tasks such as robotic exploration and infrastructure deployment. Tasks such as excavation and transport of regolith require precise pose estimation, but proposed approaches such as visual-inertial odometry (VIO) accumulate odometry drift over long traverses. Precise pose estimation is particularly important for upcoming missions such as the ISRU Pilot Excavator (IPEx) that rely on autonomous agents to operate over extended timescales and varied terrain. To help overcome odometry drift over long traverses, we propose LunarLoc, an approach to global localization that leverages instance segmentation for zero-shot extraction of boulder landmarks from onboard stereo imagery. Segment detections are used to construct a graph-based representation of the terrain, which is then aligned with a reference map of the environment captured during a previous session using graph-theoretic data association. This method enables accurate and drift-free global localization in visually ambiguous settings. LunarLoc achieves sub-cm level accuracy in multi-session global localization experiments, significantly outperforming the state of the art in lunar global localization. To encourage the development of further methods for global localization on the Moon, we release our datasets publicly with a playback module: https://github.com/mit-acl/lunarloc-data.**

I. Introduction

Autonomous systems will play a vital role in preparing humanity for a sustained presence on the Moon. Precise and robust global localization will be a key capability for enabling autonomous surface operations to operate drift-free during precision-based tasks. The ability for robotic systems to accurately determine their position and orientation within a shared lunar reference frame is critical for navigation, exploration, infrastructure deployment, and scientific discovery. Unlike Earth-based systems that benefit from Global Navigation Satellite Systems (GNSS), the Moon lacks robust GNSS support, requiring global localization techniques that integrate on-board perception and environmental understanding for accurate mapping.

Authors are affiliated with Massachusetts Institute of Technology, Cambridge, MA 02139, USA. annikat@mit.edu.

The simulation used for this research was developed as part of the Lunar Autonomy Challenge. The Lunar Autonomy Challenge is a collaboration between NASA, The Johns Hopkins University (JHU) Applied Physics Laboratory (APL), Caterpillar Inc., and Embodied AI. APL is managing the challenge for NASA.

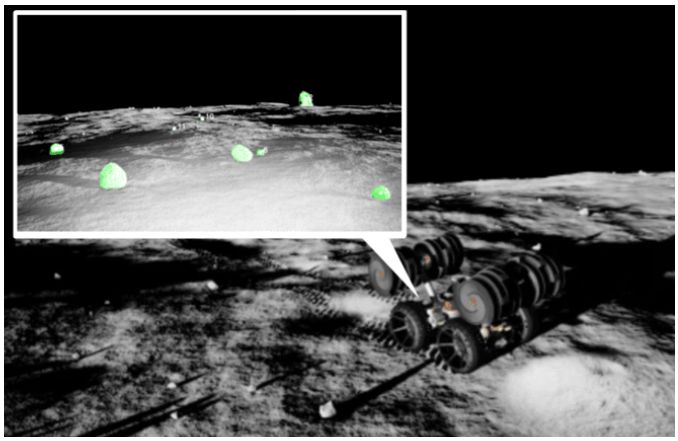


Fig. 1: LunarLoc uses open-set segmentation to generate maps through detection of boulders (shown in green), then performs object-based global localization using multi-session maps.

This ability for robots to localize without dependence on GNSS is particularly challenging in the context of lunar operations where harsh lighting conditions drastically alter the appearance of the environment. Classical approaches for localization involve visual place recognition [1, 2] followed by visual feature extraction [3, 4] and matching [5]. Place recognition methods struggle with viewpoint variation and may fail to provide descriptors that are different enough for meaningful association across places [6]. While feature-based approaches are effective for applications such as frame-to-frame tracking, they may fail to provide meaningful visual cues for long-term localization. These limitations of current global localization methods motivate the need for a separate approach to global localization on the Moon.

LunarLoc uses the underlying structure of the environment as a cue for global localization. The lunar environment has an abundance of rocks, the positions of which do not change under varying lighting conditions. Zero-shot extraction of segments within the scene provides consistent detections of rocks regardless of lighting conditions. LunarLoc provides a robust paradigm for global localization on the lunar surface, as demonstrated through experiments with a digital twin of the IPEx rover operating in a simulated lunar environment.

Our contributions include:

- Real-time object-based mapping using open-set segmentation for multi-session graph-theoretic global localiza-

tion on the lunar surface

- Experimental validation of our approach in a simulated lunar environment, achieving cm-level accuracy which significantly outperforms the state of the art for lunar global localization
- Release of our datasets as well as a simulation playback environment to encourage further development of lunar global localization

II. Related Works

Several approaches have been proposed for global localization in environments without GNSS support and unstructured environments. We review these approaches as well as landmark-based localization and planetary localization.

A. Global Localization in Environments without GNSS

Image-based localization techniques are commonly used in environments with GNSS. Place recognition approaches such as bag of words [8] and NetVLAD [1] store image representations as embeddings in a database which are used to associate new images with previous observations. These approaches are reasonably robust to changing environments from the same viewpoint. Sarlin et. al [9] built upon this concept to allow for online pose estimation using image-based localization. New images are compared to an existing library for an initial coarse estimation, before keypoint matching is applied for finer positioning. While these approaches demonstrate high recall in visually distinctive environments, their performance degrades when comparing different places that include similar features [6]. Learning-based methods such as [10] and [11] are also sensitive to perceptual aliasing in visually ambiguous settings.

B. Localization in Unstructured and Natural Environments

Unstructured and natural environments present a challenging setting for many visual SLAM systems which may assume the presence of urban-centric fixtures. To address the lack of semantic information such as lane markings and buildings, [12] integrates wheel odometry with visual tracking and [13] includes topological maps in the localization process. Approaches such as [14] and [6] specifically use the underlying structure of the environment as a cue for global localization.

C. Landmark-Based and Object-Centric Localization

To overcome issues arising from perceptual aliasing, recent works have proposed landmark-based and object-centric localization using graph theoretic alignment between objects within maps. These approaches are lightweight, lending themselves well to multi-agent scenarios as well. Ankenbauer et. al [14] proposed object-centric global localization in urban and unstructured scenarios, relying on YOLO [15] to extract objects from the environment which limits extension to settings without pre-determined object classes. To enable open-set operation, SOS-Match [6] employs zero-shot segmentation from SAM to extract segments reflective of objects in the scene. ROMAN [16] and [17] incorporate semantics into the object-based formulation to further disambiguate between geometrically similar scenes.

TABLE I: Planetary Localization Methods

Method	Modality	Locale	Accuracy [m]
GPS, [22]	Satellite Time Code	Earth	4.9
Geo-Referencing [18]	Satellite Imagery	Mars	0.5
TRN [20, 21]	Local Imagery	Mars	10.0
MAVeN [23, 24]	Aerial Imagery	Mars	3.0
LunarNav [25]	Local Imagery	Moon	5.0
ShadowNav [26]	Local Imagery	Moon	1.5
LunarLoc (Proposed)	Local Imagery	Moon	0.02

D. Planetary and Lunar Localization

Due to the lack of GNSS coverage on both the Moon and Mars, previous missions have adopted several different strategies to perform localization. While most rovers utilize wheel odometry and inertial measurement units for relative localization, these methods are prone to odometry drift over longer traverses. The Spirit and Opportunity rovers overcame this by correlating local panoramas to satellite images at regular intervals [18]. This method relies on high resolution images of the rover's operational site, and is highly prone to changes in lighting. Both additionally incorporated visual-odometry and celestial sensing, but the former is still prone to drift and the latter only provides accurate estimates for heading [19]. For the task of landing, orbital images are again used as a reference against sensor data in Terrain Relative Navigation (TRN) [20]. This technique was later shown to provide 10 meter position accuracy on the Perseverance rover [21]. A comparison of existing and proposed planetary localization methods is provided in Table I. Of these methods, LunarLoc is the only approach that provides centimeter-level accuracy.

III. Methodology

A. Open-Set Object-Based Mapping

Our mapping approach consists of performing zero-shot instance segmentation on input images using a pre-trained foundation model [7]. The segments are primarily extracted as boulders in the scene, reflecting the underlying geometric structure of the environment as shown in Fig. 2.

We perform mapping simultaneously from the front and back of the rover. Images captured from the front-left and rear-left IPeX rover cameras, as shown in Fig. 3, are used as inputs to the segmentation model.

We apply point prompting over a uniform 32×32 grid, producing a set of instance masks:

$$\mathcal{M} = \{m_i\}_{i=1}^{N}, \tag{1}$$

where each mask m_i is defined by its image-frame centroid $\mathbf{c}_i \in R^2$ and size $s_i \in R^+$:

$$m_i = (\mathbf{c}_i, s_i). \tag{2}$$

We find mask size by projecting each centroid $\mathbf{c}_i$ from the 2D image frame into 3D coordinates in the rover frame using stereo depth imagery. Assuming approximately spherical

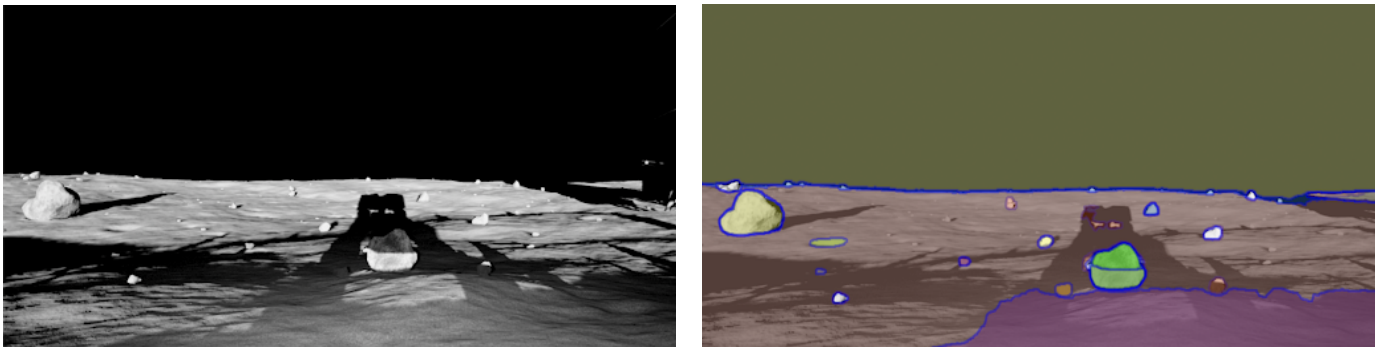


Fig. 2: Extraction of segment locations via open-set segmentation with FastSAM [7]. Images include a front-view image taken from the IPEx digital twin in the Lunar Simulator (left) and the zero-shot segmented image (right).

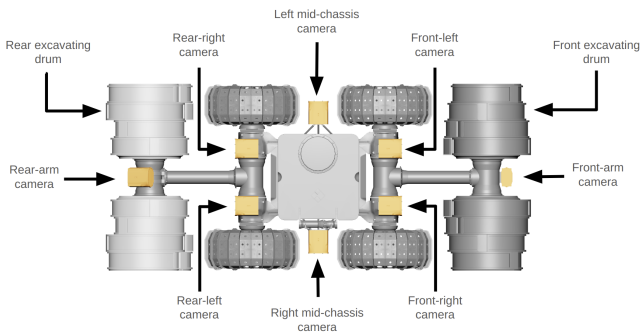


Fig. 3: IPEx rover diagram where the front-left camera and rear-left camera are used as inputs to the segmentation model and front-right and rear-right cameras provide imagery for stereo depth estimation of mapped objects, from [27]

objects, we use the standard deviation σ_i of the mask and the depth D_i at $\mathbf{c}_i$ to estimate the physical size s_i of the object:

$$s_i = \sigma_i \cdot D_i \cdot \alpha, \tag{3}$$

where α is the angular resolution (e.g., radians per pixel) of the camera. This converts 2D pixel extent into a real-world spatial scale.

We filter masks to remove those that (1) intersect the image boundary or (2) fall outside pre-defined size thresholds $s_{\text{min}} \leq s_i \leq s_{\text{max}}$. To further refine the set, we compute the 2D spatial covariance matrix of each mask and reject elongated shapes by enforcing an eigenvalue ratio constraint:

$$\frac{\lambda_{\text{max}}}{\lambda_{\text{min}}} \leq \tau_{\text{elong}}, \tag{4}$$

where λ_{max} and λ_{min} are the principal eigenvalues of the covariance matrix of the mask's pixel distribution. Masks with high anisotropy are typically shadows or terrain edges and are excluded. The filtered set is denoted:

$$\widetilde{\mathcal{M}} = \left\{ m_i \in \mathcal{M} \;\middle|\; \begin{array}{l} m_i \text{ is fully visible}, \\ s_{\text{min}} \leq s_i \leq s_{\text{max}}, \\ \lambda_{\text{max}}/\lambda_{\text{min}} \leq \tau_{\text{elong}} \end{array} \right\}. \tag{5}$$

The final set of projected 3D centroids stored for each frame is:

$$M_{\text{veh}} = \{\mathbf{C}_i\}_{i=1}^{\widetilde{N}}, \quad \mathbf{C}_i \in R^3, \tag{6}$$

where $\widetilde{N} = |\widetilde{\mathcal{M}}|$.

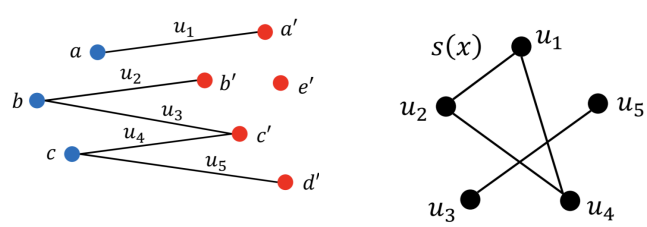
The use of FastSAM [7] allows this procedure to occur in real-time on a CPU, which is crucial for implementation of SLAM systems on the lunar surface. As the segmentation time is proportional to the prompt area, further speed-ups can additionally be achieved by making assumptions on the horizon line placement and excluding portions of the image.

B. Graph-Theoretic Data Association

The goal of the data association pipeline is to perform localization between the current map M_{veh}, computed in real time onboard the agent, and a reference map M_{ref}, collected by an agent during a prior session.

We utilize a geometric data association framework [28] to compute the largest geometrically consistent set of 3D segment correspondences between M_{veh} and M_{ref}. Geometric consistency is defined via pairwise distance preservation between associated segments. For instance, if points $\mathbf{a}, \mathbf{b} \in M_{\text{veh}}$ are matched to $\mathbf{a}', \mathbf{b}' \in M_{\text{ref}}$, then the pair is considered consistent if:

$$\|\mathbf{a} - \mathbf{b}\| \approx \|\mathbf{a}' - \mathbf{b}'\|.$$



(a) (b)

Fig. 4: Illustrative example of the consistency graph and its affinity matrix for point cloud registration, from [28]. (a) Blue and red point clouds with putative associations $u_1, \ldots, u_5$. (b) Consistency graph G, where edges indicate geometric consistency between associations.

An example of this process is shown in Fig. 4.

We represent candidate associations as nodes in a graph, where the affinity matrix $A \in \{0,1\}^{n \times n}$ defines pairwise consistency:

$$A_{u_i,u_j} = \begin{cases} 1 & \text{if associations } u_i \text{ and } u_j \text{ are consistent,} \\ 0 & \text{otherwise.} \end{cases}$$

To find the largest mutually consistent set of correspondences, we solve for the densest clique via the following optimization:

$$\begin{aligned} \max_{u \in \{0,1\}^n} \quad & \frac{u^\top A u}{u^\top u} \\ \text{subject to} \quad & u_i u_j = 0 \quad \text{if } A_{i,j} = 0. \end{aligned} \tag{7}$$

Given the resulting set of geometrically consistent segment correspondences $S = \{(\mathbf{a}_i, \mathbf{b}_i)\}_{i=1}^k$, where $\mathbf{a}_i \in M_{\text{veh}}$ and $\mathbf{b}_i \in M_{\text{ref}}$, we estimate the rigid transformation that aligns the current vehicle map to the reference map using Arun's method [29]. This produces the optimal rotation $\mathbf{R} \in \text{SO}(3)$ and translation $\mathbf{t} \in R^3$ minimizing:

$$\sum_{i=1}^{k} \|\mathbf{b}_i - (\mathbf{R}\mathbf{a}_i + \mathbf{t})\|^2 .$$

The resulting transformation from the vehicle frame to the reference map frame is:

$$\mathbf{T}_{\text{veh}}^{\text{ref}} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^\top & 1 \end{bmatrix} \in \text{SE}(3).$$

IV. Evaluation

A. Simulator

We evaluate the performance of our approach using the Lunar Simulator, an extension of the CARLA Simulator [30] developed for the Lunar Autonomy Challenge [27]. While CARLA is an open-source platform, the Lunar Simulator adapts its capabilities to replicate the unique challenges of lunar exploration with a high-fidelity, physics-based simulation of the lunar surface, including terrain, lighting conditions, and other environmental factors reflective of space exploration contexts. The Lunar Simulator includes a 27m x 27m map with 386 rocks, a lunar lander in the center, and elevation varying across approximately 1m, as shown in Fig 5.

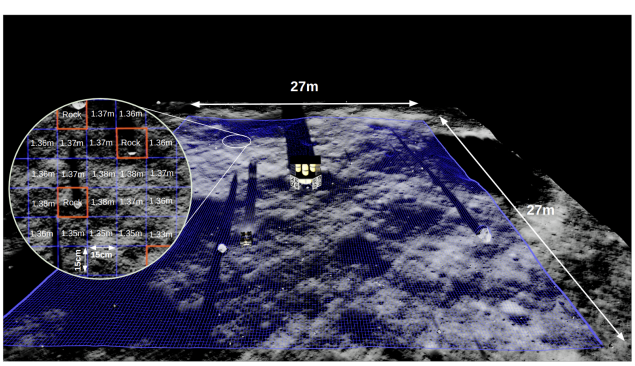


Fig. 5: The Lunar Simulator includes a lunar lander at the center of a 27m x 27m map, from [27]

TABLE II: Traverse Data

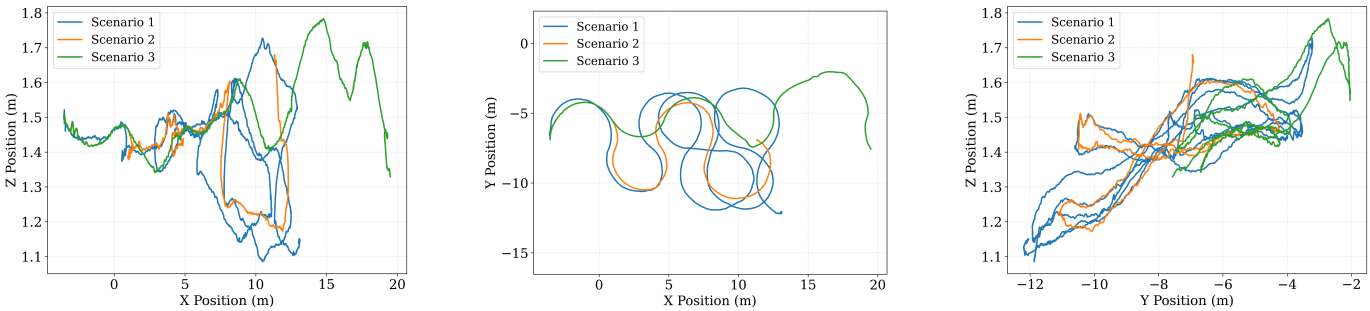
Traverse	Length [m]	Duration [min]	Object Detections
1	6.3	6	1283
2	6.5	6	1280
3	5.6	6	1261
4	17.8	16	3376
5	9.6	9	1884
6	8.9	6	1060
7	37.3	7	6263
8	15.6	12	2374
9	54.5	10	1695
10	88.2	15	2982
11	36.5	7	874
12	35.9	7	871
13	197.3	26	14494
14	33.3	6	2911
15	98.3	16	6722
16	8.7	2	755
17	47.2	8	3701

In the Lunar Simulator, we deploy a digital twin of NASA's In-Situ Resource Utilization (ISRU) Pilot Excavator (IPEx) rover. This agent reflects the physical IPEx robot at Kennedy Space Center, which is engineered to excavate and transport lunar regolith. The simulator incorporates detailed models of the IPEx's sensors, actuators, and overall dynamics. The sensor locations reflect the IPEx rover's design, as shown in Fig. 3. Cameras are arranged in two sets of stereo-pairs on the front and rear of the rover and as standalone cameras on the sides and excavating drums. All cameras are grayscale and simulated as perfect pinhole cameras.

B. Dataset

To evaluate the performance of LunarLoc, we collected stereo images and poses from short, medium, and long traverses spanning the same general region of the simulator map covering an area of approximately 1000 square meters. All 17 traverses are shown in Fig. 7, and detailed information on each traverse is provided in Table II. Subsets of longer traverses are taken to form shorter traverses for analysis. Each traverse is made available as a simple comma-separated table consisting of the IPEx position and boulder detections in each frame and as a `.lac` archive file, containing more detailed sensor information and time synchronized image data, details of which are outlined in Table III.

Each `.lac` archive is a gzip tarball. For ease of use, a Python library is provided to access synchronized camera and sensor data. The library provides direct access to underlying tabular data as well as an API mimicking the sensor interfaces provided by the Lunar Simulator. This mock API enables the playback of previous traverses outside of the simulator and without any reliance on the simulator code. Localization methods developed using the data loader can be quickly tested outside of the resource intensive simulator and then seamlessly deployed in the Lunar Simulator, accelerating development.



(a) X-Z Traverse (b) X-Y Traverse (c) Y-Z Traverse

Fig. 6: Ground truth traverses from the Lunar Simulator used to evaluate LunarLoc. Scenarios 1, 2 and 3 refer to traverses used for multi-session global localization.

Traverse data and the python data loader utility are made available on GitHub at https://github.com/mit-acl/lunarloc-data.

To collect data for each traverse, the IPEx is initialized at a common starting location from which it navigates to randomly selected way points in the immediate region until the minimum path length for the traverse is reached. This exploration strategy increases the frequency of repeated observations of rocks and other terrain features within and between traverses. Images from the front-left and front-right stereo cameras are sampled at 5Hz and used as inputs for the zero-shot segmentation and object centroid estimation pipeline. If suitable objects are detected in the scene the estimated position of the object's centroid is recorded along with the rover's ground truth position and other state information. This information is saved as an individual file for each traverse enabling post processing and performance comparison of localization methods outside of the resource intensive simulator. Through the use of the provided playback utility, it is possible to synchronize multiple separate traverses, extending beyond multi-session localization to model multi-agent systems operating in the same environment.

V. Experimental Results

We evaluate the performance of LunarLoc using five pairs of traverses from the Lunar Simulator shown in Fig. 6, including all combinations of four traverses aside from the longest traverse with the shorter portion of itself, as Traverse 3 is a subset of Traverse 12. We report the number of segments found in each submap as well as the number of inlier associations returned during data association in Table IV. We deploy LunarLoc in real time. On a NVIDIA GeForce RTX 3090, the mean detection time of FastSAM runs in real time at 25 Hz [7] which is the slowest component of LunarLoc's front-end detection module.

TABLE III: Metrics recorded in each LAC archive

Sensor Group	Recorded Metrics (frequency)
Rover State @ 20Hz	Mission Time Accelerometer Measurements Gyroscope Measurements Power Consumption Control Input (linear and angular velocity) Rover Pose (position and orientation)
Camera @ up to 10Hz	Camera Images (subsampled to 0.5Hz) Camera Enable / Disable State Camera Light Intensity Camera / Light Position

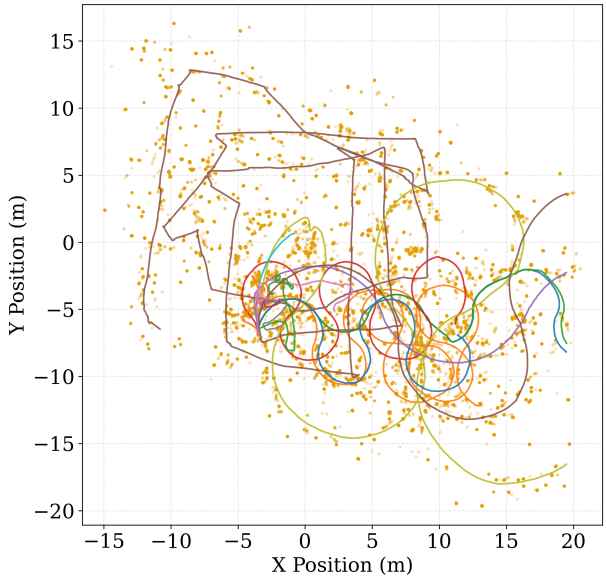


Fig. 7: All traverses available in the LunarLoc dataset shown in different colors where detected objects are shown as yellow dots.

LunarLoc uses the underlying structure of the environment, i.e. boulder locations, as a cue for global localization via a graph theoretic approach. With the number of segments detected ranging from 36 to 185 with up to 41% inlier associations, we show that rock detections are a reliable cue for data association in this setting.

To evaluate the localization performance of LunarLoc, we report the tracking performance defined by the root mean squared error of the translation component $\mathbf{t}$ between the vehicle map M_{veh} and the reference map M_{ref}. Transformations

TABLE IV: Tracking Performance on LunarLoc Dataset (RMSE ↓ [m]).

Path 1	Path 2	Segs in 1	Segs in 2	Ain	RMSE ↓ [cm]
3	5	36	62	15	0.08
3	7	36	102	22	0.61
5	7	62	102	28	1.79
5	12	62	185	18	0.32
7	12	102	185	41	0.68

returned by LunarLoc as shown in Table IV have localization error less than 2 cm in all evaluated cases, significantly outperforming the state of the art for localization on the Moon as shown in Table I.

VI. Conclusion

We propose LunarLoc, a global localization method using graph-theoretic data association of objects in the scene extracted using zero-shot segmentation. Future work includes integrating this localization framework into multi-agent settings and performing more extensive testing from different viewpoints and varying lighting conditions. We also plan to perform global localization with respect to an aerial map of the rock locations taken during landing. While place recognition approaches struggle with perceptual aliasing and feature-based methods may not be distinct enough to operate in various lighting conditions [6], LunarLoc is a promising approach using the underlying structure of the terrain as a cue for global localization on the Moon.

References

[1] R. Arandjelovic, P. Gronat, A. Torii, T. Pajdla, and J. Sivic, "Netvlad: Cnn architecture for weakly supervised place recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 5297–5307.

[2] M. Oquab, T. Darcet, T. Moutakanni, H. V. Vo, M. Szafraniec, V. Khalidov, P. Fernandez, D. Haziza, F. Massa, A. El-Nouby, R. Howes, P.-Y. Huang, H. Xu, V. Sharma, S.-W. Li, W. Galuba, M. Rabbat, M. Assran, N. Ballas, G. Synnaeve, I. Misra, H. Jegou, J. Mairal, P. Labatut, A. Joulin, and P. Bojanowski, "Dinov2: Learning robust visual features without supervision," 2023.

[3] D. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, pp. 91–, 11 2004.

[4] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "Orb: An efficient alternative to sift or surf," in *2011 International Conference on Computer Vision*, 2011, pp. 2564–2571.

[5] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[6] A. Thomas, J. Kinnari, P. C. Lusk, K. Kondo, and J. P. How, "Sos-match: Segmentation for open-set robust correspondence search and robot localization in unstructured environments," in *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2024, pp. 5613–5620.

[7] X. Zhao, W. Ding, Y. An, Y. Du, T. Yu, M. Li, M. Tang, and J. Wang, "Fast segment anything," 2023.

[8] M. Cummins and P. Newman, "Fab-map: Probabilistic localization and mapping in the space of appearance," *The International journal of robotics research*, vol. 27, no. 6, pp. 647–665, 2008.

[9] P.-E. Sarlin, C. Cadena, R. Siegwart, and M. Dymczyk, "From coarse to fine: Robust hierarchical localization at large scale," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 12 716–12 725.

[10] N. Keetha, A. Mishra, J. Karhade, K. M. Jatavallabhula, S. Scherer, M. Krishna, and S. Garg, "AnyLoc: Towards universal visual place recognition," *arXiv preprint arXiv:2308.00688*, 2023.

[11] D. DeTone, T. Malisiewicz, and A. Rabinovich, "Superpoint: Self-supervised interest point detection and description," in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 2018, pp. 224–236.

[12] M. Grimes and Y. LeCun, "Efficient off-road localization using visually corrected odometry," in *2009 IEEE International Conference on Robotics and Automation*. IEEE, 2009, pp. 2649–2654.

[13] T. Ort, L. Paull, and D. Rus, "Autonomous vehicle navigation in rural environments without detailed prior maps," in *2018 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2018, pp. 2040–2047.

[14] J. Ankenbauer, P. C. Lusk, and J. P. How, "Global localization in unstructured environments using semantic object maps built from various viewpoints," *arXiv preprint arXiv:2303.04658*, 2023.

[15] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," *arXiv preprint arXiv:1804.02767*, 2018.

[16] M. B. Peterson, Y. X. Jia, Y. Tian, A. Thomas, and J. P. How, "Roman: Open-set object map alignment for robust view-invariant global localization," *arXiv preprint arXiv:2410.08262*, 2024.

[17] K. Singh and J. J. Leonard, "Open-set semantic uncertainty aware metric-semantic graph matching," *arXiv preprint arXiv:2409.11555*, 2024.

[18] T. J. Parker, M. P. Golombek, and M. W. Powell, "Geomorphic/Geologic Mapping, Localization, and Traverse Planning at the Opportunity Landing Site, Mars," in *41st Annual Lunar and Planetary Science Conference*, ser. Lunar and Planetary Science Conference, Mar. 2010, p. 2638.

[19] M. Maimone, J. Biesiadecki, E. Tunstel, Y. Cheng, and C. Leger, "Surface navigation and mobility intelligence on the mars exploration rovers. in intelligence for space robotics," 01 2006.

[20] A. Johnson, A. Ansar, L. Matthies, N. Trawny, A. Mourikis, and S. Roumeliotis, *A General Approach to Terrain Relative Navigation for Planetary Landing*. [Online]. Available: https://arc.aiaa.org/doi/abs/10.2514/6.2007-2854

[21] J. V. Hook, R. Schwartz, K. Ebadi, K. Coble, and C. Padgett, "Topographical landmarks for ground-level terrain relative navigation on mars," in *2022 IEEE Aerospace Conference (AERO)*, 2022, pp. 1–6.

[22] F. van Diggelen and P. Enge, "The world's first gps mooc and worldwide laboratory using smartphones," in *Proceedings of the 28th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2015)*, Tampa, Florida, Sep. 2015, pp. 361–369.

[23] M. A. San Martin, D. S. Bayard, D. T. Conway, M. Mandic, and E. S. Bailey, "A Minimal State Augmentation Algorithm for Vision-Based Navigation without Using Mapped Landmarks," 2017. [Online]. Available: https://hdl.handle.net/2014/46668

[24] H. F. Grip, D. Conway, J. Lam, N. Williams, M. P. Golombek, R. Brockers, M. Mischna, and M. R. Cacan, "Flying a helicopter on mars: How ingenuity's flights were planned, executed, and analyzed," in *2022 IEEE Aerospace Conference (AERO)*, 2022, pp. 1–17.

[25] S. Daftry, Z. Chen, Y. Cheng, S. Tepsuporn, B. Coltin, U. Naam, L. M. Ma, S. Khattak, M. Deans, and L. Matthies, "Lunarnav: Crater-based localization for long-range autonomous lunar rover navigation," 2023. [Online]. Available: https://arxiv.org/abs/2301.01350

[26] D. Atha, R. Michael Swan, A. Cauligi, A. Bettens, E. Goh, D. Kogan, L. Matthies, and M. Ono, "Shadownav: Autonomous global localization for lunar navigation in darkness," *IEEE Transactions on Field Robotics*, vol. 1, p. 213–230, 2024. [Online]. Available: http://dx.doi.org/10.1109/TFR.2024.3462391

[27] Johns Hopkins University Applied Physics Laboratory, "Lunar autonomy challenge documentation," 2024, accessed: 2025-05-15. [Online]. Available: https://lunar-autonomy-challenge.jhuapl.edu/Challenge-Documentation/index.php#introduction

[28] P. C. Lusk, K. Fathian, and J. P. How, "Clipper: A graph-theoretic framework for robust data association," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 13 828–13 834.

[29] K. S. Arun, T. S. Huang, and S. D. Blostein, "Least-squares fitting of two 3-D point sets," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-9, no. 5, pp. 698–700, 1987.

[30] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, "CARLA: An open urban driving simulator," in *Proceedings of the 1st Annual Conference on Robot Learning*, 2017, pp. 1–16.