Robust Semantic Constellation Matching and SLAM for Planetary Rover Global Localization

Emma Seabrook
ISAE-Supaero
Toulouse, France
emma.seabrook@student.isae-supaero.fr

Alberto Ricci
ISAE-Supaero
Toulouse, France
alberto.ricci@student.isae-supaero.fr

Damien Vivet
ISAE-Supaero
Toulouse, France
damien.vivet@isae-supaero.fr

***Abstract*—Planetary rover localization remains a critical challenge due to the absence of Global Navigation Satellite Systems and communication latency that prevents real-time human intervention. While traditional SLAM-based methods enable local navigation, they suffer from cumulative drift and rely on loop closure, a constraint that is often impractical given the non-repetitive trajectories typical of planetary exploration. This work presents a SLAM-based localization framework that leverages a global re-localization module based on geometric constellation matching. Utilizing a YOLOv12 model, semantic rock detections are extracted from both local rover imagery and overhead orbital maps. These landmarks are matched using a triangle invariant descriptor and geometric hashing, allowing a RANSAC-based alignment process to estimate the rover's global pose. Integrated within an Extended Kalman Filter (EKF), this system demonstrates reliable global pose recovery and bounded drift correction in simulated Mars environments.**

***Index Terms*—Planetary localization, SLAM, global relocalization, map matching, RANSAC, semantic detection, EKF**

I. Introduction

Present Mars missions heavily depend on ground-based operators to periodically correct accumulated localization errors, severely limiting operational efficiency and autonomy of planetary rovers. Currently, as Martian rovers drive, they monotonically accumulate positional uncertainty over time. The technique employed by NASA to reduce this uncertainty is to manually input global coordinates for the rovers at the end of each drive [1]. While Perseverance's AutoNav system successfully piloted 88% of its 17.7 km traverse, its total autonomous range remained constrained by dead reckoning drift. Without these manual 'check-ins' to reset its global position, the unbounded growth of it's pose uncertainty would eventually exceed safe operational limits. This limitation demonstrates the need for a global-localization framework to reduce this positional uncertainty.

There exists three families of global matching algorithms: feature-based techniques, geometric techniques, and those using deep-learning. Feature-based techniques rely on extracting visual features from rover imagery which can then be matched to similar features extracted from satellite maps. For example, Wan et al. [2] used ASIFT features to create tie-points across different landing sites and bundle adjustment to localize across the sites. Geometric techniques are common in Lunar global localization where craters can be used to geo-locate against satellite maps [3]. Another common geometric technique includes horizon matching techniques where large craters and mountains seen on the horizon can be matched to Digital Elevation Maps (DEM) [4] [5]. Finally, deep-learning techniques train a neural network to match the rover's images to satellite maps directly [6] [7] [8].

Previous methods employed on the Spirit and Opportunity rovers used a bundle adjustment approach to match features across rover images and sometimes orbital maps [9]. With Curiosity, a mutual information matching approach was taken to relay MSL navcam orthomosaics to orbital HiRISE imagery [10]. Currently, scientists at JPL are employing a census transform to match panoramic NavCam images to orbital maps, achieving an accuracy of 0.36 m [1]. Though highly accurate, this technique requires the rover to stop the traverse to take a panoramic photo of its environment.

We present a novel EKF-SLAM and re-localization technique that leverages orbital satellite imagery to geo-locate planetary rovers in real-time. The technique uses a YOLO [11] model to detect rock-landmarks which are then used as measurements for an EKF-SLAM pipeline. The resulting map is matched using hash-tables and RANSAC to a global catalog of rocks created from satellite imagery. This localization pipeline can be performed in real-time thereby eliminating the need to stop the traverse to reduce the global positional uncertainty.

II. Methods

The proposed SLAM architecture leverages satellite imagery to perform global relocalization, thereby reducing the rover's positional uncertainty. The pipeline consists of three main modules: a semantic rock detector, a global geometric re-localizer, and an Extended Kalman Filter (EKF) SLAM framework.

A. Semantic Detection

Spatial landmarks are extracted using two YOLO-based object detection models. The first processes nadir-view satellite imagery to generate a global 2D catalog of rock locations and sizes, while the second operates on oblique stereo images from the rover to estimate landmarks in the local coordinate frame. Both models were trained on simulated images generated in Gazebo.

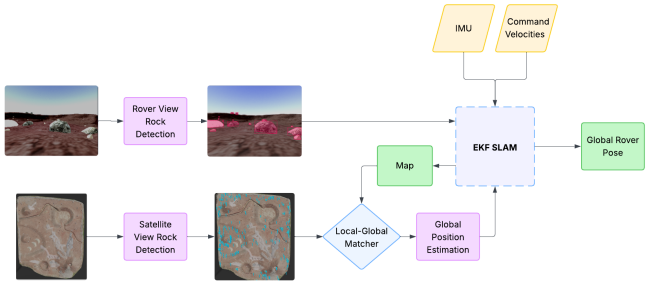


Fig. 1. Overview of the global pipeline. Two rock detectors are used to provide both local and global object detections. The onboard detections are then used in a classical EKF to provide SLAM-based localisation. By aligning the local SLAM map with the a-priori orbital map, the system can estimate the rover's absolute global pose.

Detection performance is summarized in Table I. Results indicate that the model performs significantly better in the rover viewpoint than in the satellite viewpoint. In both cases, bounding box detections (B) achieve higher precision and recall than segmentation masks (M). Overall, the best performance is obtained for rover-based bounding box detection, highlighting the advantage of closer, higher-resolution observations for reliable rock identification.

TABLE I
ROCK DETERCTION MODEL CHARACTERISATION. *B* AND *M* STAND FOR *bounding boxes* AND *masks* RESPECTIVELY

Model	# images	Precision	Recall	mAP50	mAP50-0.95
Satellite View (B)	491	0.440	0.364	0.379	0.223
Satellite View (M)	491	0.413	0.350	0.350	0.154
Rover View (B)	556	0.744	0.692	0.737	0.545
Rover View (M)	556	0.738	0.631	0.695	0.457

B. Global Relocalization

To correct odometry drift, the rover's global pose is estimated by matching a local sub-map of features detected during the rover's traverse with an overhead global map of landmarks extracted from satellite imagery. The matcher works by creating a hash table of landmark triangle invariants and using the RANSAC [12] algorithm to determine the best match in the global map.

1) Invariant Descriptors: The rock identifications from the global map are grouped into triangles and encoded with a scale and permutation invariant descriptor. This descriptor includes side length ratios, internal angles, and normalized area. A geometric hash table is then created with these descriptors to accelerate candidate matching.

2) RANSAC Map Alignment: A KD-tree is used to sample a rock with two of it's nearest neighbours from the local map. This triangle is then matched to the closest bin in the hash tree. A 2D similarity transform is then estimated between the two using a Procrustes alignment.

3) Validation: Before computing the transformation, candidate matches are rapidly filtered by comparing the physical dimensions of the local and global rocks, rejecting those that exceed a predefined size tolerance.

4) Two-stage scoring: The local map is then transformed using the estimated transformation and the number of inliers is obtained. If the transformation meets the inlier threshold, a new transformation is estimated using the inliers. This new transformation is then applied to the local map and re-scored.

5) RMS Tie-breaker: In the event that two (or more) transformations result in the same number of inliers, the transformation with the lower Root Mean Square Error (RMSE) is kept.

C. EKF-SLAM Integration

The semantic detection and global re-localization are then integrated into an EKF-SLAM framework. The EKF takes as inputs command velocities, IMU data, semantic rock detections, and global relocalization updates, to produce an estimate of the rover's pose over time.

1) State & Prediction: The state vector tracks the rover's 6-DOF pose alongside the 3D positions of local landmarks. The prediction step is driven by command velocities and IMU orientation increments, utilizing dynamic process noise.

2) Landmark Update: Rocks detected by the stereo-semantic detector are used as visual landmarks. The 3D-position of the rocks relative to the rover are used in the EKF update step.

3) Global Updates: Relocalization events act as planar (SE(2)) position fixes. To ensure stability and reject outliers, these updates are validated via a distance/yaw gating check and integrated as Kalman update steps.

4) Latency Handling: A state-history queue dynamically rolls back and replays odometry to seamlessly integrate delayed global measurements at their correct temporal location.

III. EXPERIMENTAL VALIDATION

To rigorously assess the performance and robustness of the integrated EKF-SLAM and global map-matching pipeline, a high-fidelity simulation campaign was conducted across diverse planetary topographies. The primary objective of these experiments was to evaluate the system's ability to mitigate cumulative odometry drift and maintain global localization accuracy in feature-sparse environments.

A. Experimental Setup

The pipeline is tested in multiple simulation environments to validate both the EKF and global-matching algorithms. The architecture was developed in Python and deployed using ROS 2 Jazzy on Ubuntu 24. All simulations were executed using the Gazebo Harmonic simulator.

1) Simulation Mars Yards: Testing was conducted across four diverse Martian topographies. Three environments were originally developed for the European Rover Challenge remote task by Sowa et al. [13], and the fourth is a 3D reconstruction of a physical Mars Yard. Table II describes the physical characteristics of the simulated Mars yards.

2) Rover Platform & Sensor Payload: The simulated platform is the NASA JPL Open Source Rover (OSR) [14]. The sensor payload included a 2304x1296pixel stereo camera updating at 30 Hz and an IMU sensor updating at 30 Hz.

3) Satellite Data Generation: To generate the global reference catalogs required for global localization, "satellite" maps were synthesized directly within Gazebo. A camera was spawned above the origin of each world, capturing a 6000x6000 pixel bird's-eye-view image.

TABLE II
GAZEBO SIMULATION ENVIRONMENT CHARACTERISTICS. *Note: MY indicates 'MarsYard'*

Map	Size ($X \times Y \times Z$) [m]	Format	# Rocks
MY2020	43.32 × 36.29 × 3.01	3D Mesh	2715
MY2021	47 × 47 × 6.58	Heightmap	740
MY2022	50 × 50 × 4.82	Heightmap	294
CNES MY	57 × 83 × 7.81	3D Mesh	468

B. Evaluation Metrics

The pipeline's accuracy is evaluated globally using Absolute Trajectory Error (ATE) and locally via Relative Pose Error (RPE). Relocalization efficacy is specifically measured through the success rate, threshold accuracy, yaw/position error distributions, and overall filter consistency (3σ bounds).

C. Results

The results of the pipeline testing can be found in Tables III and IV. To ensure consistency, rosbags were collected with the rover motions and the raw rock detections from the YOLO pipeline. The results for both the standard EKF-SLAM and the one with the global relocalization module enabled were collected using the same rosbag.

D. Global Matching Accuracy:

The global relocalization module demonstrated high reliability, producing global matches at an average rate of 0.089 Hz, and achieving an average positional error of 1.2m across the four simulation environments.

E. SLAM Trajectory & Drift:

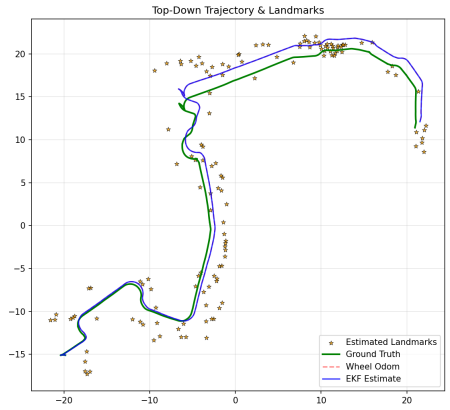
Integrating the global relocalization module effectively bounded the Absolute Trajectory Error (ATE). In contrast, the open-loop EKF baseline diverged significantly over time. Time-resolved 3-sigma consistency analysis (see Figure 3) confirmed that successful relocalization events systematically reduced the predicted covariance of the pose, highlighting the stabilizing effect of the global updates.

F. Relocalization Events

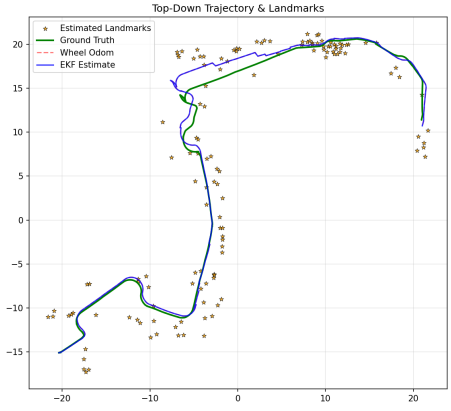
Figure 1 presents the ground truth, wheel odometry, and estimated paths of a traverse in the 2021 MarsYard with the relocalization events marked along the EKF trajectory. The relocalization events are infrequent in the beginning of the traverse but increase in frequency over time. As the landmark map grows over time, the global matcher has more landmarks to work with and can therefore make more frequent and more accurate matches.

This indicates that in regions with sparse features, or during short traverses, the pipeline will struggle to geo-locate. However, as the traverse increases in length, the re-localization events will become more frequent as more landmarks are added to the map.

The simulated traverses are relatively short when compared to the ones performed by real Mars rovers. In an application scenario, the map size should be bounded to prevent the EKF from being overloaded or the noise present in the map from preventing good matches.



(a) 2D-Trajectory Reconstruction for 2021 MY without Global Relocalization

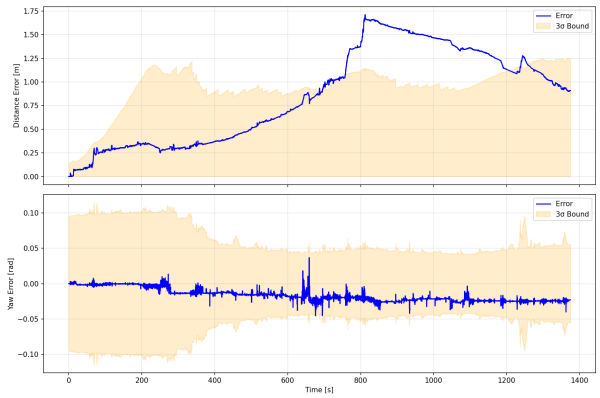


(b) 2D-Trajectory Reconstruction for 2021 MY with Global Relocalization.

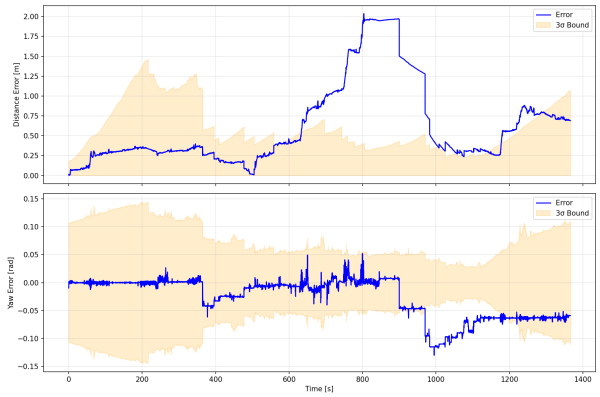
Fig. 2. Comparative analysis of 2D trajectory reconstruction in the XY plane. The EKF estimate (blue) maintains high fidelity to the ground truth (green) throughout the mission.

IV. DISCUSSION

Semantic landmarks provide higher-level features that reduce drift compared to purely geometric approaches; how-



(a) EKF error compared to 3-σ bound without global relocalization.



(b) EKF error compared to 3-σ bound with global relocalization.

Fig. 3. EKF position error compared to estimated 3 σ bounds. The solid blue line represents the absolute positional error of the rover relative to ground truth. The dashed red lines denote the filter's internal 3σ covariance envelope.

TABLE III
EKF-ONLY EVALUATION ACROSS SIMULATION ENVIRONMENTS

Metric	**CNES MY**	**MY2020**	**MY2021**	**MY2022**
Distance [m]	70.66	87.05	90.67	85.51
Duration [s]	496.9	905.57	1374.8	1046.0
EKF ATE [m]	1.93	1.70	0.91	1.21
EKF RMSE [m]	1.27	0.57	0.97	0.49

TABLE IV
FULL PIPELINE EVALUATION ACROSS SIMULATION ENVIRONMENTS

Metric	**CNES MY**	**MY2020**	**MY2021**	**MY2022**
EKF ATE [m]	1.79	1.42	0.91	0.82
EKF RMSE [m]	1.17	0.53	0.97	0.41
ATE Improvement [%]	7.8	19.7	0.0	47.561

ever, this reduction remains limited by the accuracy of the rock detection model. In this work, detection performance is constrained by the use of lightweight models trained on a small dataset, leading to increased landmark uncertainty and degraded SLAM accuracy. Future work will focus on improving these models and increasing the diversity of the dataset to diversify the environments and situations this framework can be used in.

Experiments are conducted in simulation, where wheel odometry exhibits minimal slip, resulting in optimistic drift estimates relative to real planetary conditions. In sparse environments, the reduced number of detectable landmarks further limits the robustness of the RANSAC-based matching. Future work will deploy this EKF-SLAM global-relocalization framework on physical robotic platforms to evaluate its real-world efficacy.

The detection model employed in this work is intentionally lightweight, reflecting the computational constraints of the target onboard hardware available for planned real-world deployment. Although experimental validation was ultimately conducted in simulation, these constraints guided the model design. This choice entails a trade-off, as reduced model capacity leads to lower detection accuracy and robustness compared to state-of-the-art approaches. Empirically, this manifests as increased noise in the predicted bounding boxes and segmentation masks, introducing uncertainty in landmark localization. Given the sensitivity of the EKF-SLAM framework to measurement quality, this noise propagates through the estimation process and degrades overall system performance.

Despite these limitations, global re-localization effectively compensates for accumulated drift by anchoring the local map to orbital data, providing a promising avenue for global-relocalization in GPS-denied environments.

V. CONCLUSIONS AND FUTURE WORK

This work demonstrates the viability of integrating semantic perception with geometric matching to achieve absolute localization for planetary rovers in GPS-denied environments.

Future work will focus on improving the robustness and accuracy of the perception pipeline, particularly for rock landmark detection under computational constraints. Enhancing detection quality is critical to reducing uncertainty in the SLAM estimate and improving matching reliability.

From a state estimation perspective, transitioning from an EKF-SLAM formulation to a smoothing-based approach could better handle delayed and asynchronous measurements. Additionally, extending the perception system to incorporate larger-scale features, such as craters and dune formations, would improve scalability to orbital datasets such as HiRISE imagery.

Finally, validation in real-world conditions will be necessary to account for factors not captured in simulation, including wheel slip, terrain variability, and sensor noise.

REFERENCES

[1] V. Verma, J. Nash, L. Saldyt, Q. Dwight, H. Wang, S. Myint, J. Biesiadecki, M. Maimone, A. Tumbar, A. Ansar, G. Kubiak, and R. Hogg,

"Enabling long precise drives for the perseverance mars rover via onboard global localization," in *2024 IEEE Aerospace Conference*, pp. 1–18, 2024.

[2] W. Wan, Z. Liu, K. Di, B. Wang, and J. Zhou, "A Cross-Site Visual Localization Method for Yutu Rover," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XL-4, pp. 279–284, 2014.

[3] A. Cauligi, R. M. Swan, H. Ono, S. Daftry, J. Elliott, L. Matthies, and D. Atha, "Shadownav: Crater-based localization for nighttime and permanently shadowed region lunar navigation," in *2023 IEEE Aerospace Conference*, pp. 1–12, 2023.

[4] S. Chiodini, M. Pertile, S. Debei, L. Bramante, E. Ferrentino, A. Villa, I. Musso, and M. Barrera, "Mars rovers localization by matching local horizon to surface digital elevation models," pp. 374–379, 06 2017.

[5] F. Cozman, E. Krotkov, and C. Guestrin, "Outdoor visual position estimation for planetary rovers," *Autonomous Robots*, vol. 9, pp. 135–150, 09 2000.

[6] I. Moreno, M. Azkarate, L. Gerdes, and C. Bahnsen, *Improving Global Localization Algorithms for Mars Rovers with Neural Networks*. PhD thesis, 06 2021.

[7] B. Wu, P. Ross W. K., P. Ludivig, A. S. Chung, and T. Seabrook, "Absolute localization through orbital maps and surface perspective imagery: A synthetic lunar dataset and neural network approach," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, p. 3262–3267, IEEE Press, 2019.

[8] V. Franchi and E. Ntagiou, "Planetary rover localisation via surface and orbital image matching," in *2022 IEEE Aerospace Conference (AERO)*, pp. 1–14, 2022.

[9] M. Maimone, Y. Cheng, and L. Matthies, "Two years of visual odometry on the mars exploration rovers," *Journal of Field Robotics*, vol. 24, no. 3, pp. 169–186, 2007.

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

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

[12] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, pp. 381–395, June 1981.

[13] B. Sowa and J. Hernas, "leo_gz_simulation (ROS 2)," 2024. GitHub Repository.

[14] NASA Jet Propulsion Laboratory, "Nasa jpl open source rover." https://github.com/nasa-jpl/open-source-rover, 2018. Accessed: 2025-06-16.