AI-Based Navigation for Lunar Satellites in VLLO

Gleisson Bezerra
ITA Space Center
Technological Institute of Aeronautics
São José dos Campos, Brazil
gleisson@ita.br

Pedro Varela
ITA Space Center
Technological Institute of Aeronautics
São José dos Campos, Brazil
pedropafvl@ita.br

Felipe Oliveira Tavares
ITA Space Center
Technological Institute of Aeronautics
São José dos Campos, Brazil
felipeo.tavares@ita.br

Renan Menezes
ITA Space Center
Technological Institute of Aeronautics
São José dos Campos, Brazil
renang@ita.br

Lidia Sato
ITA Space Center
Technological Institute of Aeronautics
São José dos Campos, Brazil
shibuya@ita.br

Luis Loures
ITA Space Center
Technological Institute of Aeronautics
São José dos Campos, Brazil
loures@ita.br

***Abstract*—This paper presents the design and preliminary experiments of the Lunar Intelligent Navigation via Neural Architecture (LINNA) payload, a technology demonstration integrated into the SelenITA mission—Brazil's first lunar CubeSat. The experiment is motivated by the future needs for autonomous state estimation in the Very Low Lunar Orbit (VLLO) regime, where Mascon-induced gravitational anomalies challenge traditional inertial-only navigation. The primary goal of LINNA is to evaluate the feasibility of conducting onboard, vision-based positioning as a future auxiliary system for autonomous Guidance, Navigation, and Control (GNC). LINNA operates as a standalone technological experiment, using a nadir-pointing optical sensor to capture lunar surface data at altitudes between 30 km and 100 km. The methodology focuses on an onboard processing pipeline where a lightweight Neural Network (NN) identifies surface landmarks to provide stable references for a Visual SLAM (Simultaneous Localization and Mapping) algorithm. High-fidelity simulations using synthetic lunar imagery are employed to assess the neural network's reliability under varying lighting conditions without compromising the mission's primary scientific objectives. By quantitatively measuring the vision-based system's ability to provide independent position and altitude updates, the project seeks to establish foundational flight-heritage for AI-driven payloads. The expected outcome is a comprehensive characterization of the system's potential to bound inertial drift in perturbed environments. These insights are essential for the future development of scalable and autonomous small satellite missions, contributing to the long-term roadmap of lunar exploration.**

***Index Terms*—Deep Learning, Machine Learning, Optical Engineering, Embedded AI, Computer Vision, Automation, Lunar Orbit**

I. Introduction

Autonomous navigation in lunar environments has evolved from inertial-based systems to advanced vision-based navigation (VBN) approaches capable of supporting precision operations. During the Apollo Program, navigation relied primarily on inertial sensors and dead-reckoning techniques. The Lunar Roving Vehicle navigation system, for instance, combined gyroscopic heading with wheel odometry, achieving acceptable short-range performance but exhibiting cumulative drift due to the absence of external corrections [1].

The introduction of Terrain Relative Navigation (TRN) concept as a subarea of VBN represented a major paradigm shift. TRN systems correlate onboard sensor data with pre-existing well-known references to estimate spacecraft position. [2] demonstrated that LIDAR-based TRN could achieve precision on the order of tens of meters, even under large initial state uncertainties. This transition marked the shift from purely propagative navigation to map-referenced estimation, which is essential for operations in gravitationally perturbed environments such as the Moon.

Subsequent developments extended TRN concepts to passive optical sensing. The OSIRIS-REx mission demonstrated Natural Feature Tracking (NFT), where onboard imagery is matched against known surface landmarks to refine navigation estimates in real time [3]. This architecture laid the groundwork for modern VBN pipelines.

Recent advances focus on crater-based optical navigation, leveraging the abundance, stability, and global distribution of lunar craters. The success of SLIM [4] mission demonstrated the first high-precision landing ($\approx$100 m accuracy) using VBN, providing strong flight heritage. Complementary work by [5] shows that neural-network-assisted crater detection can significantly improve navigation accuracy in descent scenarios.

The primary objective of LINNA payload is to evaluate the feasibility of a vision-based pipeline for position and altitude estimation during continuous orbital operations in the VLLO regime. By leveraging deep learning for robust landmark extraction, this work seeks to demonstrate a scalable approach to mitigate inertial drift in small satellite missions. The remainder of this paper is organized as follows: Section II discusses the state of the art and research gaps; Section III details the mission concept and system architecture; Section IV describes the end-to-end image processing pipeline; Section V presents the deep learning framework, including synthetic dataset generation and model selection; Section VI outlines the state estimation backend; Section VII analyzes the experimental results; and Section VIII concludes with lessons learned and future work directions.

II. State of the Art and Related Work

Current implementations are largely limited to descent and landing phases. The application of VBN to continuous orbital operation in VLLO, where mascon-induced perturbations significantly degrade inertial solutions remains insufficiently explored.

A. Crater Detection Strategies

Crater detection is a fundamental enabler of lunar VBN. The evolution of crater-based navigation is closely tied to the availability of high-quality global crater databases.

A major milestone in the development of global reference datasets is the work of [6], which introduced a comprehensive catalog of lunar impact craters larger than 1–2 km. This database contains over two million craters and provides accurate geometric and spatial information derived from multiple lunar missions. However, variability in human crater identification and regional inconsistencies highlight the need for automated, consistent, and navigation-oriented detection methods [7].

The transition to data-driven approaches has significantly improved detection robustness. Deep learning frameworks such as DeepMoon [8] demonstrated near-human performance in crater identification, while more recent architectures incorporate multi-modal inputs (e.g., imagery and topography) to enhance detection accuracy. However, these approaches are often designed for offline analysis and require adaptation for real-time navigation.

Recent work has increasingly focused on integrating crater detection directly into navigation pipelines. The LunaNet architecture [9] combines neural crater detection with crater identification and Extended Kalman Filtering (EKF), demonstrating substantial improvements in navigation accuracy.

In this context, "ai-enabled crater-based navigation for lunar mapping" [10] leverages deep learning not only for crater detection but also for mapping and localization consistency, emphasizing the role of AI in generating coherent spatial representations from sequential observations. A key contribution of this work is the demonstration that neural networks can be trained to produce navigation-relevant features, rather than purely geometric detections, improving robustness to illumination changes and partial crater visibility.

Complementing this perspective, "Crater navigation and timing for autonomous lunar orbital operations in small satellites" [11] extend crater-based navigation from descent scenarios to orbital operations, which is particularly relevant for VLLO. Their work introduces a framework in which crater observations are used not only for spatial localization but also for timing and orbit determination, effectively coupling landmark detection with orbital dynamics. By integrating crater measurements into filtering architectures, they demonstrate the feasibility of reducing orbital state uncertainty in small satellite missions without continuous ground support.

Overall, the field has evolved from heuristic detection methods to integrated, AI-driven navigation frameworks, supported by large-scale crater databases and advanced estimation techniques. The combination of global datasets, deep learning-based detection, and orbital navigation frameworks defines the current state of the art. However, challenges remain in achieving real-time onboard implementation, ensuring robustness across varying environmental conditions, and validating these systems in operational deep-space mission

B. Synthesis and Research Gap

Most studies still focus on descent, and there is limited validation of continuous VBN in orbital regimes under realistic constraints. In addition, the use of AI-based detection onboard spacecraft raises practical challenges related to computational resources, radiation effects, and real-time data handling. Although crater databases provide a solid foundation, their efficient onboard use for matching and state estimation is not yet fully resolved, and experimental validation in flight conditions is still limited.

The LINNA payload is designed to study the feasibility of onboard vision-based state estimation as a navigation approach for spacecraft operating in VLLO, while also evaluating the performance of lightweight neural networks under deep-space edge-computing constraints. Unlike much of the previous work, the focus here is on continuous orbital operation rather than descent-only scenarios, along with the integration of neural landmark detection and onboard filtering within limited computational resources.

III. Mission Concept and System Architecture

SelenITA is a 16U CubeSat mission designed to operate in low lunar orbit regimes, including VLLO during its science phases. The mission concept considered in this work includes commissioning in a near-circular orbit at approximately 150 km mean altitude, threshold science operations with a periapsis altitude of approximately 70 km, and enhanced science operations with a periapsis altitude of approximately 30 km.

LINNA is conceived as a technology demonstrator for AI-based optical navigation in LLO and VLLO conditions. Its operation is based on scheduled observation windows, during which a nadir-pointing optical sensor acquires lunar surface images for TRN and Visual SLAM experiments. The resulting navigation outputs are recorded for post-facto comparison with independent reference solutions and are not used as authoritative inputs for spacecraft guidance, navigation, or control.

Figure 1 summarizes the candidate physical architecture of the LINNA system. The onboard segment includes the camera module, processing unit, flight software, navigation application, temporary data storage, housekeeping and performance monitoring, and fault-detection support. The figure also shows the main spacecraft interfaces: the Electrical Power System provides power, the ADCS provides attitude and pointing context through the spacecraft OBDH, and the TM/TC subsystem supports command and telemetry exchange. On the ground side, mission and payload operations provide planning and configuration inputs, while the LINNA ground-processing environment validates the downlinked products.

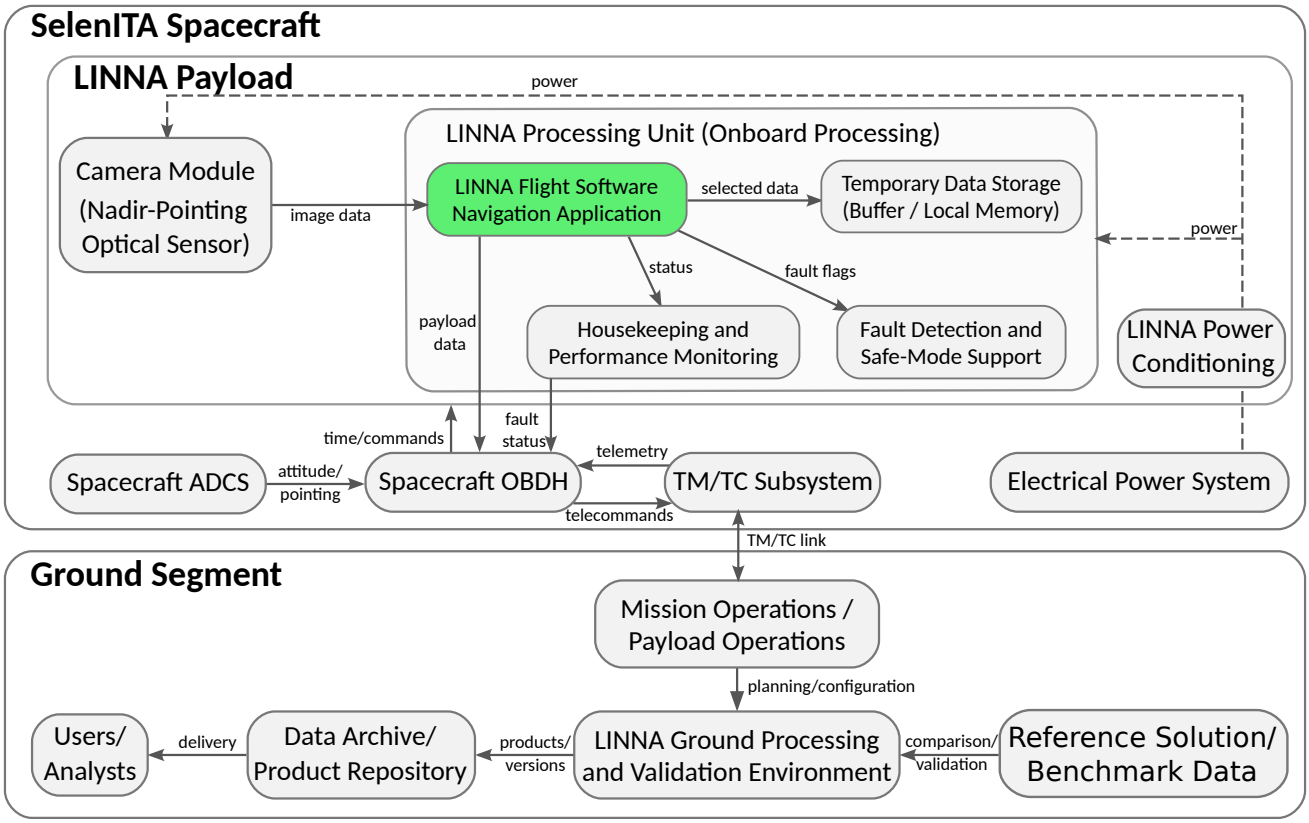


Fig. 1. The diagram shows the onboard payload elements, including the nadir-pointing camera module, processing unit, flight software, temporary data storage, housekeeping and performance monitoring, and fault-detection support. It also identifies the main interfaces with the SelenITA spacecraft.

IV. Image Processing Pipeline

While the following workflow is ultimately intended for edge-computing deployment, the current study focuses on the functional validation of the core algorithms rather than low-level hardware optimization.

The end-to-end target architecture is illustrated in Figure 2, providing the necessary system context through four fundamental stages: **image acquisition, pre-processing, crater detection model, and the state estimation backend.**

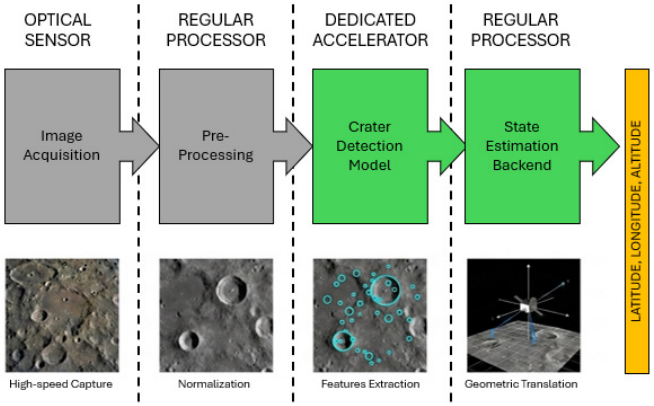


Fig. 2. Image Processing Stages.

Although the entire pipeline is outlined for architectural clarity, the primary focus of this work is the development and benchmarking of the **AI-based crater detection** and the subsequent **position and altitude estimation algorithm**. These stages represent the critical path for this navigational experiment and are described in detail in the specific sections.

A. Image Acquisition

The pipeline begins with the Image Acquisition stage, where a nadir-pointing optical sensor captures imagery of the lunar surface for experimental purposes. These captures provide the fundamental environmental data required to evaluate the robustness and performance of the proposed feature-extraction algorithms under representative orbital conditions.

At the conclusion of this stage, the optical data is formatted as a pixel matrix in **RGB (Red, Green, Blue)** format. Although the subsequent detection algorithms may rely on simplified inputs, maintaining the multi-channel color information at this stage is a deliberate architectural choice. By preserving the full bandwidth of the sensor at the point of acquisition, the experiment can better characterize how varying surface albedos and solar angles impact the initial digital representation of the terrain.

B. Pre-processing

The primary objective of this stage is to standardize the visual information, mitigating the inherent environmental chal-

lenges of the VLLO. Given that the lunar surface presents an extreme dynamic range, where saturated direct sunlight often coexists with deep shadows, this module performs **Image Normalization**. This process balances the exposure across the frame, ensuring that feature extraction remains consistent despite the harsh lighting conditions and varying solar incidence angles.

Following normalization, a **Grayscale Conversion** is executed to reduce the data dimensionality from the original multi-channel RGB matrix to a single-intensity channel. This step is fundamental for the experimental methodology of this work, as it simplifies the input features that the subsequent Deep Learning models must process.

C. Crater Detection Model

The third stage of the pipeline is a neural network architecture specifically selected for its ability to recognize lunar features under varying illumination conditions. The model's primary function is **Crater Landmark Extraction**, a process that translates the visual patterns in the normalized grayscale frames into structured data. By interpreting topographic textures and shadow gradients, the network identifies the **centroids and diameter ranges**, effectively isolating these stable landmarks from the complex background of the lunar regolith.

This capability is essential for establishing a reliable and continuous stream of reference points, ensuring that the landmarks identified are consistent enough to support the subsequent positioning algorithms.

D. Estate Estimation Backend

The final stage of the pipeline is the algorithm, where the geometric features extracted by the neural network are converted into navigational telemetry. This module utilizes a **Visual SLAM (Simultaneous Localization and Mapping)** framework to perform Geometric Estimation, correlating the detected crater landmarks—specifically their pixel coordinates and apparent diameters—with the camera's intrinsic parameters and the spacecraft's orbital geometry. This mathematical translation is the core of the **TRN experiment**, transforming two-dimensional visual data into a three-dimensional spatial context.

Through Landmark Mapping, the backend calculates the spacecraft's relative position and altitude. After that it estimates the absolute position using reference anchors. By comparing the observed geometry of the craters against the expected perspective from a Very Low Lunar Orbit (VLLO), the system provides vision-based state updates consisting of **Latitude, Longitude, and Altitude**.

V. Deep Learning for Crater Detection

The **Crater Detection Model** is a core component of the LINNA payload, responsible for extracting navigation-relevant landmarks from the acquired lunar imagery. Since no publicly available dataset adequately represents the viewing geometry, altitude range, and illumination conditions of the SelenITA mission in VLLO, a dedicated synthetic dataset was developed to support both training and evaluation.

A. Synthetic Dataset Generation

The core rendering pipeline was implemented using the Blender 3D software, which processes environmental data to simulate realistic orbital views and enables the generation of thousands of images at scale. Topographic information was sourced from the LRO LOLA SLDEM dataset (256 pixels/degree), covering 0° to 360°E longitude and -60° to 60°N latitude [12]. The surface texture was derived from the LROC WAC Empirically Normalized Mosaic at the 643 nm band, selected for its superior resolution of 304 pixels/degree [13].

1) Coordinate System and Spherical Mesh Generation: The rendering pipeline operates within a Moon-Centered Moon-Fixed (MCMF) Cartesian coordinate system, assuming a constant lunar radius of 1737.4 km. For each generated patch, a right-handed East-North-Up (ENU) local frame is established at the camera's nadir point, and a rotation matrix is constructed to transform coordinates between the local ENU frame and the global MCMF system. To intrinsically account for planetary curvature, the terrain was modeled as a spherical cap mesh rather than a flat plane: a uniform base grid of 65×65 vertices is analytically projected onto the lunar sphere at uniformly spaced latitude and longitude intervals within the camera's view. These vertices are then converted back to the local ENU frame for manipulation within the rendering engine, and the mesh undergoes adaptive subdivision at render time, dynamically refining the geometry based on a screen-space pixel density threshold of 0.5 pixels/polygon to optimize computational resources without pre-allocating excessive geometry.

2) Patch Extent and Curvature Correction: The visible ground extent required for each render was computed using pinhole camera geometry, applying separate latitude and longitude half-extents to correct for longitude compression at higher latitudes. The calculation accounts for the camera altitude, tilt angle, horizontal and vertical fields of view (FoV), and a safety margin applied to the final footprint. The angular bounds defining the mesh footprint are bounded between minimum and maximum values to prevent numerical degeneracies near the poles and to cap the extent at physically meaningful limits. Additionally, to mitigate rendering artifacts at light-shadow boundaries on the curved, displaced geometry, a shadow terminator offset was applied during the rendering process.

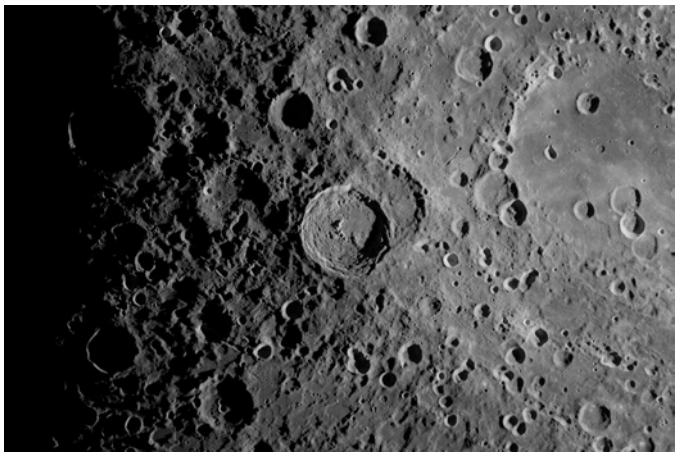
3) Surface Displacement and Seam Correction: The SLDEM data was ingested as floating-point tiles with elevation values in kilometers. For each generated patch, the corresponding elevation data is extracted and normalized to a $[0,\ 1]$ range, producing a displacement texture that is then mapped back to physical elevation offsets relative to the patch midpoint. Within the rendering engine, this normalized texture drives true geometric displacement of the spherical mesh, ensuring that topographic features such as crater rims, ridges, and basins are represented as actual surface geometry rather than mere texture shading. This approach preserves the physical fidelity of the terrain, which is essential for producing realistic shadow

patterns and parallax effects consistent with the spacecraft's viewing geometry.

4) Materials and Illumination: The material shading system utilizes a physically based Principled BSDF shader configured to mimic the micro-scale scattering behavior of the lunar regolith. The surface was modeled with high roughness and a constrained specular index of refraction, resulting in a highly diffuse, non-reflective appearance consistent with the Moon's optical properties. The base color was driven by the radiometrically calibrated WAC EMP reflectance map, which is processed with the same geographic reprojections and seam corrections applied to the SLDEM. Illumination is provided by a single directional light source mimicking the Sun, configured with a realistic angular diameter to produce soft shadows, and the world background strength is set to zero to accurately simulate the lunar vacuum and the absence of atmospheric scattering.

5) Final Render: In order to validate the visual and geometric fidelity of the proposed rendering pipeline, a comparative evaluation is presented using the Vavilov crater region, with Figure 3a showing a real photograph captured during the April 2026 Artemis II mission and Figure 3b illustrating its synthetic counterpart.

(a) Real photograph of the Vavilov crater captured during the Artemis II mission.



(b) Synthetic render generated by the proposed pipeline

Fig. 3. Qualitative comparison between empirical and synthetic imagery.

Image synthesis was performed using the Cycles path-tracing engine with adaptive sampling enabled to balance render quality and computational cost. The output undergoes Filmic tone mapping and OpenImageDenoise post-processing using the RGB, Albedo, and Normal passes, and was then exported as 16-bit PNG files. In parallel, the pipeline generates paired JSON metadata files containing the precise geographic coordinates of the image corners, the camera's spatial pose (latitude, longitude, altitude, tilt, and azimuth), and the solar illumination vectors.

B. Orbital Simulation and Dataset Compilation

The construction of the dataset required coupling the mission's orbital trajectory with the rendering pipeline to produce georeferenced synthetic images at each spacecraft position. This subsection details the three components of this process: the orbital parameters and propagation model used to generate the spacecraft ephemeris, the camera configuration adopted to maximize crater visibility, and the crater catalog and annotation strategy employed to produce ground truth labels.

1) Orbital Parameters: Further details on the SelenITA mission orbit design and requirements are described in [14]. For the purpose of this work, orbital ephemeris data were generated for a one-month period spanning from 7 August 2029 to 6 September 2029 using the Systems Tool Kit (STK) software. The initial orbital state is defined by the classical Keplerian elements listed in Table I, expressed in the Moon Inertial reference frame. The propagated state provides time-tagged altitude, latitude, and longitude referred to the MCMF reference frame at a temporal resolution of 30 s.

TABLE I
INITIAL ORBITAL ELEMENTS FOR THE SELENITA SPACECRAFT AT EPOCH 7 AUGUST 2029 00:00:00.000 UTCG, EXPRESSED IN THE MOON INERTIAL REFERENCE FRAME.

Parameter	Value
Semi-major axis	1768.180 km
Eccentricity	0.006908
Inclination	89.076°
RAAN	3.164°
Argument of perilune	333.604°
True anomaly	8.882°

The orbit was propagated using a 7th-order Runge–Kutta–Fehlberg integrator with 8th-order error control. The perturbation model includes third-body gravitational accelerations from the Sun and Earth, computed from planetary ephemeris tables, as well as the lunar gravitational field modeled through an expansion of Legendre polynomials up to degree and order 100, using coefficients derived from the GRAIL Lunar Gravity Model. Additionally, impulsive Hohmann-based maneuvers were incorporated to correct the orbital eccentricity to 0.006 whenever the periselene altitude decreased below 15 km, ensuring long-term orbit maintenance throughout the simulation window. Over the simulated period, the resulting trajectory exhibits a mean altitude of 30.8 km, with a minimum of 14.0 km and a maximum of 48.6 km.

2) Crater Catalogue and Annotation: The crater ground truth used for training and evaluation is derived from the global lunar impact crater database presented by [6], which constitutes a comprehensive census of over 2 million craters. For this work, craters with diameters between 5 and 30 km were selected from the catalogue. The lower bound of 5 km was chosen because smaller craters, which account for over 1.2 million entries in the database, occupy very few pixels at the mission's orbital altitudes and wide FoV, making them difficult to resolve reliably even by visual inspection. Retaining them would also dominate the class distribution and impose a finer network output resolution, increasing computational cost without proportional benefit. The upper bound of 30 km was set because craters above this threshold are scarce in the database (3,674 entries), and including them would further aggravate the class imbalance without meaningful gains in spatial coverage.

The selected range was divided into three diameter classes, as illustrated in Figure 4: Class 0 (5–10 km, 58,649 craters), Class 1 (10–15 km, 12,830 craters), and Class 2 (15–30 km, 7,908 craters). Although the distribution remains imbalanced, this partitioning enables class-aware training strategies and provides the detection model with explicit size discrimination, which benefits both detection performance and the accuracy of the subsequent position estimation stage.

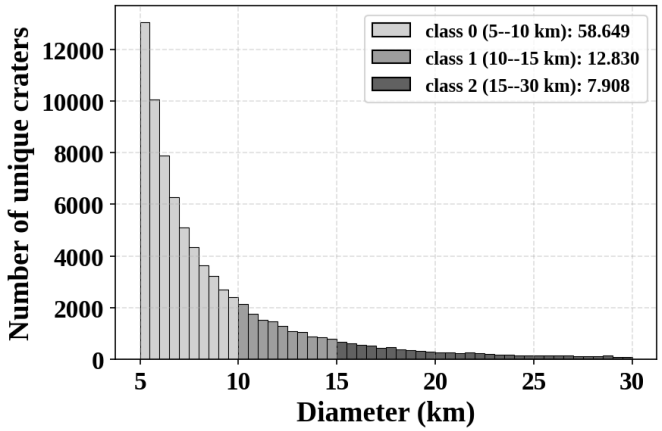


Fig. 4. Histogram of unique lunar crater diameters from the [6] catalogue (5–30 km), split into three classes: Class 0 (5–10 km), Class 1 (10–15 km), and Class 2 (15–30 km), shown in grayscale from light to dark.

Annotations were generated automatically by projecting the crater's catalogue onto each synthetic image using a local equirectangular approximation.

3) Camera Parameters: Since the spacecraft latitude, longitude, and altitude are fully determined by the orbital ephemeris and the camera is assumed to be nadir-pointing, the only free parameter governing the synthetic image generation is the camera FoV. To select an appropriate value, a statistical analysis was conducted over all orbital positions by varying the FoV from 10° to 160° and computing, for each configuration, the probability of capturing at least N craters from the catalogue (5–30 km diameter range) within the resulting ground footprint.

A FoV of 120° was selected as the minimum value that maintains a probability greater than 90% of capturing at least three crater per image. At this setting, the camera covers a horizontal ground extent of approximately 48.5 km at the minimum altitude of 14.0 km and 168.4 km at the maximum altitude of 48.6 km, providing sufficient spatial coverage for crater detection across the full range of orbital conditions.

4) Dataset Specifications: An image was rendered at every orbital position along the one-month ephemeris, excluding positions where the sub-satellite latitude exceeded 60° or fell below -60°, since the SLDEM topographic dataset does not provide coverage beyond these latitudes. All images were generated at a resolution of 512×512 pixels, matching the input size expected by the neural network model. To introduce illumination diversity and improve the robustness of the trained detector, the solar parameters were randomized for each render: the Sun elevation was uniformly sampled between 5° and 15°, and the Sun azimuth between -45° and 45°. This process yielded a total of 51,924 annotated images, with crater labels distributed across classes as follows: Class 0 (5–10 km) with 956,110 instances (72.8%), Class 1 (10–15 km) with 221,860 instances (16.9%), and Class 2 (15–30 km) with 135,098 instances (10.3%). Figure 5 shows samples from the synthetic dataset with annotated crater labels.

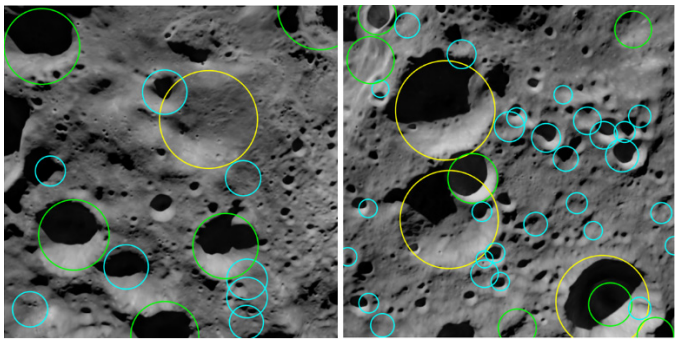


Fig. 5. Samples from the synthetic dataset with annotated crater labels (cyan: class 0, lime: class 1, yellow: class 2).

C. Training and Validation Metrics

The full set of 51,924 annotated images was partitioned into training, validation, and test subsets following an 80/15/5 split. To provide a realistic assessment of generalization along the orbital trajectory, the test set was extracted as a contiguous orbital segment, whereas the training and validation samples were drawn from the remaining shuffled images.

After **15 hours of training campaign using a RTX 4060 GPU**, a combination of models (MobileNetV3 [15] and CenterNet [16]) showed that detection performance improves with crater size, as larger craters occupy more pixels and exhibit more prominent shadow gradients, as detailed in Table II, facilitating their identification by the network. Class 2 achieves the highest precision (0.81), while Class 0, despite

being the most numerous, presents the lowest detection performance due to the limited pixel footprint of small craters at the mission's orbital altitudes. The localization error remains below 5 pixels for all classes, which is consistent with the spatial discretization imposed by the network stride.

TABLE II
PER-CLASS DETECTION PERFORMANCE ON THE TEST SET: PRECISION (P), RECALL (R) AND MEAN LOCALIZATION ERROR (LOC).

Class	P	R	loc (px)
Class 0 (5–10 km)	0.7021	0.3248	3.50
Class 1 (10–15 km)	0.7362	0.3586	4.11
Class 2 (15–30 km)	0.8131	0.4659	4.75

A qualitative assessment of the detection behavior is provided in Figure 6, where the predicted center points (marked with crosses) are overlaid against the corresponding ground truth annotations (dashed circles, drawn with the mean radius of each class). In these samples, no false positives are observed, and all detected craters are clearly visible on the lunar surface. Furthermore, the localization error is sufficiently small that the predicted centers consistently fall within the ground truth crater region, supporting the use of these detections as reliable landmarks for the subsequent vision-based state estimation stage.

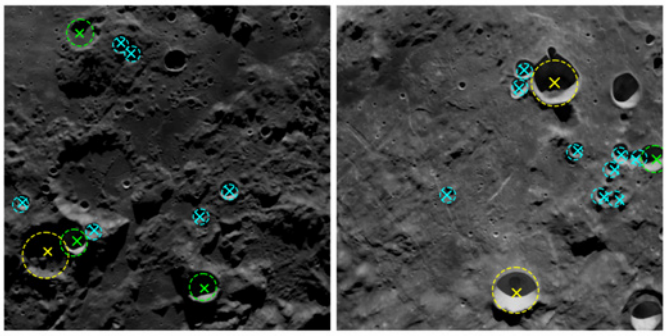


Fig. 6. Detection samples from the test set. Crosses (×) denote predictions and dashed circles denote ground truth annotations.

VI. VISION-BASED STATE ESTIMATION

The estimator avoids perspective-n-point matching against a crater catalogue and instead **uses frame-to-frame visual odometry** under three experimental assumptions:

- **Nadir camera, fixed yaw** — the optical axis points along the local radial-in direction at all times.
- **Polar near-circular orbit** — the satellite's instantaneous velocity is along a meridian.
- **Classified detections only** — the detector outputs no continuous radius, so the per-class median radius (in km) acts as a coarse scale prior.

The per-frame state has three unknowns: latitude, longitude and altitude h. The linearized pinhole projection of a stationary world point at altitude h relates each matched crater pair $(u_t, v_t) \rightarrow (u_{t+1}, v_{t+1})$ to the inter-frame translation $(\Delta x_{east}, \Delta y_{north}, \Delta h)$. fx, fy, cx and cy are intrinsic camera parameters:

$$h\Delta u = -f_x \Delta x_{east} - (u_t - c_x)\Delta h$$
$$h\Delta v = \quad f_y \Delta y_{north} - (v_t - c_y)\Delta h$$

Stacking $2N$ such rows from N matched pairs gives an over-determined linear system, solved by least squares. The recovered translation is converted to spherical $(\Delta\text{lat}, \Delta\text{lon}, \Delta h)$ and integrated [17].

A. Landmark-to-State Mapping

Crater identities are unknown — the detector returns image positions and a coarse size class only. Frame-to-frame matching exploits a single property: **all real correspondences share a common pixel translation**, while pairs between unrelated craters scatter in motion-space.

For every same-class candidate pair (p, q) with p in frame t and q in frame $t+1$, the implied ground shift is:

$$\boldsymbol{\Delta}_{pq} = (\Delta_{east}, \Delta_{north})_{p \rightarrow q} \quad \text{(km, scaled by the GT altitude)}$$

A translation specialization of geometric hashing [18]/RANSAC [19] picks the densest cluster:

$$c(p, q) = \left| \{ (p', q') : \|\boldsymbol{\Delta}_{pq} - \boldsymbol{\Delta}_{p'q'}\| < \tau \} \right|,$$
$$(p^\star, q^\star) = \arg\max_{(p,q)} c(p, q)$$

The cluster centre $\boldsymbol{\Delta}_{p^\star q^\star}$ is the recovered shift; agreeing pairs within τ are then assigned greedily one-to-one. If the cluster carries fewer than $N_{\min}$ agreeing pairs, the matcher falls back to **singleton-class** pairs — classes with exactly one detection in each frame, unambiguous by construction.

Pre-clustering motion bounds (orbital prior, scaled to the inter-frame sample rate δt):

$$|\Delta_{east}| < B_e, \qquad B_n^{\min} < \Delta_{north} < B_n^{\max}$$

Concrete values used in the run analysed below (at $\delta t = 30\,\text{s}$, scaled linearly with the actual cadence):

- $B_e \approx 2.5$ km/frame (east-motion bound)
- $[B_n^{\min}, B_n^{\max}] \approx [35, 65]$ km/frame (north-motion bounds)
- $\tau \approx 4$ km (cluster tolerance in shift space)
- $N_{\min} = 2$ (minimum agreeing pairs)

Working in km space rather than pixels keeps the bounds invariant to altitude and FoV variations across the dataset.

B. Altitude Estimation

The 3-DOF least-squares system uses the **ground-truth altitude** h as the scale factor in the projection equations. Using the integrated estimate would couple altitude drift back into in-plane recovery; the integrated altitude is therefore *reported* but not *fed back*.

For N matched pairs the per-frame system reads:

$$\underbrace{\begin{bmatrix} -f_x & 0 & -(u_t^{(1)} - c_x) \\ 0 & f_y & -(v_t^{(1)} - c_y) \\ \vdots & \vdots & \vdots \\ -f_x & 0 & -(u_t^{(N)} - c_x) \\ 0 & f_y & -(v_t^{(N)} - c_y) \end{bmatrix}}_{A \in \mathbb{R}^{2N \times 3}} \begin{bmatrix} \Delta x_{east} \\ \Delta y_{north} \\ \Delta h \end{bmatrix} = h \underbrace{\begin{bmatrix} \Delta u^{(1)} \\ \Delta v^{(1)} \\ \vdots \\ \Delta u^{(N)} \\ \Delta v^{(N)} \end{bmatrix}}_{b \in \mathbb{R}^{2N}}$$

Two regimes follow directly from the row count:

- **Multi-pair match** ($N \geq 2$, so $2N \geq 4$ rows against three unknowns) — solve $A\mathbf{x} = h\mathbf{b}$ in least-squares sense; all three components, Δh included, are recovered.
- **Singleton fall-back** ($N = 1$, so only two rows) — underdetermined; lock $\Delta h = 0$ and solve the remaining 2×2 system in closed form for $(\Delta x_{east}, \Delta y_{north})$.

The altitude estimate therefore tracks ground truth whenever multi-pair matches are available, and stays flat through stretches dominated by single-pair fall-backs.

C. Error Bounding Strategy

Loose matching produces spectacular failures when detector recall is sparse — random low-magnitude clusters can out-vote the real (small but consistent) one. Five layers prevent that:

1) **Hard motion bounds** at the candidate-pair stage:

$$|\Delta_{east}| < B_e, \qquad B_n^{\min} < \Delta_{north} < B_n^{\max}.$$

2) **Direction-of-motion safety filter**: A polar pass advances monotonically in latitude; reject the matcher result if the recovered north component flips sign:

$$\text{reject if} \quad \operatorname{sign}\big(\Delta_{north}^{(t)}\big) \cdot \operatorname{sign}\big(\Delta_{north}^{(t-1)}\big) < 0.$$

3) **Constant-velocity prediction** [20] when the matcher fails — extrapolate the last successful motion for at most K frames (here $K = 20$):

$$\hat{x}_{t+1} = \hat{x}_t + \boldsymbol{\Delta}_{t^\star}, \qquad t - t^\star \leq K.$$

4) **Time-gap reset**: Re-anchor when consecutive frames are more than T_{gap} apart (the dataset omits polar passes, producing minute-scale gaps):

$$\text{reset if} \quad t_{i+1} - t_i > T_{\text{gap}} \ (\approx 60\,\text{s}).$$

No-match timeout (optional): Re-anchor when no real match has succeeded for T_{timeout} seconds — bounds long pure-prediction stretches:

$$\text{reset if} \quad t - t_{\text{last match}} > T_{\text{timeout}}.$$

VII. Experimental Results

This section documents the estimation pipeline simulation session using a simulator (figure 9) built to run all experiments.

The session analyzed below consists in **5,196 images with 15 seconds interval** and covers roughly **24 hours of a simulated orbit with an average 30 km altitude**. Every plot's time axis uses the **actual simulation clock time** (parsed from the dataset's frame identifiers, which encode hour, minute, second and date) rather than an elapsed offset.

A. Visual Odometry Clusters

Every logged frame falls into one of six matcher outcomes:

- **Anchored match** — reset event to the true pose (initialization, time-gap, or safety timeout);
- **Multi-pair match** — several agreeing crater pairs converged on a single shift;
- **Singleton match** — only one unambiguous pair survived but the orbital prior validated it;
- **Prediction-only** — no measurement that frame, but the previous successful motion was extrapolated forward;
- **Missed** — no measurement and the predictor had nothing left to extrapolate from;
- **Rejected** — a candidate cluster was discarded by the safety filter that catches motion in the wrong direction;

As showed in figure 7, a dominant **multi-pair match** slice indicates the matcher is finding the orbital motion reliably. A heavy **prediction-only** or **miss** tail signals that detector recall (or class stability) is below what the matcher needs, and a non-trivial **rejected** count means phantom clusters at the wrong shift were trying to mislead the matcher.

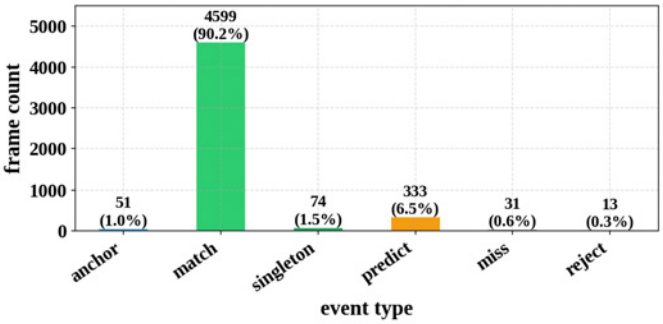


Fig. 7. Types and distribution of events during clusters tracking cycles.

Zoomed to a one-hour slice (figure 8), the cause-and-effect relationship between resets and the position error becomes resolvable. Gray dashed lines are annotated with the reset trigger (initialization, time-gap, manual step, or safety timeout); expect the trace to drop sharply right after one of them, since the integrated state was reset to ground truth at that instant.

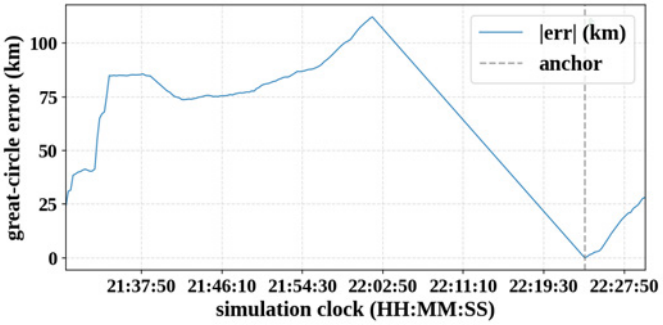
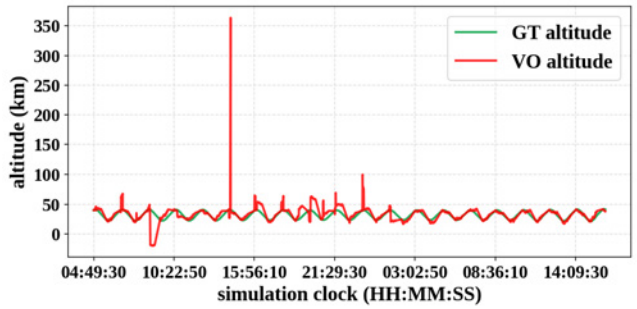


Fig. 8. Accumulated latitude and longitude error over time.

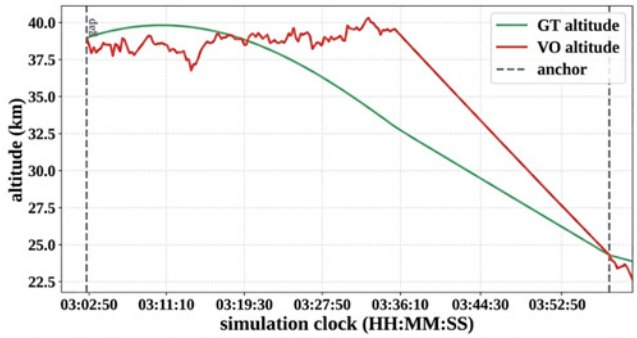
Fig. 9. SelenITA spacecraft simulator generating LINNA estimates according to a test dataset that was not part of the training.

B. Visual Odometry Altitude

The red line below (figure 10) estimate visually overlays the green ground truth for most of the run, demonstrating that the 3-DOF least-squares recovers the orbital altitude profile faithfully whenever measurements are available.



(a) Altitude estimation over time during 24 hours simulation.



(b) Zoomed to a one-hour slice of simulation clock time, the individual reset events become resolvable.

Fig. 10. Detailed view of altitude predictions without reset events.

Each dashed line is annotated with the trigger that fired it:

- **Initialization** (the very first frame, before any history exists);
- **time-gap reset** (a multi-minute gap between consecutive frames — the dataset omits polar passes, so this triggers automatically once per orbital pass);
- **manual step** (the operator paused or stepped the live visualizer, so consecutive frames are no longer continuous in time);
- **safety timeout** (the optional watchdog forced a re-anchor because no real match had succeeded for too long).

VIII. Conclusion and Future Work

This paper presented the preliminary design and experimental validation of the LINNA payload, a vision-based navigation experiment for the SelenITA mission.

A key distinction between this work and the current state of the art in lunar vision-based navigation lies in the focus on the Very Low Lunar Orbit (VLLO) regime, where traditional gravity models fail and optical feature stability is compromised by high orbital velocities. Unlike existing solutions that often focus on higher lunar orbits or rely on high-performance computing clusters, the LINNA architecture was designed from its inception to be modular. By validating this pipeline through a streamlined approach—prioritizing feature relevance over raw complexity—this study establishes a robust foundation for a future transition to embedded flight hardware. The modularity of the framework allows for subsequent integration of hardware-specific optimizations without compromising the core geometric stability of the experiment.

Despite the promising performance of the current pipeline, several refinements are planned for subsequent development

phases. First, the AI **Crater Detection Model** requires further expansion to include a more diverse set of object classes beyond generic craters. By categorizing landmarks by size, age, or morphological distinctness, the model's overall detection precision can be significantly enhanced, reducing ambiguity during the feature matching process.

Furthermore, while this experiment utilized Ground Truth (GT) resets to maintain the navigational lock, an operational version will necessitate the integration of a **minimum anchoring catalog**. Such a baseline dataset will allow for periodic recalibration of the state estimates, eliminating the reliance on simulated resets and providing a robust reference for absolute positioning.

Finally, future research will focus on retraining the neural network with a refined precision metric to better balance **False Positives (FPs)**. Preliminary observations indicate that many FPs are, in fact, real lunar craters not present in the current annotations; these "hidden" landmarks contribute valuable data to the state estimation backend. However, as the model still identifies non-crater features as landmarks, a strategic retraining approach is required to embrace valid unannotated detections while suppressing actual noise. This refinement is crucial for stabilizing the estimation logic and ensuring the reliability of vision-based navigation for the next generation of lunar CubeSats.

ACKNOWLEDGMENT

The authors would like to thank the **ITA Space Center and the Aeronautics Institute of Technology (ITA)** for the technical support and infrastructure provided during this research. This work was partially supported by financial support provided by the project FINEP 01.22.0488.00 (1308/22).

REFERENCES

[1] E. Smith and L. Mastin, "Lunar roving vehicle navigation system performance review," NASA, Technical Report, 1973.

[2] A. Johnson and T. Ivanov, "Analysis and testing of a lidar-based approach to terrain relative navigation for precise lunar landing," NASA, Technical Report, 2011.

[3] D. A. Lorenz, R. Olds, A. May, C. Mario, M. E. Perry, E. E. Palmer, and M. Daly, "Lessons learned from osiris-rex autonomous navigation using natural feature tracking," in *2017 IEEE Aerospace Conference*, 2017, pp. 1–12.

[4] T. Ishida, S. Fukuda, K. Kariya, H. Kamata, K. Takadama, H. Kojima, S. Sawai, and S. Sakai, "Vision-based navigation and obstacle detection flight results in slim lunar landing," *Acta Astronautica*, vol. 226, pp. 772–781, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0094576524006490

[5] S. Andolfo, A. Genova, F. V. Buonomo, A. M. Gargiulo, M. El Awag, P. Federici, R. Teodori, R. La Grassa, C. Re, and G. Cremonese, "Neural network-aided optical navigation for precise lunar descent operations," *Aerospace*, vol. 12, no. 3, 2025. [Online]. Available: https://www.mdpi.com/2226-4310/12/3/195

[6] S. J. Robbins, "A new global database of lunar impact craters greater than 1–2 km: 1. crater locations and sizes, comparisons with published databases, and global analysis," *Journal of Geophysical Research: Planets*, vol. 124, no. 4, pp. 871–892, 2019. [Online]. Available: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018JE005592

[7] S. J. Robbins, I. Antonenko, M. R. Kirchoff, C. R. Chapman, C. I. Fassett, R. R. Herrick, K. Singer, M. Zanetti, C. Lehan, D. Huang, and P. L. Gay, "The variability of crater identification among expert and community crater analysts," *Icarus*, vol. 234, pp. 109–131, 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0019103514001080

[8] A. Silburt, M. Ali-Dib, C. Zhu, A. Jackson, D. Valencia, Y. Kissin, D. Tamayo, and K. Menou, "Lunar crater identification via deep learning," *Icarus*, vol. 317, pp. 27–38, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0019103518301386

[9] L. Downes, T. J. Steiner, and J. P. How, "Deep learning crater detection for lunar terrain relative navigation," in *AIAA SciTech 2020 Forum*, 2020, p. 1838.

[10] S. McLeod, C.-K. Chng, M. Rodda, and T.-J. Chin, "Ai-enabled crater-based navigation for lunar mapping," *arXiv preprint arXiv:2509.20748*, 2025.

[11] Z. R. McLaughlin, R. E. Gold, S. G. Catalan, B. A. Jones, and R. Zanetti, "Crater navigation and timing for autonomous lunar orbital operations in small satellites," in *Proceedings of the 44th Annual American Astronautical Society Guidance, Navigation, and Control Conference, 2022*, M. Sandnas and D. B. Spencer, Eds. Cham: Springer International Publishing, 2024, pp. 155–171.

[12] M. Barker, E. Mazarico, G. Neumann, M. Zuber, J. Haruyama, and D. Smith, "A new lunar digital elevation model from the lunar orbiter laser altimeter and SELENE terrain camera," *Icarus*, vol. 273, pp. 346–355, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0019103515003450

[13] A. Boyd, M. Robinson, and H. Sato, "Lunar Reconnaissance Orbiter wide angle camera photometry: An empirical solution," in *43rd Annual Lunar and Planetary Science Conference*, no. 1659, 2012, p. 2795.

[14] T. Matos, L. Loures, L. Sato, V. Rodrigues, T. Carvalho, R. Menezes, B. Schuster, D. Vieira, P. Acosta, D. Arena *et al.*, "Engineering challenges of a cubesat mission around the moon: First steps on the path to selenita," in *37th Annual Small Satellite Conference*, 2023.

[15] A. G. Howard, M. Sandler, G. Chu, L.-C. Chen, B. Chen, M. Tan, W. Wang, Y. Zhu, R. Pang, V. Vasudevan, Q. V. Le, and H. Adam, "Searching for mobilenetv3," *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 1314–1324, 2019.

[16] X. Zhou, D. Wang, and P. Krähenbühl, "Objects as points," *arXiv preprint arXiv:1904.07850*, 2019.

[17] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge University Press, 2004.

[18] Y. Lamdan and H. Wolfson, "Geometric hashing: A general and efficient model-based recognition scheme," in *[1988 Proceedings] Second International Conference on Computer Vision*, 1988, pp. 238–249.

[19] 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, no. 6, p. 381–395, Jun. 1981. [Online]. Available: https://doi.org/10.1145/358669.358692

[20] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with applications to tracking and navigation: theory algorithms and software*. John Wiley & Sons, 2001.