



HAL
open science

MoonAnything: A Vision Benchmark with Large-Scale Lunar Supervised Data

Clémentine Grethen, Yuang Shi, Simone Gasparini, Géraldine Morin

► To cite this version:

Clémentine Grethen, Yuang Shi, Simone Gasparini, Géraldine Morin. MoonAnything: A Vision Benchmark with Large-Scale Lunar Supervised Data. MMSys '26: ACM Multimedia Systems Conference 2026, Apr 2026, Hong Kong Hong Kong, Hong Kong SAR China. pp.457-463, <10.1145/3793853.3799814>. <hal-05585915>

HAL Id: hal-05585915

<https://hal.science/hal-05585915v1>

Submitted on 9 Apr 2026

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

MoonAnything: A Vision Benchmark with Large-Scale Lunar Supervised Data

Clémentine Grethen¹, Yuang Shi^{1,2,3}, Simone Gasparini^{1,3}, Géraldine Morin^{1,3}

¹IRIT - Université de Toulouse, France ²National University of Singapore, Singapore ³IPAL, IRL2955, Singapore

Abstract

Accurate perception of lunar surfaces is critical for modern lunar exploration missions. However, developing robust learning-based perception systems is hindered by the lack of datasets that provide both geometric and photometric supervision. Existing lunar datasets typically lack either geometric ground truth, photometric realism, illumination diversity, or large-scale coverage. In this paper, we introduce MoonAnything, a unified benchmark built on real lunar topography with physically-based rendering, providing the first comprehensive geometric and photometric supervision under diverse illumination with large scale. The benchmark comprises two complementary sub-datasets : i) LunarGeo provides stereo images with corresponding dense depth maps and camera calibration enabling 3D reconstruction and pose estimation; ii) LunarPhoto provides photorealistic images using a spatially-varying BRDF model, along with multi-illumination renderings under real solar configurations, enabling reflectance estimation and illumination-robust perception. Together, these datasets offer over 130K samples with comprehensive supervision. Beyond lunar applications, MoonAnything offers a unique setting and challenging testbed for algorithms under low-textured, high-contrast conditions and applies to other airless celestial bodies and could generalize beyond. We establish baselines using state-of-the-art methods and release the complete dataset along with generation tools to support community extension: <https://github.com/clementinegrethen/MoonAnything>.

CCS Concepts

• **Computing methodologies** → **Reconstruction; Neural networks; Reflectance modeling.**

Keywords

Autonomous exploration, 3D reconstruction, Vision based navigation, BRDF, Dataset, Lunar reflectance

1 Introduction

Accurate perception of lunar surfaces is fundamental to modern lunar exploration. With renewed global interest from NASA’s Artemis program, China’s Chang’E missions, and commercial ventures [4, 18, 27], three perception capabilities become critical: *Terrain Relative Navigation* (TRN) for precise landing without GPS, *Hazard Detection and Avoidance* (HDA) for identifying unsafe terrain during descent, and *3D surface reconstruction* for rover navigation [8, 10, 11]. These capabilities rely on computer vision algorithms that operate reliably under the Moon’s challenging conditions.

The lunar environment presents challenges for visual perception. The lack of atmosphere creates pitch-black shadows with extreme

contrast [2], while the textureless regolith surface causes traditional feature detectors to fail [5, 21]. These factors create a substantial domain gap, thus deep learning models trained on Earth imagery fail on lunar scenes, producing unreliable geometry estimations with flat reconstructions and inconsistent relief [5].

High-quality datasets are essential for developing robust lunar perception systems. Recent 3D reconstruction models such as DUST3R [26], MAST3R [14], and VGGT [25] achieve remarkable performance on human-scale imagery precisely because they are trained on millions of real-world and synthetic images spanning indoor, outdoor, and object-centric scenes. However, these models do not generalize to lunar imagery without domain-specific adaptation. Although fine-tuning on lunar data has proven effective [6], such adaptation requires comprehensive datasets that capture the unique characteristics of lunar perception. Effective lunar vision datasets must provide three interrelated information:

- *Geometry supervision* through stereo pairs and dense ground truth depth maps enables training of 3D reconstruction models that are then able to infer lunar terrain structure.
- *Appearance modeling* through an accurate surface reflectance model is essential for generating photorealistic simulations that support algorithm development and validation.
- *Illumination variation* is critical because lunar missions must operate across diverse lighting conditions, from the harsh shadows of polar regions to varying sun angles throughout the lunar day. Datasets that capture only single lighting conditions produce models that fail when deployed under different illumination.

Existing lunar datasets address these requirements only partially. Real mission data from NASA’s Lunar Reconnaissance Orbiter (LRO) and China’s Chang’E program provide valuable orbital and surface imagery, but lack the dense geometric ground truth required for supervised learning [24, 27]. Synthetic datasets offer controlled supervision but have significant limitations. Generating photorealistic lunar imagery requires accurate modeling of surface reflectance through a Bidirectional Reflectance Distribution Function (BRDF), which describes how light reflects from a surface as a function of illumination and viewing angles. For the Moon, this is particularly important because lunar regolith exhibits unique optical properties, including opposition surge and anisotropic scattering [23], that differ substantially from terrestrial materials. Furthermore, real lunar surfaces are not homogeneous: reflectance varies spatially across the terrain, requiring Spatially-Varying BRDF (SVBRDF) models for accurate rendering. For example, LuSNAR [15] provides stereo pairs with depth maps and semantic labels using Unreal Engine, but lacks physically accurate BRDF, limiting photometric realism. The Polar Optical Lunar Analog Reconstruction (POLAR) stereo dataset [29] offers laboratory-controlled imagery with precise ground truth, but covers only a limited size terrain mockup under fixed lighting.



This work is licensed under a Creative Commons Attribution 4.0 International License.

Table 1: Comparison of lunar perception datasets. MoonAnything is the first to provide comprehensive geometric and photometric supervision with diverse illumination.

Dataset	Type	Region	Stereo	Depth	Reflect.	Multi-Light	Phys. Render	Scale
LRO NAC [24]	Real/Orbital	Global	×	×	×	×	N/A	2M images
Chang'E [27]	Real/Surface	Global	×	×	×	×	N/A	7.5K images
DLR's TRON [12]	Lab	Mockup	✓	×	×	×	×	7.2K images
POLAR [29]	Lab	Mockup	✓	✓	×	×	×	2.6K pairs
LuSNAR [15]	Synthetic	Simulated	✓	✓	×	×	×	13K images
StereoLunar [5]	Synthetic	South Pole	✓	✓	×	×	✓	50K pairs
LunarG2R [6]	Real	Tycho Crater	×	×	✓	×	✓	83K pairs + DEM
MoonAnything	Real&Synth.	Tycho&S. Pole	✓	✓	✓	✓	✓	130K pairs + DEMs + SVBRDF

Our prior works partially addressed these challenges through two datasets: StereoLunar [5] provides stereo imagery with depth supervision for the South Pole, while LunarG2R [6] provides geometry and reflectance pairs for Tycho crater. However, these datasets remain isolated. StereoLunar lacks reflectance supervision, LunarG2R lacks depth ground truth, and neither provides multi-illumination variation. Critically, *no existing dataset provides comprehensive geometric and photometric supervision under systematic illumination variation.*

In this paper, we introduce MoonAnything, a comprehensive benchmark for lunar surface perception that unifies and extends our previous works to address the limitations discussed above. MoonAnything comprises two complementary sub-datasets:

- LunarGeo is designed for geometric perception, providing stereo imagery with dense depth supervision across diverse viewing conditions and lunar regions. It extends LunarStereo [5] by adding stereo imagery of the Tycho crater region, complementing the original South Pole coverage and enabling cross-region generalization studies.
- LunarPhoto is designed for appearance modeling, providing learned reflectance parameters and multi-illumination renderings that capture spatially-varying surface properties. It extends LunarG2R [6] with multi-illumination renderings generated using learned SVBRDF parameters, supporting photometric stereo and illumination-robust training.

MoonAnything provides the first unified benchmark offering large-scale geometric and photometric supervision under diverse illumination (see Table 1). The dataset comprises over 130K pairs with corresponding DEMs and SVBRDF models (*an order of magnitude larger than existing lunar datasets*), establishing a new standard for data-driven lunar perception research. Beyond lunar tasks, MoonAnything serves broader computer vision research. The SVBRDF data enables material appearance modeling, while the textureless, high-contrast imagery provides a challenging benchmark for evaluating algorithm robustness under edge-case conditions applicable to other airless celestial bodies and low-feature environments. Experiments establish baselines using state-of-the-art methods and demonstrate the dataset's utility for advancing lunar perception.

The rest of the paper is organized as follows. Section 2 reviews existing lunar datasets and their limitations. Section 3 details the data generation pipeline. Section 4 presents the dataset statistics and structure. Finally, we demonstrate baseline experiments using SOTA methods in Section 5 and conclude in Section 6.

2 Related Work

Lunar perception datasets fall into three main categories: i) synthetic simulations, ii) real mission imagery, and iii) laboratory mockups. We review each category and identify the gaps in the following.

Synthetic simulation datasets use graphics engines or planetary renderers to generate labeled lunar scenes with controlled ground truth. The Artificial Lunar Landscape dataset [20] employs Terragen to create approximately 10,000 photorealistic images with pixel-wise semantic labels for sky, small rocks, and large rocks. However, it provides no geometric ground truth, such as depth maps or camera poses. LuSNAR [15] uses Unreal Engine to generate high-resolution stereo pairs along with depth maps, semantic labels, and simulated LiDAR/IMU data, offering a more comprehensive sensor suite. Space-specific tools such as PANGU [19] and SurRender [13] provide physically-based rendering capabilities tailored to planetary environments. Despite their utility, these synthetic datasets typically lack real lunar topography, relying instead on procedurally generated or artist-designed terrain. Viewpoint diversity is often restricted to ground-level or near-nadir trajectories. Most critically, simplified lighting and reflectance models fail to capture the complex photometric behavior of lunar regolith.

Real lunar image datasets provide authentic lunar appearance but lack geometric supervision. The Chang'E dataset [27] aggregates $\approx 7,500$ descent-phase images from the Chang'E-3 and Chang'E-4 landers. While these images provide an authentic representation of the lunar surface, they lack accurate 3D ground truth and calibrated camera poses. Chandrayaan-2's Orbiter High Resolution Camera acquired stereo triplet sequences at ≈ 0.3 m/px resolution [9], yet image calibration parameters, precise poses, and aligned terrain models remain unpublished. NASA's Lunar Reconnaissance Orbiter (LRO) Narrow Angle Camera (NAC) delivers very high-resolution panchromatic imagery down to 0.5 m/px [22], but these narrow swaths have sparse overlap and limited stereo baselines. A fundamental limitation of real lunar datasets is that orbital and descent imagery is predominantly nadir-viewing with minimal lateral baseline, making these datasets unsuitable for supervised 3D perception because of the lack of 3D ground truth.

Laboratory mockup datasets provide controlled imagery with precise ground truth over physical terrain mockups. The German Aerospace Center (DLR)'s TRON facility captures images over a 4×2 m lunar surface mockup using a robotic arm with millimeter-accurate positioning [12]. NASA's POLAR dataset [29] provides $\approx 2,500$ HDR stereo pairs captured over a regolith simulant scene

under low-angle illumination mimicking polar lighting conditions. The Synthetic Lunar Terrain (SLT) dataset [17] introduces neuro-morphic event camera data alongside conventional imagery, expanding the sensor modalities available for lunar perception research. While laboratory datasets offer precise geometric ground truth and repeatable conditions, they suffer from limited scale and diversity, with fixed or limited lighting conditions that do not represent the diversity of real lunar landscapes.

Summary and Positioning. Table 1 summarizes the existing datasets. Synthetic datasets provide scalable labeled data but lack physical realism. Real datasets offer an authentic appearance but no geometric supervision. Laboratory datasets deliver precise ground truth but limited diversity. None provides comprehensive geometric and photometric supervision under illumination variation. MoonAnything addresses this gap by using the real lunar Digital Elevation Models (DEMs). Our proposed dataset provides two different settings: i) LunarGeo provides stereo imagery with dense depth maps rendered using physically-based ray tracing over real lunar topography, spanning diverse trajectories over two geographic regions; and ii) LunarPhoto provides images rendered with SVBRDF parameters learned from real satellite imagery, along with multi-illumination renderings for material appearance modeling research.

Together, MoonAnything offers geometric realism grounded in actual lunar terrain, dense supervision for both geometry and illumination, and diverse lighting conditions. This comprehensive coverage enables training and evaluation of learning-based approaches for vision-based navigation, including terrain-relative navigation, hazard detection and avoidance, and 3D surface reconstruction.

3 Data Generation

3.1 Common Pipeline Tools

Both LunarGeo and LunarPhoto share a common foundation of terrain data, rendering infrastructure, and reflectance modeling, as shown in Figure 1. We first describe these shared components before detailing the specific generation procedures for each sub-dataset.

Geographic Coverage. MoonAnything spans two lunar regions: the *South Pole* and the *Tycho crater*. The South Pole is the target for Artemis III and presents the most challenging lighting conditions for lunar perception, with the Sun never rising more than 7° above the horizon. The Tycho crater, a well-studied mid-latitude impact site, offers dense multi-angle LRO imagery required for BRDF estimation. This geographic split mirrors real mission constraints, where perception methods must generalize across regions.

Terrain Data Sources. Both sub-datasets derive geometry from high-resolution Digital Elevation Models (DEMs):

- *South Pole:* A DEM at 5 m/px resolution derived from LOLA laser altimeter data captured by NASA’s LRO mission [1]. This region features polar craters with elevations ranging from $-4,350$ m to $1,850$ m relative to the lunar mean radius, including permanently shadowed basins and illuminated ridges.
- *Tycho Crater:* A DEM at 1 m/px resolution covering the Tycho crater region, produced using Airbus Pixel Factory from LRO stereo imagery [22]. This young impact crater exhibits pronounced geomorphological features, including a prominent central peak, terraced inner walls, and extensive ejecta deposits. The

Table 2: Altitude versus effective GSD for LunarGeo.

		South Pole									
Alt. (km)		3.5	6.2	9.5	12.8	16.1	19.4	22.7	26.0	29.2	30.5
GSD (m/px)		5.7	10.0	15.4	20.7	26.0	31.4	36.7	42.1	47.2	49.3
		Tycho Crater									
Alt. (km)		3.5	4.1	4.7	5.3	5.9	6.5	7.1	7.7	8.3	8.9
GSD (m/px)		3.7	4.3	4.9	5.5	6.2	6.8	7.4	8.1	8.7	9.3

covered area spans approximately $95 \text{ km} \times 90 \text{ km}$, with elevations ranging from $-3,570$ m to $1,856$ m.

Rendering Framework. As illustrated in Figure 1, both sub-datasets use a physically-based rendering pipeline built on the SurRender [13]. This framework integrates terrain geometry, surface reflectance, solar illumination, and camera models to produce imagery with full geometric supervision and calibrated metadata.

Base Reflectance Model. Surface reflectance is modeled using the Hapke BRDF [7], a physically-grounded formulation for airless bodies that captures opposition surge and anisotropic scattering. For the Tycho crater, we extend the standard constant-parameter Hapke model with a more elaborate SVBRDF model [6] for improved photorealism. LunarGeo uses constant Hapke parameters for the South Pole and learned SVBRDF for Tycho; LunarPhoto uses the SVBRDF both for multi-lighting rendering and as ground truth for reflectance estimation.

3.2 LunarGeo for Geometric Perception

LunarGeo provides stereo image pairs with dense depth supervision for 3D reconstruction research and training. It is generated using real Lunar data that covers both the South Pole and Tycho crater regions, using the common rendering pipeline with region-specific camera, illumination, and reflectance configurations. For the South Pole, we render using the Hapke BRDF with constant albedo, providing consistent photometric behavior across the region. For Tycho crater, we employ both the base Hapke model and the SVBRDF by using LunarG2R framework [6].

Camera Configuration. Each stereo pair is rendered at 512×512 px, with optical blur simulated via Gaussian point-spread function sampling. We use 45° FoV for South Pole and 30° FoV for Tycho (narrower to mitigate DEM artifacts). We simulate three trajectory types inspired by lunar descent phases, informed by analysis of Chang’E-3 mission data and other descent studies [3, 28, 30]:

- **Nadir:** Vertically downward-looking cameras simulating controlled descent. Stereo baselines range from 2% to 10% of altitude, with both cameras at equal height.
- **Oblique:** Tilted cameras with viewing angles between 20° and 35° , simulating lateral motion or target-centered reorientation. Camera altitudes may be equal or offset.
- **Dynamic:** Challenging configurations with altitude variation up to $\pm 15\%$, roll angles up to $\pm 10^\circ$, and stereo baselines ranging from 2% to 22% of the camera altitude.

Stereo pairs span region-specific altitude bands, yielding a range of ground sampling distances (GSDs), as summarized in Table 2. For the South Pole region, camera altitudes are sampled from 3.5 km to 30.5 km, covering a broad set of observation scales. For the Tycho

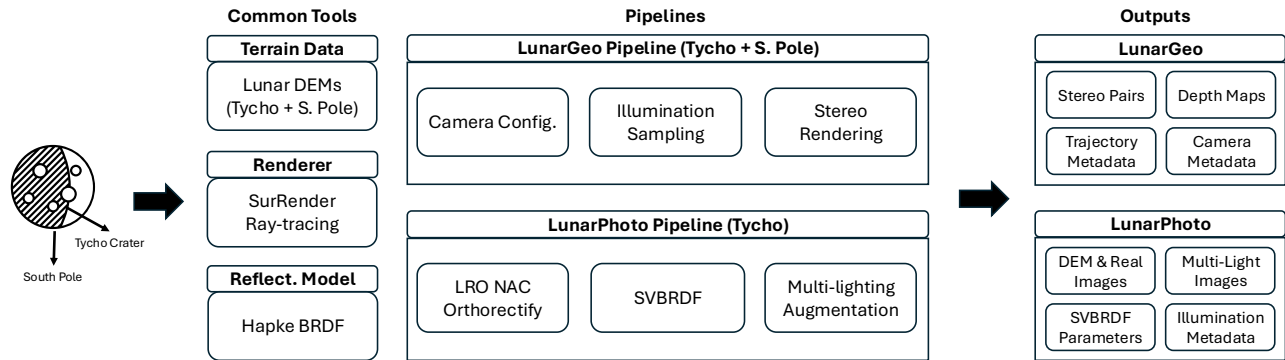


Figure 1: Overview of the MoonAnything dataset generation pipeline.

crater region, we instead consider 10 discrete altitude bands ranging from 3.5 km to 9.5 km. This reduced altitude range limits the spatial footprint of each rendered image, preventing excessively large ground coverage and thereby mitigating the influence of DEM artefacts and interpolation errors on the rendered stereo pairs. Camera positions are uniformly sampled across each region’s spatial extent to ensure terrain diversity.

Illumination Variation. To enable the study of illumination effects on reconstruction quality, each scene is rendered under three distinct realistic solar configurations to produce different shadow patterns. For the South Pole, we simulate the low-angle conditions typical of polar regions. For Tycho crater, we sample a broader range of illumination angles, reflecting its mid-latitude location.

Ground Truth and Annotations. Each stereo pair includes comprehensive geometric and calibration annotation: i) *Camera intrinsics*: Focal length, principal point, and sensor dimensions; ii) *Camera extrinsics*: 6-DOF poses in Moon-fixed reference frame; iii) *Dense depth maps*: Per-pixel depth along camera rays; iv) *Stereo geometry*: Inter-camera baseline translation and rotation; and v) *Trajectory metadata*: Altitude, GSD, and geographic coordinates.

Geometric Perception Tasks. The stereo pairs with dense depth supervision support stereo matching, multi-view 3D reconstruction, and camera pose estimation. The diverse viewing configurations (nadir, oblique, dynamic) and altitude ranges make LunarGeo suitable for evaluating algorithms under realistic descent conditions. These capabilities directly support mission-critical applications, such as *hazard detection and avoidance*, where accurate depth estimation enables the identification of rocks, craters, and unsafe slopes, and *terrain-relative navigation*.

3.3 LunarPhoto for Photometric Perception

LunarPhoto provides paired geometry and reflectance data for appearance modeling research. It extends the original LunarG2R dataset [6], which provided geometry and reflectance pairs but only single-illumination observations per sample. To the best of our knowledge, this is the first dataset providing spatially-varying BRDF parameters with diverse multi-illumination supervision. The sub-dataset covers the Tycho crater region, combining real LRO observations with physically-grounded multi-lighting augmentation.

Real Observation Extraction. Each sample consists of a DEM crop of 128×128 px at 5 m/px (0.4 km^2) from the Tycho DEM (the same terrain source and resolution used in LunarGeo). For each crop, a corresponding LRO NAC [22] image (0.5 m/px to 2 m/px) is orthorectified onto the DEM and cropped to exactly match the spatial extent, producing a ground-truth appearance image aligned with the topography. Metadata, including camera pose, Sun illumination direction, and geographic footprint, is given for each pair.

To construct valid training samples, we compute validity masks for both DEM and orthoimage sources, then erode the merged mask using a kernel equal to half the crop footprint to ensure complete spatial coverage. This defines the set of *valid positions*, *i.e.*, pixels that can safely serve as centers of crops whose full spatial extent contains only valid DEM and appearance data. Crop centers are randomly sampled from valid positions within a validity mask, with the number of crops per orthoimage proportional to its valid surface area. Since orthoimages exhibit non-negligible spatial overlap, we perform geographic splitting, where the global DEM is divided into tiles, and all pairs whose centers fall within the same tile are assigned to the same train/val/test split, to avoid data leakage.

SVBRDF. While the Hapke model provides physically-grounded reflectance, real lunar surfaces exhibit spatially-varying photometric properties not captured by constant-parameter assumptions. We employ a learned SVBRDF model that captures opposition surge and backscattering effects, thus improving photorealism [6].

Multi-Lighting Augmentation. Real LRO observations capture each terrain location under only limited illumination conditions. To support tasks requiring diverse lighting (*e.g.*, photometric stereo, intrinsic decomposition, and relighting), we augment the dataset with appearance images rendered under varied Sun directions.

We simulate Sun positions throughout a lunar day using the SPICE toolkit¹. For each DEM patch, we sample 9 distinct solar hours spanning the local daytime period, converting UTC timestamps into Sun position vectors in the Moon-fixed reference frame. This guarantees that sampled illuminations correspond to physically plausible orbital configurations. Combined with the learned BRDF, we render appearance images under these solar configurations while preserving realistic shadow boundaries, shading gradients, and photometric consistency.

¹<https://github.com/AndrewAnnex/SpicePy>

DEM Normalization. Although lunar relief spans elevations from -5.6 km to 7.5 km globally [16], each DEM crop covers only a small region where height range varies significantly with local morphology. To ensure consistent input representations, we normalize each DEM crop by subtracting its mean elevation and scaling by the dataset-wide standard deviation. This makes models agnostic to absolute altitude while preserving relative terrain structure.

Ground Truth and Annotations. Each geometry–appearance sample includes: i) *DEM patch*: Normalized 128×128 elevation grid at 5 m/px; ii) *Real appearance image*: Orthorectified LRO observation aligned with the DEM; iii) *Depth map*. Per-pixel depth expressed in the real image camera view and aligned with the appearance image; iv) *Surface normal map*: Per-pixel surface normals expressed in the real image camera frame; v) *Multi-lighting images*: Nine rendered appearances under SPICE-sampled solar positions; vi) *Illumination metadata*: Sun azimuth, elevation, and UTC timestamp for each image; vii) *BRDF parameters*: Learned reflectance coefficients for the local terrain; viii) *Geographic coordinates*: Patch center location in selenographic coordinates;

Photometric Perception Tasks. The geometry vs. appearance pairs of LunarPhoto enable training and evaluation of monocular algorithms for depth estimation, crater detection, and semantic segmentation. The multi-lighting renderings support normal estimation via photometric stereo and serve as principled data augmentation for training illumination-robust models. Additionally, the pixel-wise correspondence between real LRO observations and local DEM geometry provides supervision for reflectance estimation, enabling evaluation and improvement of BRDF prediction methods.

4 MoonAnything Dataset

Data Description. We analyze the characteristics of the sub-datasets in the following.

LunarGeo is a synthetic lunar stereo dataset covering two regions: the South Pole and the Tycho crater, with three acquisition trajectories (oblique, dynamic, nadir). Stereo pairs for the South Pole use the classical Hapke BRDF. For Tycho, pairs are rendered with both Hapke and a spatially-varying BRDF (SVBRDF), enabling more realistic, crater-specific reflectance. Figure 2 shows example stereo pairs and 3D scenes for the different trajectories and models. Each pair includes a dense depth map (EXR) and camera parameters (intrinsic and extrinsic in NPZ), allowing full 3D pose recovery. Table 4 summarizes dataset size and splits: about 38K pairs for the south pole and 20K for Tycho, split 80%/10%/10% for train/val/test.

LunarPhoto is composed of 84K samples, split into 67K training, 8.5K validation, and 8.5K test samples. Each sample represents a local lunar surface patch and combines geometric information, real orbital imagery, and physically-based renderings under varying illumination conditions. For every sample, nine distinct Sun configurations are provided, resulting in a total of 750K rendered images per BRDF model. Table 5 summarizes the overall dataset statistics and the content provided for each sample.

Qualitative Examples. Examples are shown in Figure 3. Each data sample includes a DEM crop (b) and an orthorectified real lunar image acquired by the LRO NAC (a). Depth (c) and surface normal maps (d) are provided in the viewpoint of the real image camera, obtained by rendering the image onto the DEM using the

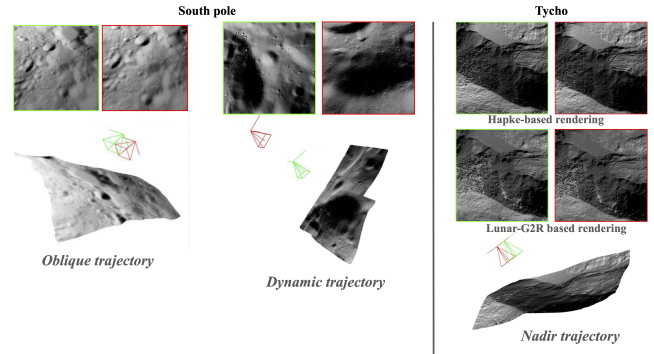


Figure 2: Examples of LunarGeo illustrating the three trajectory types. The first row shows stereo pairs and their associated 3D scene over the lunar South Pole, rendered using the Hapke BRDF with constant parameters. The second row presents two stereo pairs from the Tycho crater: the first rendered using the classical Hapke BRDF, and the second using SVBRDF from [6].

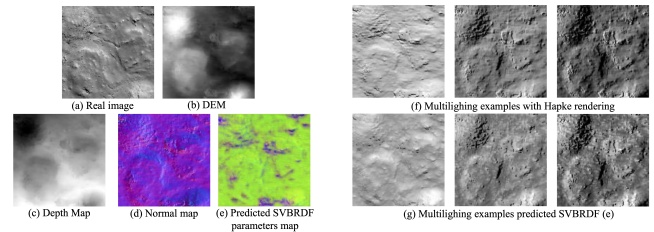


Figure 3: Samples from LunarPhoto.

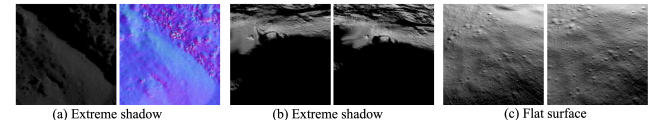


Figure 4: Examples of challenging cases: (a) extreme shadowing observed in the multi-lighting data from LunarPhoto, (b) a LunarGeo pair from the Tycho under raking lighting, and (c) a flat stereo pair from the S. Pole captured by LunarGeo.

corresponding camera parameters. In addition, per-pixel SVBRDF parameters (e) are supplied together with multi-illumination renderings generated using both the classical Hapke reflectance model (f) and the SVBRDF (g), all stored in TIFF format. Metadata files include camera parameters, Sun directions, and geolocation information, enabling reproducible rendering and downstream tasks.

Challenging Scenarios. MoonAnything deliberately includes a wide range of challenging visual conditions to promote robustness in downstream vision applications. In particular, extreme illumination cases with large shadowed regions are common in low-Sun-angle configurations, as illustrated in 4 (a) and (b), where strong contrast and partial observability significantly complicate both feature extraction and geometric reasoning. In addition to illumination variability, the dataset covers diverse terrain types, ranging from

Table 3: Stereo 3D reconstruction results on LunarGeo. Best results are highlighted in green and second-best in orange.

Dataset	Method	Nadir			Oblique			Dynamic		
		ACC.(m)↓	Compl.(m)↓	Chamfer(m)↓	ACC.(m)↓	Compl.(m)↓	Chamfer(m)↓	ACC.(m)↓	Compl.(m)↓	Chamfer(m)↓
Tycho (Unseen)	MASt3R	60	58	59	128	88	108	49	45	47
	VGGT	39	39	40	95	78	87	75	51	63
	VGGT FT	35	35	35	76	60	68	47	52	50
	MASt3R FT	37	37	37	72	57	64	34	32	33
S. Pole (Seen)	MASt3R	236	235	236	385	259	322	289	270	279
	VGGT	225	228	226	222	210	216	323	353	300
	VGGT FT	43	43	43	64	63	64	69	69	70
	MASt3R FT	103	97	100	141	147	144	109	114	111

Table 4: LunarGeo dataset statistics and train/val/test splits.

Region	Total	Train (80 %)	Val (10 %)	Test (10 %)
South Pole	38K	30K	4K	4K
Tycho	20K	16K	2K	2K

Table 5: LunarPhoto statistics and per-sample content.

Item	Description
Total samples	84K samples
Split	67K train / 8.5K val / 8.5K test
Sun positions (multi-lighting)	9 per sample
Rendered images	750K per BRDF model
Geometry	dem.tif, depth.tif, normal.tif
Real data	real_image.tif (orthorectified LRO NAC image)
Metadata	metadata.json (camera, Sun, geolocation)
BRDF	brdf_map.tif (per-pixel SV-BRDF parameters)
Renderings	Multi-illumination renderings using Hapke and learned BRDF models (TIFF)

highly structured areas such as crater rims, slopes, and ejecta fields to flatter surfaces with limited texture and fewer distinctive features, as shown in (c). This combination of lighting and terrain diversity is essential to evaluate and train models capable of operating reliably in realistic lunar exploration scenarios.

5 Stereo 3D Reconstruction as a Showcase

We evaluate the practical relevance of MoonAnything using stereo 3D reconstruction as a downstream task. MASt3R [14] and VGGT [25] are fine-tuned on the South Pole subset of LunarGeo and evaluated on stereo sequences acquired under three trajectory configurations: nadir, oblique, and dynamic. We report 3D reconstruction metrics using completeness, accuracy, and Chamfer distance for geometry. In addition to in-domain evaluation on the South Pole, performance is also assessed on the Tycho crater to evaluate cross-region generalization. Quantitative results are reported in Table 3.

Fine-tuning on a limited but physically consistent portion of MoonAnything leads to substantial improvements over pretrained models across all trajectory configurations. On the *seen* South Pole split, both models significantly reduce reconstruction errors, with VGGT FT (where FT denotes fine-tuning) performing best for nadir views, while MASt3R FT shows increased robustness under oblique and dynamic trajectories.

Importantly, similar gains are observed on the *unseen* Tycho split, despite its markedly different terrain morphology. MASt3R FT

achieves the lowest errors for oblique and dynamic trajectories, demonstrating strong generalization to out-of-distribution lunar landscapes. Overall, these results show that fine-tuning on a restricted subset of MoonAnything is sufficient to learn transferable, geometry-aware representations, validating both the realism of the proposed rendering pipeline and the suitability of MoonAnything for downstream planetary 3D vision tasks.

6 Conclusion

We presented MoonAnything, a unified benchmark for lunar surface perception that addresses the lack of comprehensive datasets combining geometric and photometric supervision. By unifying and extending our prior works [5, 6], MoonAnything provides stereo imagery of **large** dataset size with dense depth maps across two lunar regions (LunarGeo) and spatially-varying BRDF parameters with multi-illumination renderings (LunarPhoto). Together, these sub-datasets offer the first benchmark enabling research on 3D reconstruction, reflectance estimation, and illumination-robust perception within a consistent lunar context. The Moon specific setting offered a calibrated (e.g. a unique, known light source, the sun) and large dataset/ground truth that calls for generalization in more complex, general setting.

Acknowledgments

This work was carried out with the support of the European Space Agency (ESA) under contract No: 4000140461/23/NL/GLC/my.

References

- [1] M.K. Barker, E. Mazarico, G.A. Neumann, M.T. Zuber, J. Haruyama, and D.E. Smith. 2016. A new lunar digital elevation model from the Lunar Orbiter Laser Altimeter and SELENE Terrain Camera. *Icarus* 273 (July 2016), 346–355. doi:10.1016/j.icarus.2015.07.039
- [2] Valentin Tertius Bickel, Ben Moseley, I Lopez-Francos, and M Shirley. 2021. Peering into lunar permanently shadowed regions with deep learning. *Nature communications* 12, 1 (2021), 5607.
- [3] Joel Getchius, Devin Renshaw, Daniel Posada, Troy Henderson, Lillian Hong, Shen Ge, and Giovanni Molina. 2024. Hazard Detection and Avoidance for the Nova-C Lander. In *AASG&C*. doi:10.1007/978-3-031-51928-4_53
- [4] JN Goswami and M Annadurai. 2009. Chandrayaan-1: India’s first planetary science mission to the Moon. *Current science* (2009), 486–491.
- [5] Clémentine Grethen, Simone Gasparini, Géraldine Morin, Jeremy Lebreton, Lucas Marti, and Manuel Sanchez-Gestido. 2025. Adapting Stereo Vision From Objects To 3D Lunar Surface Reconstruction with the StereoLunar Dataset. In *ICCVW*.
- [6] Clémentine Grethen, Nicolas Menga, Roland Brochard, Géraldine Morin, Simone Gasparini, Jeremy Lebreton, and Manuel Sanchez Gestido. 2026. Lunar-G2R: Geometry-to-Reflectance Learning for High-Fidelity Lunar BRDF Estimation. *arXiv:2601.10449* (2026).
- [7] Bruce Hapke. 1993. *Theory of reflectance and emittance spectroscopy*. CUP.

- [8] Steffen Herbot and Christian Wöhler. 2011. An introduction to image-based 3D surface reconstruction and a survey of photometric stereo methods. *3D Research* (2011).
- [9] Indian Space Research Organisation (ISRO). [n. d.]. Chandrayaan-II. <https://pradan.issdc.gov.in/ch2/>.
- [10] Andrew E Johnson, Andres Huertas, Robert A Werner, and James F Montgomery. 2008. Analysis of on-board hazard detection and avoidance for safe lunar landing. In *AeroConf*.
- [11] Andrew E Johnson and James F Montgomery. 2008. Overview of terrain relative navigation approaches for precise lunar landing. In *AeroConf*. IEEE, 1–10. doi:10.1109/AERO.2008.4526302
- [12] Jérémy Lebreton, Ingo Ahrens, Roland Brochard, Christoph Haskamp, Hans Krüger, Matthieu Le Goff, Nicolas Menga, Nicolas Ollagnier, Ralf Regele, Francesco Capolupo, and Massimo Casasco. 2024. Training Datasets Generation for Machine Learning: Application to Vision Based Navigation. arXiv:2409.11383 [cs.CV]
- [13] Jérémy Lebreton, Roland Brochard, Nicolas Ollagnier, Matthieu Baudry, Adrien Hadj Salah, Grégory Jonniaux, Keyvan Kanani, Matthieu Le Goff, and Aurore Masson. 2024. High performance Lunar landing simulations. arXiv:2409.11450 (2024).
- [14] Vincent Leroy, Yohann Cabon, and Jérôme Revaud. 2024. Grounding Image Matching in 3D with MAST3R. In *ECCV*. doi:10.1007/978-3-031-73220-1_5
- [15] Jiayi Liu, Qianyu Zhang, Xue Wan, Shengyang Zhang, Yaolin Tian, Haodong Han, Yutao Zhao, Baichuan Liu, Zeyuan Zhao, and Xubo Luo. 2024. LuSNAR: A Lunar Segmentation, Navigation and Reconstruction Dataset based on Multi-sensor for Autonomous Exploration. arXiv 2407.06512 (2024). doi:10.48550/ARXIV.2407.06512
- [16] LRO Project. 2008. *A Standardized Lunar Coordinate System for the Lunar Reconnaissance Orbiter*. Technical Report. NASA.
- [17] Marcus Mörtens, Kevin Farries, John Culton, and Tat-Jun Chin. 2024. Synthetic Lunar Terrain: A Multimodal Open Dataset for Training and Evaluating Neuro-morphic Vision Algorithms. arXiv 2408.16971 (2024). doi:10.48550/ARXIV.2408.16971
- [18] Paul B Niles and Ryan Stephan. 2025. The Commercial Lunar Payload Services Initiative. In *AeroConf*.
- [19] SM Parkes, Iain Martin, Martin Dunstan, and D Matthews. 2004. Planet surface simulation with pangu. In *SpaceOps*.
- [20] Romain Pessia, Genya Ishigami, and Quentin Jodelet. 2019. Artificial Lunar Landscape Dataset. <https://www.kaggle.com/datasets/romainpessia/artificial-lunar-rocky-landscape-dataset>.
- [21] Daniel Posada and Troy Henderson. 2024. Dense Feature Matching for Hazard Detection and Avoidance Using Machine Learning in Complex Unstructured Scenarios. *Aerospace* 11, 5 (2024), 351.
- [22] Mark Robinson. 2011. LRO Moon LROC 5 RDR v1.0. NASA Lunar Reconnaissance Orbiter Camera (LROC) data product.
- [23] Hiroyuki Sato, Mark S. Robinson, Bruce Hapke, Brett W. Denevi, and Angela K. Boyd. 2014. Resolved Hapke parameter maps of the Moon. *J. Geophys. Res.* 119, 8 (2014).
- [24] Richard Vondrak, John Keller, Gordon Chin, and James Garvin. 2010. Lunar Reconnaissance Orbiter (LRO): Observations for lunar exploration and science. *Space science reviews* 150, 1 (2010), 7–22. doi:10.1007/s11214-010-9631-5
- [25] Jianyuan Wang, Minghao Chen, Nikita Karaev, Andrea Vedaldi, Christian Rupprecht, and David Novotný. 2025. VGGT: Visual Geometry Grounded Transformer. In *CVPR*. doi:10.1109/CVPR52734.2025.00499
- [26] Shuzhe Wang, Vincent Leroy, Yohann Cabon, Boris Chidlovskii, and Jérôme Revaud. 2024. DUST3R: Geometric 3D Vision Made Easy. In *CVPR*. doi:10.1109/CVPR52733.2024.01956
- [27] Yanbo Wang, Ting Yuan, Chuankai Liu, Qi Wu, and Jiuchao Qian. 2024. The Real Chang’e Lunar Landscape Dataset. doi:10.21227/9Y8Q-DX27
- [28] Edward C. Wong, Gurkirpal Singh, and James P. Masciarelli. 2006. Guidance and Control Design for Hazard Avoidance and Safe Landing on Mars. *J. Spacecr. Rockets* (2006). doi:10.2514/1.19220
- [29] Uland Wong, Ara Nefian, Larry Edwards, Xavier Buoysounouse, P Michael Furlong, Matt Deans, and Terry Fong. 2017. Polar optical lunar analog reconstruction (polar) stereo dataset. *NASA Ames Research Center* (2017).
- [30] Ping Yu, Yu Zhao, Ji Li, XiaoWen Zhang, DaYi Wang, XiangYu Huang, Jun Liang, YiFeng Guan, HongHua Zhang, and Wei Yang. 2014. Guidance navigation and control for Chang’E-3 powered descent. *Sci. Sin. Technol.* 44 (2014). doi:10.1360/092014-43