MMoEDTM: Improving the Generalization of Multi-terrain Lunar DTM and Optimization of Rocky Surface Generation

Yi Zheng, Tong Xia, Niutao Liu*, *Member, IEEE*, and Ya-Qiu Jin, *Life Fellow, IEEE*

***Abstract*—In recent years, deep learning has been widely applied to the reconstruction of high-resolution lunar digital terrain models (DTMs) from single-view orbital optical images. However, existing deep learning-based methods have limited generalization across different terrains. Their results are often unstable in highly rugged areas and in regions with dense rocks. The mixture-of-experts (MoE) architecture improves generalization through collaboration among multiple expert networks, which has been proven effective in large language models. In contrast, conventional dense small models exhibit limitations in generalization due to constraints in parameter scale and training data. To address this problem, an end-to-end lightweight MoE framework with a Markov random field (MRF) constraint, termed MMoEDTM, is proposed. This framework improves the generalization ability of small models and supports high-precision DTM reconstruction across different lunar terrains. This framework uses MRF to classify terrains and distribute samples to different expert networks. It realizes decoupled modeling and separate representation for terrain features, and sustains model robustness in undulating regions. This work develops a Coarse Adapter module. It embeds coarse digital elevation models (DEM) as terrain prior into reconstruction, and narrows the representation gap between deep features and original elevation information. Current methods face limitations in the reconstruction of rough rocky surfaces. This paper adopts a slope constraint loss function. With shadow detection, the rock shadow error is evaluated to be around 10%. This method improves generalization in lunar environments and restores surface details. A quantitative analysis of rock roughness is performed using the generated DTM of the low-latitude Galilei-E crater region. The results are systematically compared with SAR backscatter intensity and show high consistency with SAR observations. In addition, reconstruction experiments are performed in the permanently shadowed region of Shackleton crater.**

***Index Terms*—Super Resolution, Digital Terrain Model, Rock, Mixture of Experts, Permanently Shadowed Region**

I. INTRODUCTION

IN lunar exploration missions, DEMs are critical. High-precision and high-resolution lunar DEMs support key applications, including lunar surface feature identification, landing-site selection, and rover navigation and positioning [1], [2].

There are two methods for constructing high-resolution lunar DEM, including Stereo-photogrammetry (SPG) and laser altimetry-based interpolation [3]. SPG is mature and reliable for flat lunar surface regions, but has not achieved full global coverage of the Moon. Laser altimetry-based interpolation is independent of solar illumination and is not affected by shadowed regions. Lunar Orbiter Laser Altimeter (LOLA) data provide a resolution of 30 m/pixel at low latitudes and 5 m/pixel in polar regions. However, elevation points on the lunar surface are unevenly distributed and have large sampling intervals. As a result, most lunar regions still lack meter-scale DEMs [4].

In recent years, research has increasingly focused on extracting lunar elevation information from single images and constructing DEMs [5] at low latitude. Traditional single-image DEM reconstruction methods are mainly based on surface reflectance modeling. Typical approaches include Shape from Shading (SFS) and Shape and Albedo from Shading (SAFS) [6]. These methods utilize the intensity variation characteristics of the lunar surface to estimate parameters such as the relative height or surface normal vector at each point. However, the roughness and reflectance parameters of the lunar surface are unknown and variable, introducing uncertainties into terrain generation. Moreover, a single generation function cannot adapt to diverse lunar landforms, further increasing uncertainties, especially in rocky and rough regions [7], [8], [9].

Deep learning has been applied to high-resolution reconstruction of planetary DTM from single-view orbital images [8], [10], [11], [12], [13]. It has become an important complementary approach to SPG and SFS. Model training usually uses existing high-resolution DTMs generated by SPG or orthorectified images. The model then learns the relationship between optical images and DTMs [14]. However, existing methods still face several major challenges. Generalized reconstruction across multiple lunar terrain types are difficult. A model must preserve smooth continuity in flat regions while also restoring detailed undulations in rough terrains. Current

Manuscript received …… This work was supported in part by the National Natural Science Foundation of China under Grant 62495032 and Grant 62201154, and in part the Chenguang Program of Shanghai Education Development Foundation and Shanghai Municipal Education Commission.

Yi Zheng, Tong Xia, Niutao Liu and Ya-Qiu Jin are with the Key Laboratory of Information Science of Electromagnetic Waves (MoE), Fudan University, Shanghai 200433, China (corresponding e-mail to Niutao Liu: ntliu@fudan.edu.cn).

Digital Object Identifier…..

DTM datasets used for model training are mostly generated by methods such as SFS. These datasets are not suitable for rocky areas. If the model fully learns SFS-related features, its outputs may deviate from the real height distribution of lunar rocks. In addition, stripe artifacts in high-resolution DTMs can interfere with model training.

The MoE [15] architecture relies on multi-expert division of labor and dynamic gating routing mechanisms to capture differentiated features, which has substantially boosted the cross-scene generalization capability of large language models. In contrast, small-scale models constrained by limited training data and parameter budgets. Directly introducing the MoE paradigm into such small models tends to trigger parameter redundancy, feature redundancy and degraded inference efficiency, thereby making the dense architecture the mainstream alternative. Characterized by full global parameter sharing, dense structures fail to decouple lunar topographic features with distinct complexities and adapt to feature variations in rugged landscapes and abrupt terrain regions. Consequently, these models exhibit limited generalization capacity, leading to severe degradation in the continuity and stability of terrain reconstruction results.

To address the aforementioned challenges, this study proposes Markov Random Fields (MRF) [16] based Mixture of Experts DTM (MMoEDTM), a deep learning model based on the MoE architecture. This method takes a single high-resolution optical image and a low-resolution DEM as inputs. It uses MRF to classify DTM data into flat, gently undulating, and crater terrains, and adopts targeted training for each category to improve generalization, achieving state-of-the-art (SOTA) performance across multiple evaluation metrics. To address the mismatch between existing data and the real distribution and morphology of rocks in the generation of the rocky region. A portion of large-scale flat-terrain data with detailed undulation features is added into the training dataset. The ability to characterize flat terrain is transferred to crater regions. This makes the generated rocks closer to real topographic features. The effects of these parameter settings are also evaluated experimentally. The proposed method is applied to the rock-rich Galilei-E crater. Roughness indices are extracted from the generated DTM to describe the spatial distribution of lunar surface roughness. These results are then verified by comparison with SAR images [17]. Furthermore, to enable the model to adapt to special regions with large-angle illumination, such as Permanently Shadowed Regions (PSRs), large-angle illumination data are incorporated into the training dataset. The generalization is further verified by generating the DTM of Shackleton crater.

The remainder of this paper is organized as follows. Section 2 describes the dataset preparation procedure, the network architecture, and the modules of MMoEDTM. Section 3 validates the generalization performance through experiments on flat terrains and crater walls, as well as on discrete and dense rocks. Section 4 presents the application of the proposed method to the rocky Galilei-E crater at low latitudes and the Shackleton crater in PSRs. Section 5 discusses and analyzes the influence of the solar incidence and azimuth angles on generation performance, and examines how varying rock proportions in the training set affect the generated results. Finally, Section 6 provides the conclusion.

II. Dataset Preparation Procedure

This reconstruction framework uses single-view high-resolution optical images and corresponding low-resolution DEM as inputs. A deep learning network is then used to infer high-resolution DTMs at the pixel level. The Narrow Angle Camera (NAC) onboard the Lunar Reconnaissance Orbiter (LRO) has acquired meter-scale orbital images of the lunar surface and achieved near-global coverage of the lunar surface [18]. The Selene and Lunar Orbiter Laser Altimeter DEM (SLDEM) [19] is a fusion product of LOLA measurements and a photogrammetric model derived from terrain camera images of the Japan Aerospace Exploration Agency's SELenological and ENgineering Explorer (SELENE) mission, covering latitudes between 60°S and 60°N. Its spatial resolution is around 60 m/pixel. SLDEM exhibits higher accuracy and fewer stripe artifacts than the LOLA DEM. Therefore, SLDEM is used as the input low-resolution DEM. Meanwhile, the LRO NAC DTM [20] is selected as the ground truth reference data in this study. The data processing procedure of this paper is as follows:

(1) Data Preparation：

The training dataset contains samples from three typical lunar terrains, including flat, mountain, and crater terrain. Flat terrain mainly includes lunar maria with gentle topography. Mountain terrain exhibits moderate elevation variation and is defined by a slope range intermediate between those of flat terrain and crater terrain. Crater terrain includes highly rugged areas, such as crater walls. The model inputs include Lunar Reconnaissance Orbiter Camera (LROC) NAC images with a spatial resolution of 2-5 m/pixel and SLDEM2015 at approximately 60 m/pixel. The LROC NAC DTM has the same resolution as the NAC images. To ensure spatial alignment, SLDEM is first co-registered with the NAC images and DTM. It is then resampled to the resolution of NAC image using bilinear interpolation. In addition, approximately 10% of the NAC images acquired under large solar incidence angles are included. These samples are used to improve adaptation to extreme illumination conditions in PSRs observed by ShadowCam.

(2) Image Registration and Cropping:

DEM and NAC use different sampling schemes. These differences can cause topographic misalignment during data fusion. Therefore, high-precision registration is required before terrain modeling. A stepwise registration strategy is adopted in this study. Initial registration is conducted using the software called Integrated Software for Imagers and Spectrometers Version 3. Manual pixel-level refinement is then performed in ArcGIS Pro. This process provides a reliable data basis for subsequent high-resolution DEM generation.

To meet the model input requirements, NAC image pairs and the corresponding SLDEM and NAC DTM are cropped into overlapping patches. Each patch is 256 × 320 pixels. A 100-pixel overlap is used in both the horizontal and vertical directions. Samples with severe artifacts in the NAC DTM are removed during data cleaning. In total, 9061 valid pairs are obtained, including 3947 from flat terrain, 2459 from mountain terrain, and 2655 from crater terrain.

(3) Image Normalization:

SLDEM and LROC NAC have different data resolutions. To align them, SLDEM is upsampled to a consistent resolution using bilinear interpolation. The original LROC NAC data values change with exposure conditions, so their dynamic range fluctuates. To ensure consistent statistical properties for network input images, the min–max normalization is used during preprocessing. This method linearly maps the grayscale of optical images to the range [0, 255]. It also maps elevation values from the low-resolution SLDEM and high-resolution NAC DTM to the range [0, 50][10].

$$Z_{\text{norm}} = N \cdot \left(\frac{Z_{origin} - \min\left(Z_{origin}\right)}{\Delta Z_{origin}} \right) \quad (1)$$

Where Z_{norm} denotes the normalized DEM elevation. N denotes the normalization range, which is set to 50 here. Z_{origin} denotes the original DEM value. $\min(Z_{origin})$ denotes the minimum elevation value in Z_{origin}, and ΔZ_{origin} denotes the value of $\max(Z_{origin}) - \min(Z_{origin})$.

(4) Data Augmentation:

To improve the network's generalization, several data augmentation strategies are adopted in this study. These include rotation and flipping of training samples, as well as the use of LROC NAC images with varying solar incidence and azimuth angles. This process generates approximately 40,000 augmented training pairs.

(5) DEM Mosaic Generation:

After multiple high-resolution DTM tiles are generated, tile mosaicking is required. This process includes two steps, including scale recovery and DEM mosaicking. Scale recovery is performed by applying the inverse transformation of DEM normalization. It is expressed as follows:

$$Z_{sr} = Z_{pred} \cdot \frac{\max(Z) - \min(Z)}{N} + \min(Z) \quad (2)$$

Where Z represent DEM data. Z_{sr} and Z_{pred} stand for the scale-restored DEM data and the model-generated DTM tiles, respectively. Adjacent DTM tiles have a 100-pixel overlap. They are fused using a linear gradient-weighted strategy. In the overlapping area, some pixels have valid values in both tiles. These pixels are interpolated linearly based on gradient weights. Other pixels are valid in only one tile. Their values are preserved. By processing each tile one by one, all small DTM tiles are stitched together. The result is a seamless and smooth DTM mosaic. This mosaic covers the entire large-scale orbital image.

The input size of the deep learning network is determined by both the model architecture and the GPU memory. In this study, all experiments are conducted on a single NVIDIA RTX 4090 GPU with 24 GB of video memory. The batch size is set to 4. The learning rate is linearly increased from 0 to 10^{-4} over the first 1000 iterations. It then follows a cosine decay schedule until the end of training. The total number of training iterations is 50,000.

III. MMoEDTM Network

This study develops an MRF-based MoE network, named MMoEDTM, for the high-resolution reconstruction of low-resolution DEM. The overall architecture is presented in Fig. 1. First, the large-scale DTM is divided into separate tiles. Absolute elevation differences are extracted from these tiles as core features to build super-pixel units for local terrain characterization, which are further aggregated to generate a global landform feature map. Afterwards, the feature map is imported into the MRF module. Serving as a feature-guided router, this module realizes fine-grained landform classification into flat, mountain and crater terrains. It also constructs a one-to-one mapping between super-pixel units and landform characteristics. Driven by the above routing mechanism, low-resolution DTM patches and co-registered high-resolution optical images are adaptively distributed to terrain-oriented expert networks customized for distinct landform types. Ultimately, the outputs of multiple expert networks are fused to produce high-resolution DTM results.

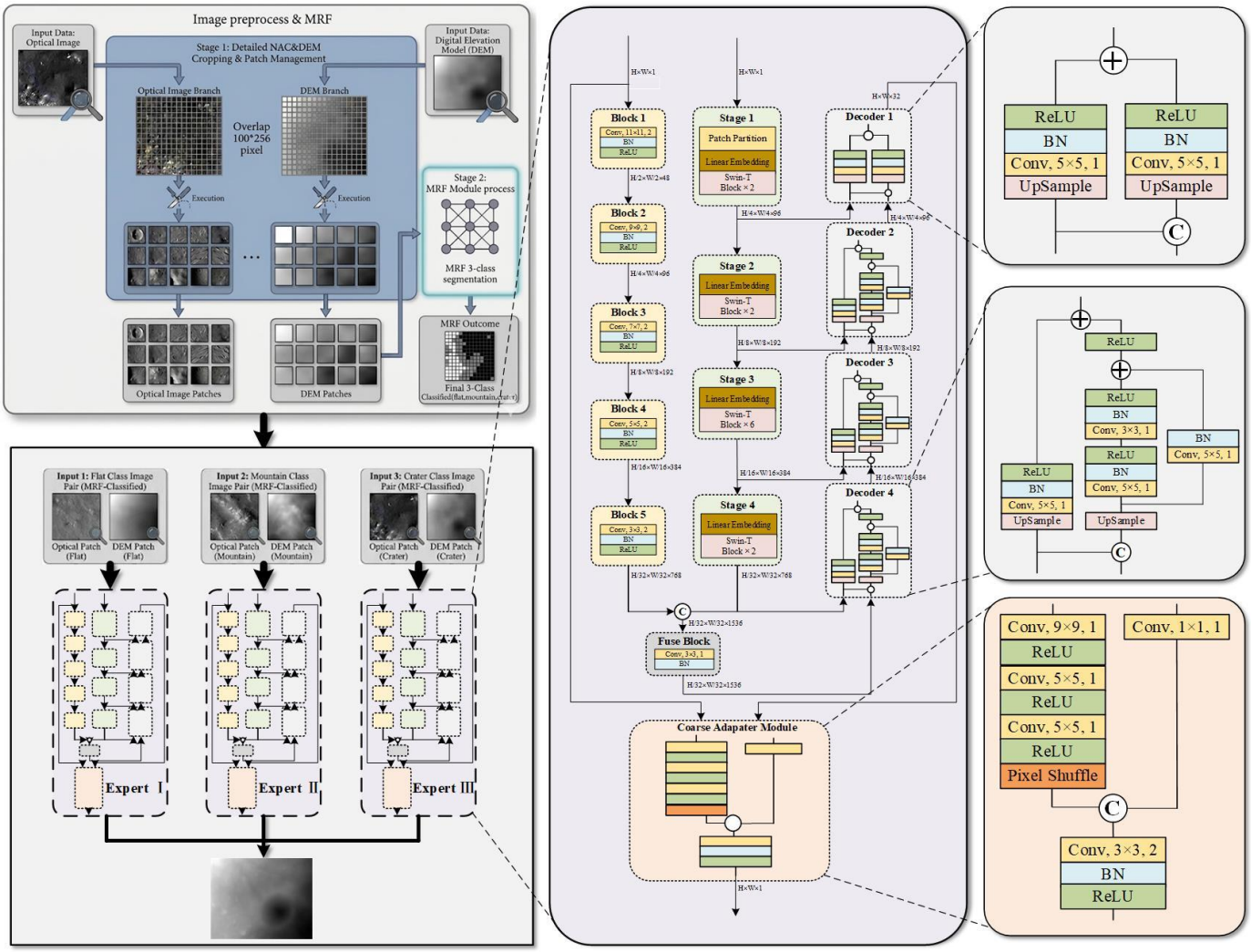
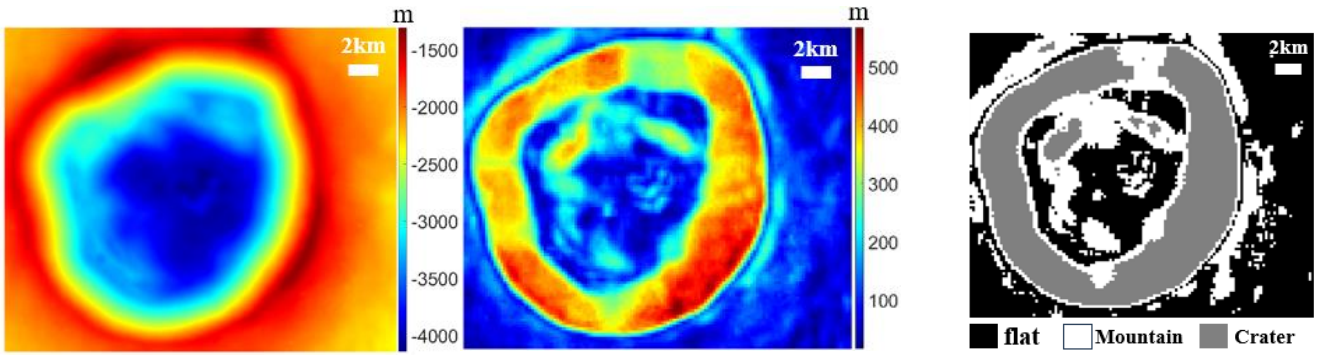


Fig. 1. Architecture of the MMoEDTM network.

A) MRF Classification Router Module

This study uses the MRF method to partition input samples during preprocessing. It divides the data into three subsets, including flat, mountain, and crater terrain. As a result, each expert network can focus on its own specific data subset throughout training and inference. The classified results are then sent to the corresponding expert network. MRF has been widely applied in computer vision tasks including image classification, segmentation, and semantic understanding. Its core theoretical foundation is the inherent spatial correlation among image pixels. This follows the locality assumption, where a pixel's posterior probability of belonging to a specific category mainly depends on the category states of its neighboring pixels. At the same time, it remains independent of non-neighboring pixels at greater distances. This characteristic aligns with the intrinsic regularity of spatial aggregation of pixels belonging to the same target in natural images.

Fig. 2(a) shows the SLDEM of the target region at 31.85°N , 67.59°W. The DEM is divided into 256 × 320-pixel image patches. Adjacent patches feature a 100-pixel horizontal and vertical overlap to enable seamless stitching after reconstruction. These overlapping pixels do not interfere with the classification task and are only used for spatial alignment and boundary smoothing. The difference between the maximum and minimum values within each individual SLDEM image patch is taken as the feature representation of the patch, which is regarded as super-pixel. These super-pixels are reconstructed by mapping each to its corresponding spatial position in the original SLDEM to form a low-resolution feature image with a spatial extent consistent with the target region. Fig. 2(b) shows the reconstructed image. A three-category classification is then performed on this image using the MRF. The result is presented in Fig. 2(c). In this figure, gray super-pixels indicate crater wall terrain, white super-pixels indicate mountain terrain, and black super-pixels indicate flat terrain.



(a) (b) (c)

Fig. 2. MRF processing stages. (a) The DEM image of Lichtenberg crater. (b) The image of super-pixel reconstruction in Lichtenberg crater. (c) The three-category MRF classification map of (b).

To verify the classification results of the MRF algorithm, Fig. 3 presents the probability density functions of flat, mountain, and crater terrains. The horizontal axis denotes the DEM elevation difference, which corresponds to super-pixel size. This value is calculated as the difference between the maximum and minimum DEM values within each 320 × 256 patch, with μ representing its mean value. It is observed that the three terrain types exhibit distinct differences in the probability density distributions. The mean μ of the absolute elevation difference is 85.65 for flat terrains, 203.54 for mountain terrains, and 396.32 for crater terrains. The MRF enables adaptive adjustment regardless of the original DEM resolution. The three distributions are well separated, with the overlap between any two adjacent regions accounting for less than 10% of their respective areas. A low overlap provides a foundation for the effective separation of the three terrain types.

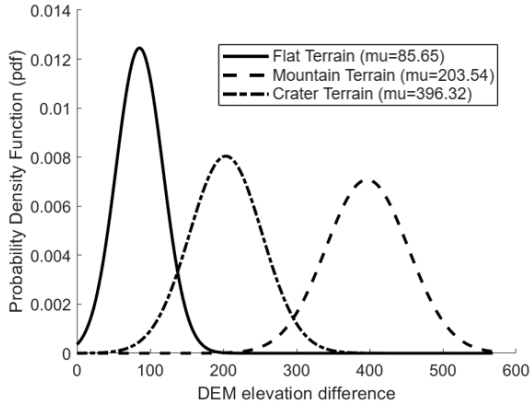


Fig. 3. Probability density function of the three-category MRF classification.

B) Coarse Adapter Module

To mitigate the progressive weakening of structural priors inherent in the coarse DEM during deep encoding, cross-modal fusion, and hierarchical decoding processes, a Coarse Adapter (CA) Module is introduced at the final reconstruction stage. Specifically, the CA Module directly takes the coarse DEM as input. It then learns a shallow adaptive mapping function. This function transforms coarse elevation cues into structure-aware features. These features are compatible with the output of the decoder.

While the main network branch is primarily dedicated to recovering high-frequency terrain details from high-resolution imagery, the CA Module explicitly preserves the low-frequency elevation trends and large-scale topographic consistency inherited from the coarse DEM. Notably, in the absence of the CA Module, the network tends to over-rely on image texture cues, which inevitably leads to local elevation drift, spurious terrain artifacts, and compromised global topographic consistency.

Thus, the CA Module serves as a complementary pathway that balances the enhancement of fine-grained terrain details and the preservation of structural fidelity, ensuring the reliability and accuracy of the reconstructed DEM.

C) Hybrid Loss Function

This dataset is constructed from high-resolution optical imagery acquired by orbiters. It encompasses diverse scenarios with significant morphological and scale differences between regular and irregular terrain features. These variations pose challenges for training deep learning models capable of effective DEM estimation. To address the above-mentioned challenges and achieve accurate vertical predictions, this study designs a hybrid loss function. This loss function guides the model to minimize prediction residuals in both horizontal and vertical dimensions. It further ensures that terrain feature contours are clear and morphologically regular horizontally.

Hybrid loss function are as follows:

$$l\left(z,z_{gt}\right)=\rho l_{norm}\left(z,z_{gt}\right)+\varphi 1_{grad}\left(z,z_{gt}\right)+\omega 1_{slope}\left(z,z_{gt}\right) \quad (3)$$

where $\rho=1,\varphi=1,\omega=5$, z and z_{gt} represent the generated high-resolution DEM map and the ground truth respectively.

$$l_{\text{norm}}=\frac{1}{n}\sum_{p=1}^{n}\left(1-\frac{\langle z_{gt},z\rangle}{\sqrt{\langle z_{gt},z\rangle}\sqrt{\langle z_{gt},z\rangle}}\right) \quad (4)$$

z and z_{gt} in loss function l_{norm} denote the height values in the generated DEM image and the ground truth respectively. $\langle z_{gt},z\rangle$ denotes the inner product of vector of z_{gt} and z . By quantifying the angle between the surface normal vectors of the generated DEM and the ground truth, this loss places more emphasis on the consistency of slope directions compared with l_{slope} .

$$l_{\text{grad}}=\frac{1}{n}\sum_{p=1}^{n}\left(\left|g_{gt\_x}-g_x\right|+\left|g_{gt\_y}-g_y\right|\right) \quad (5)$$

The mean of the absolute gradient differences along the x-axis and y-axis in loss function l_{grad} , where g_{gt_x} and g_{gt_y}

represent the horizontal and vertical gradients of the ground truth respectively. g_x and g_y denote the horizontal and vertical gradients of the predicted DEM respectively.

$$l_{\text{slope}}(z, z_{gt}) = \frac{1}{n}\sum_{i,j}\left|\arctan\left(\sqrt{\left(\frac{\Delta_x z}{r}\right)^2+\left(\frac{\Delta_y z}{r}\right)^2+10^{-6}}\right)-\arctan\left(\sqrt{\left(\frac{\Delta_x z_{gt}}{r}\right)^2+\left(\frac{\Delta_y z_{gt}}{r}\right)^2+10^{-6}}\right)\right| \quad (6)$$

The angle between the surface normals of the generated DTM and the ground truth DTM in loss function l_{slope}, which is used to measure the slope difference between the generated image z and the ground truth image z_{gt} .This loss function is derived from the physically based slope map calculation formula and exhibits high sensitivity to subtle surface undulations by comparing the predicted slope map with the ground truth. Compared with l_{norm}, it places greater emphasis on the consistency of slope magnitudes. $\Delta_x z = z_{i:j+1} - z_{i,j}$ denotes the horizontal gradient operator of the generated DTM image. $\Delta_y z = z_{i+1,j} - z_{i,j}$ denotes the vertical gradient operator of the generated DTM image. $\Delta_x z_{gt} = z_{gt(i:j+1)} - z_{gt(i,j)}$ denotes the horizontal gradient operator of the ground truth. And $\Delta_y z_{gt} = z_{gt(i+1,j)} - z_{gt(i,j)}$ denotes the vertical gradient operator of the ground truth image. r represents the resolution of the ground truth. 10^{-6} is a numerical stability regularization term used to avoid numerical anomalies in the square-root operation when gradients are zero in flat image regions. Its magnitude is much smaller than the typical values of image gradients and thus does not affect the physical meaning of the slope.

IV. RESULTS

A) Multi-Terrain Metric Validation

To evaluate the reconstruction network proposed in this study, three additional regions distinct from the training areas were selected as test sites, with high-resolution NAC DTM available for quantitative assessment. The three regions are APOLLO 11, Mare Imbrium, and Lichtenberg crater. Each test region is named according to the corresponding NAC DTM [20].

In this section, this study compare the proposed MMoEDTM with SOTA single-view deep learning methods for lunar digital terrain model reconstruction. Table 1 summarizes the reconstruction error (RE), mean absolute error (MAE), and root mean square error (RMSE) of the DTMs generated based on SPG.

In the APOLLO 11 region, MMoEDTM achieves the best performance on most core metrics. Specifically, it outperforms other methods on RE<2m at 63.94%, RE<4m at 91.43%, MAE of 1.83 meters, and RMSE of 2.43 meters. Compared with the second-best ELunarDTMNet [10], MMoEDTM improves RE<2m by 4.83%, RE<4m by 2.27%, MAE by 0.15 m, and RMSE by 0.13 m.

In the Mare Imbrium region, MMoEDTM achieves an RE<2m value of 52.43%, which is marginally higher than DLunarDTMNet [8] at 52.41%. RE<4m value is 82.81%, slightly lower than DLunarDTMNet at 83.17% but still outperforms ELunarDTMNet at 81.69%. The MAE of 2.43 m and RMSE of 3.63 m of MMoEDTM are within an acceptable range compared with ELunarDTMNet with MAE of 2.37 m and RMSE of 3.06 m, demonstrating the stability of the proposed method in mountain regions.

In the Lichtenberg crater region, MMoEDTM achieves the best performance across all metrics. Specifically, RE<2m at 25.22%, RE<4m at 47.47%, RE<10m at 86.43%, MAE of 5.25 m, and RMSE of 6.74 m are all superior to other methods. Compared with the second-best ELunarDTMNet, MMoEDTM improves RE<2m by 1.91%, RE<4m by 2.10%, RE<10m by 3.23%, MAE by 0.64 m, and RMSE by 1.09 m, demonstrating the robustness of the proposed method in crater terrain.

Overall, the proposed method outperforms or matches existing SOTA approaches on most key metrics across test regions. It exhibits distinct advantages in error control especially over complex terrains, which validates the effectiveness and generalization ability of the method presented in this study.

TABLE I
Comparisions Of RES Of MMoEDTM With SOTA DL Methods

Area name	Method	RE < 2m(%)	RE < 4m(%)	RE < 10m(%)	MAE(m)	RMSE(m)
APOLLO11 (Flat)	LDEMGAN [11]	48.32	73.06	91.54	2.76	3.02
	DLunarDTMNet [8]	58.68	88.83	99.71	2.02	2.62
	ELunarDTMNet [10]	59.11	89.16	**99.79**	1.98	2.56
	Ours	**63.94**	**91.43**	99.71	**1.83**	**2.43**
Imbrium (Mountain)	LDEMGAN	42.27	70.47	85.67	3.03	3.64
	DLunarDTMNet	52.41	**83.17**	99.38	**2.36**	3.09
	ELunarDTMNet	51.76	81.69	**99.64**	2.37	**3.06**
	Ours	**52.43**	82.81	98.59	2.43	3.63
Lichtenberg （Crater）	LDEMGAN	**27.52**	45.11	68.89	9.26	10.52
	DLunarDTMNet	21.27	41.47	80.12	6.47	8.71
	ELunarDTMNet	23.31	45.37	83.20	5.89	7.83
	Ours	25.22	**47.47**	**86.43**	**5.25**	**6.74**

The best results are in **blod**. The performances reported by LDEMGAN、DLunarDTMNet and ELunarDTMNet are from their original paper.

B) Generation of Flat Terrain

Fig. 4 shows the detailed reconstruction performance of a local flat terrain region in the Apollo 11 landing zone centered at 0.71° N, 23.42° E. Fig. 4(a) shows the NAC optical image at 2 m/pixel. Fig. 4(b) shows the low-resolution SLDEM slope map at 60 m/pixel. Fig. 4(c) shows the high-resolution DEM slope map at 2 m/pixel generated in this study. Fig. 4(d) shows the high-resolution NAC DTM slope map at 2 m/pixel. As a

typical lunar basaltic plain geomorphic unit, this region is characterized by low-relief and micro-textured surfaces.

In terms of the geometric morphology characterization of large craters, the reconstruction reproduces crater contour features consistent with the real terrain. For key geomorphic units, such as hundred-meter-scale craters in box 1, the reconstruction results of the proposed method are consistent with the NAC DTM. The major axis is defined as the longest diameter of the circumscribed ellipse of the crater's planar outline, and the minor axis is the shortest diameter of the ellipse. Both axis are core parameters characterizing the non-circularity of the crater's planar morphology. Based on the quantitative evaluation with parameters from box 1, the reconstructed terrain achieves a major-to-minor axis ratio of 1.704, with an error of 0.026 compared to the real ratio of 1.678.

Constrained by spatial resolution, SLDEM inherently lacks fine details, such as meter-scale craters. To verify the proposed method's generation capability for meter-scale craters, the region in box 2 within the reconstruction area is selected for quantitative comparative analysis. Quantitative statistical results indicate that, due to insufficient sampling resolution, no small craters with diameters less than 5 m are identifiable in the above region of the SLDEM data in Fig. 4(b). This limitation is mainly caused by insufficient sampling resolution. After reconstruction, 19 small craters of the same scale can be clearly identified in box 2 in Fig. 4(c). These recovered meter-scale craters show complete and clear outlines.

In box 3, the low-resolution SLDEM slope map in Fig. 4(b) contains banding artifacts. After processing by the proposed method, these artifacts are removed in Fig. 4(c). This improvement is mainly related to the training data and input images. The NAC DTM data used for training do not contain clear banding patterns. The input optical images also do not contain such striping artifacts. Therefore, the generated DTM avoids banding distortions. The elevation RMSE between the corrected result and the NAC DTM is reduced from 3.6373 m to 1.8562 m. The error is reduced to 51.03% of the original value.

In box 4 of Fig. 4(d), the existing high-resolution NAC DTM used as the reference baseline still suffers from artifact noise. Although the NAC DTM is often regarded as the ground truth for terrain analysis due to its high resolution, non-geomorphic artifact signals still remain in the data. These artifacts are affected by imaging sensor noise, data registration errors, radiometric distortion, and other factors. They manifest as irregular grayscale anomalies and false elevation fluctuations in local areas. In contrast, the geomorphic data of this region generated by the proposed method have effectively eliminated the interference from such artifact noise in Fig. 4(c). This is because the proportion of ground truth with artifacts in the training dataset is extremely small and its distribution is accidental, so the model did not capture them during training.

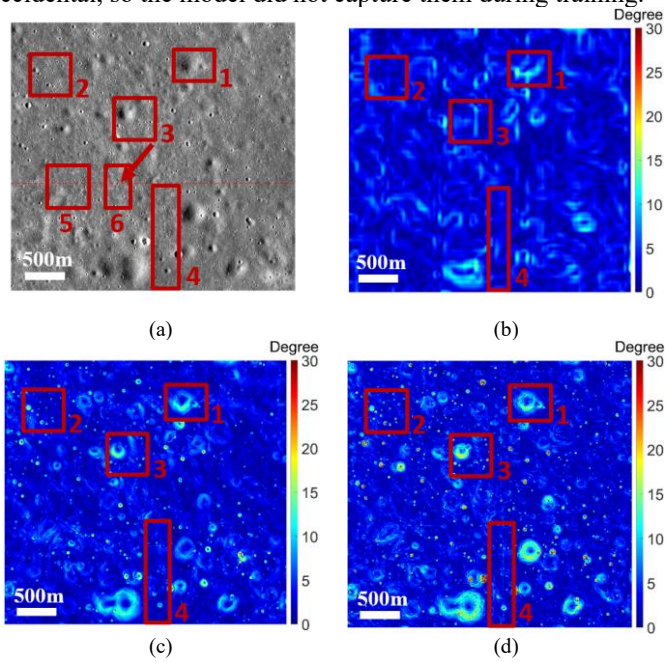


Fig. 4. The region of APOLLO 11 centered at 0.71° N, 23.42° E. (a) Optical image at resolution of 2 m/pixel. (b) SLDEM slope image at resolution of 60 m/pixel. (c) Our slope image at resolution of 2 m/pixel. (d) NAC DTM slope image at resolution of 2 m/pixel.

Fig. 5(a) and Fig. 5(b) present the digital elevation profiles along the dashed lines in Box 5 and Box 6 of Fig. 4(a), respectively. Box 5 represents a flat area with gentle topographic relief, while Box 6 focuses on small craters to evaluate fine-scale elevation reconstruction.

To illustrate details such as the absolute elevation of small crater morphology, Fig. 5(a) presents the SLDEM, NAC DTM as ground truth, and the reconstructed DTM. Over the entire region, the total elevation difference is approximately 8 m, with a horizontal span of about 500 m in real-world distance. Compared with the ground truth, the reconstructed DTM achieves a MAE of less than 0.2 m and a maximum error of approximately 0.5 m, while preserving fine-scale topographic details. Fig. 5(b) has an elevation difference of approximately 4 m over a horizontal span of 200 m in real-world distance. Compared with the ground truth data, the DTM generated by the proposed method yields a MAE of less than 0.4 m and a maximum error of approximately 1 m. Moreover, it effectively captures the undulating features of the rocky crater walls, whereas such fine topographic details are absent in the SLDEM dataset.

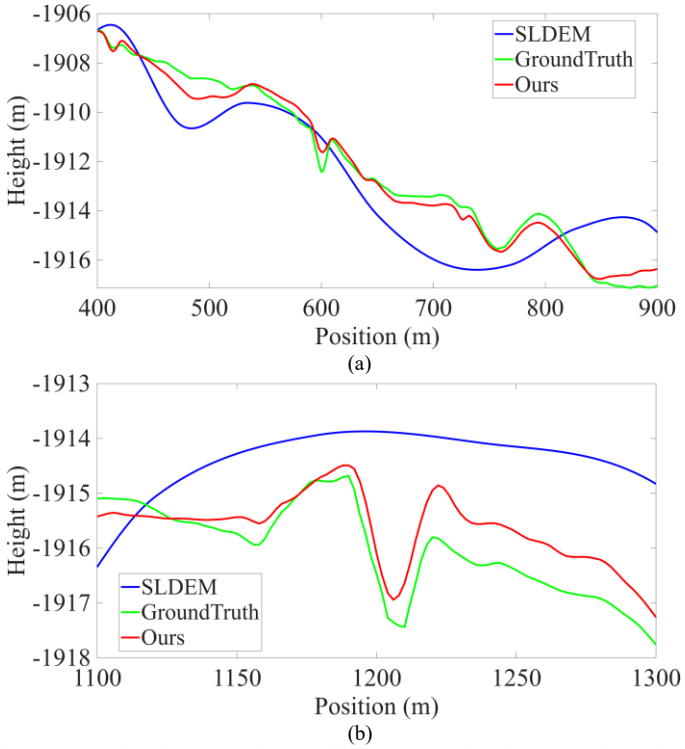


Fig. 5. DEM elevation comparison profiles (a) DEM elevation comparison of box 5 at resolution of 2 m/pixel. (b) DEM elevation comparison of box 6 at resolution of 2 m/pixel.

Fig. 6 shows the reconstruction result of a locally flat geomorphic region centered at 45.46° S and 177.52° E within the Chang'E-4 landing area in the Von Karman region. Fig. 6(a) is the NAC optical image with a resolution of 1.4 m/pixel. Fig. 6(b) is the 1.4 m/pixel DEM hillshade map generated in this study. Fig. 6(c) is the 1.4 m/pixel DEM hillshade produced by MADNet [13]. Fig. 6(d) is the DEM hillshade at the resolution of 1.4 m/pixel derived from NAC DTM. Fig. 6(e) is the DEM hillshade at the resolution of 60m/pixel generated by SLDEM. Fig. 6(f) is SLDEM, and Fig. 6(g)–(j) are the slope maps of Figs. 6 (b)–(e) respectively.

As indicated by the arrows, the reconstruction result of large craters shown in Fig. 6(b) is superior to that of MADNet in Fig. 6(c). MADNet reconstructs large crater regions. These regions exhibit obvious flattening. As a result, MADNet fails to accurately reproduce their original topographic relief characteristics. In contrast, the proposed method can restore the geomorphic morphology of large craters. It can also restore their depth information.

Box 1 in Fig. 6 demonstrates the ability of the proposed method to remove artifacts from low-resolution data. The original SLDEM data shown in Fig. 6(e) has a resolution of 60 m/pixel and exhibits 3 distinct vertical artifacts. The DEM generated by our method in Fig. 6(b) eliminates these artifact stripes.

Box 2 in Fig. 6 focuses on the detailed reconstruction of meter-scale craters around large craters. As shown in Fig. 6(b), the proposed method captures the crater edges, depths, and spatial distribution of small craters. In contrast, MADNet in Fig. 6(c) exhibits relatively weak performance in depicting meter-scale craters near large craters. The slope of the reconstructed small craters are less than 10°, with the terrain becoming overly flat and losing fine geomorphic features.

Box 3 in Fig. 6 reveals the accuracy limitations of the MADNet in small crater reconstruction. Although MADNet generates a dense distribution of small craters, the crater slopes in Fig. 6(h) are too large and exceed 15°. This result shows clear distortion when compared with the NAC DTM slope in Fig. 6(i). In contrast, the proposed method preserves the detailed morphology of small craters in Fig. 6(g). The crater contours are close to those in the reference terrain. The slope remain below 10°, which is consistent with Fig. 6(i).

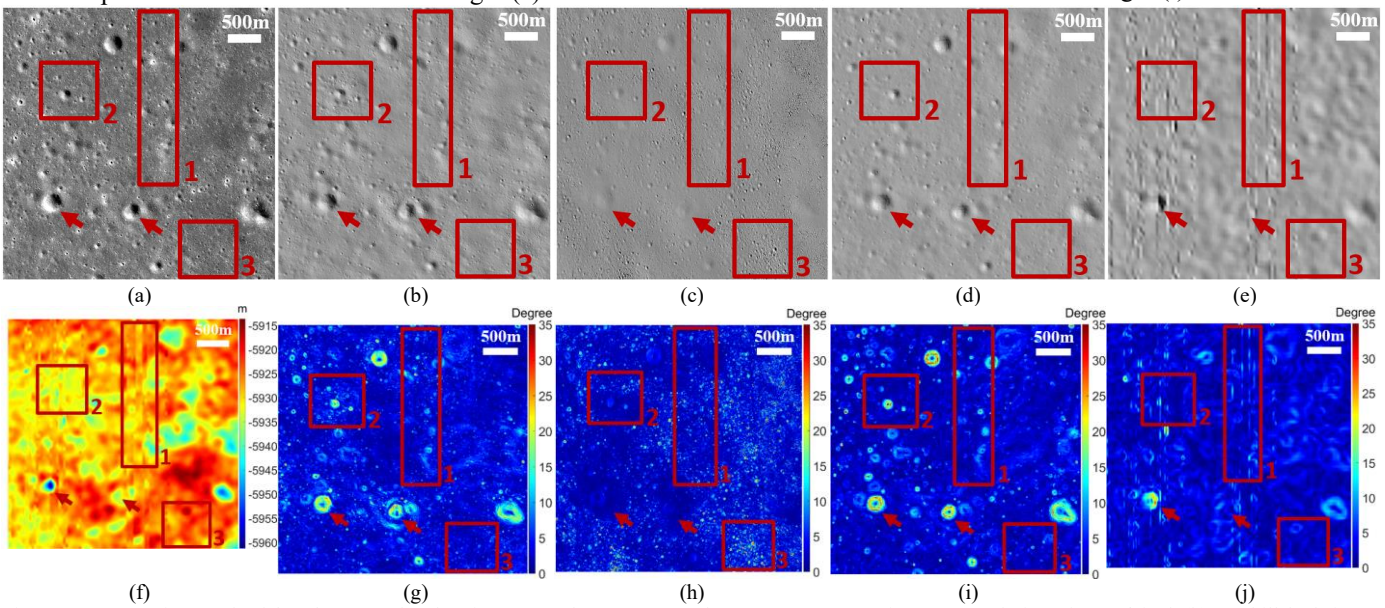


Fig. 6. Reconstruction result of the Chang'E-4 local region centered at 45.46° S and 177.52° E. (a) NAC image at resolution of 1.4 m/pixel. (b) Our hillshade image at resolution of 1.4 m/pixel. (c) MADNet hillshade image at resolution of 1.4 m/pixel. (d) NAC DTM hillshade image at resolution of 1.4 m/pixel. (e) SLDEM hillshade image at resolution of 60 m/pixel. (f) SLDEM image at resolution of 60 m/pixel. (g) Our slope image at resolution of 1.4 m/pixel. (h) MADNet slope image at resolution of 1.4 m/pixel. (i) NAC DTM slope image at resolution of 1.4 m/pixel. (j) SLDEM slope image at resolution of 60 m/pixel.

C) Generation of Crater Terrain

The region shown in Fig. 7 is located on the crater wall of the Lichtenberg crater, with a central geographic coordinate of 31.57° N, 67.86° W. Fig. 7(a) shows the high-resolution optical image at 0.5 m/pixel. Fig. 7(b) shows the high-resolution hillshade image at 0.5 m/pixel generated in this study. Fig. 7(c) shows the high-resolution NAC DTM hillshade map at 0.5 m/pixel. Fig. 7(d) shows the low-resolution SLDEM image at 60 m/pixel. Fig. 7(e) shows the high-resolution DEM slope map generated by the proposed method at 0.5 m/pixel. Fig. 7(f) shows the high-resolution NAC DTM slope map at 0.5 m/pixel.

As indicated by the arrows, the slope map derived from NAC DTM in Fig. 7(c) exhibits horizontal artifacts, but this study's result in Fig. 7(b) doesn't have artifacts. The reasons why MMoEDTM can eliminate these stripes will be explained in Section 5. Box 1 in Fig. 7(c) and Fig. 7(f) show that NAC DTM is difficult to generate meter-scale, densely distributed small craters, resulting in a homogenized terrain in this region that deviates significantly from the real terrain texture in the NAC image. However, the proposed method in Fig. 7(b) and Fig. 7(e) enables high-fidelity reconstruction of dense small craters, and its output is consistent with the features of the NAC image in Fig. 7(a) and Fig. 7(d). As shown in box 2, Fig. 7(b) can characterize the nested crater-in-crater structure. It also matches the corresponding geomorphic details in Fig. 7(a). By contrast, Fig. 7(c) shows blurred crater boundaries and lacks accurate topographic information for this nested structure.

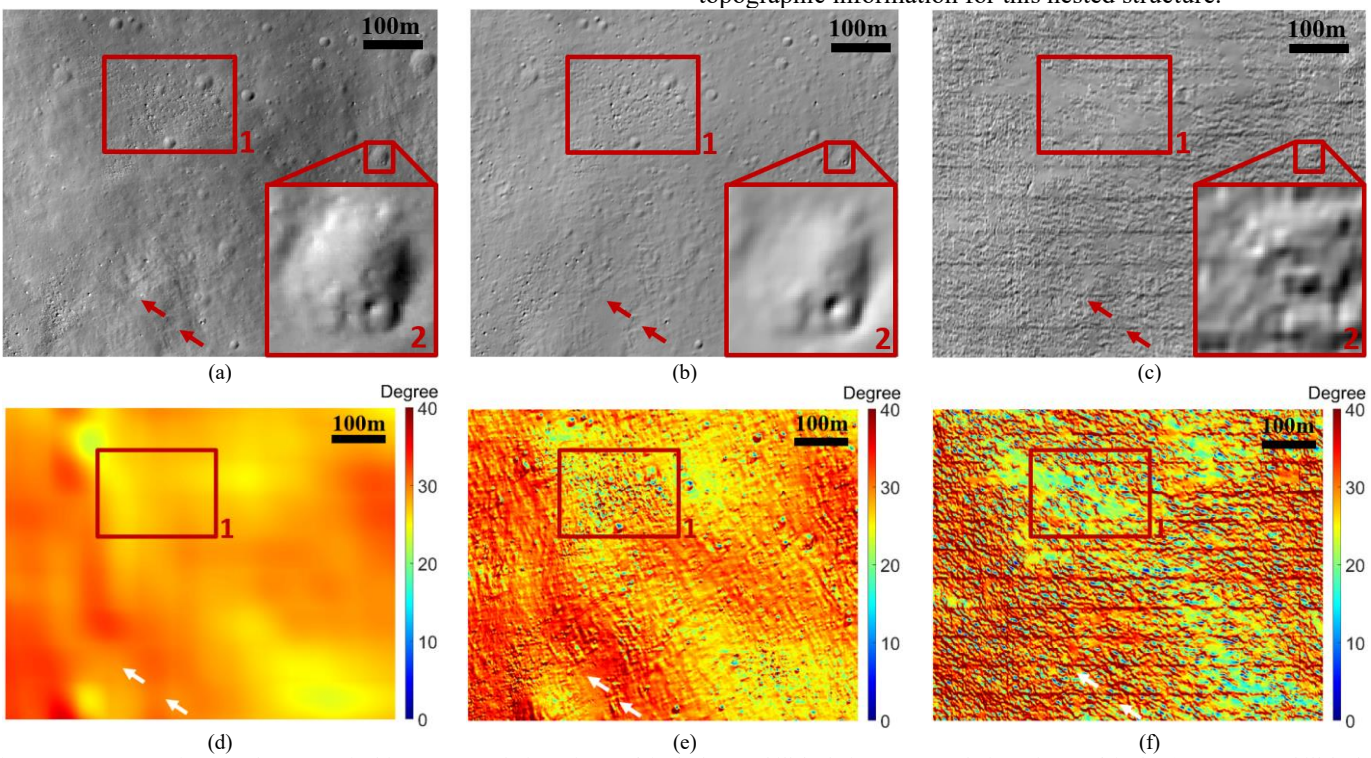


Fig. 7. Crater Generation Results. (a) Optical image at resolution of 1.2 m/pixel. (b) Our hillshade image at resolution of 1.2 m/pixel. (c) NAC DTM hillshade image at resolution of 1.2 m/pixel. (d) SLDEM slope image at resolution of 60 m/pixel. (e) Our slope image at resolution of 1.2 m/pixel. (f) NAC DTM slope image at resolution of 1.2 m/pixel.

D) Generation of Discrete Rock Regions

To evaluate the performance of generating a discrete rock region, a relevant region on a crater wall at 23.684° N, 313.117° E was selected, covering an area of 1020 m × 720 m with an absolute elevation difference of 369 m. Fig. 8(a) presents the high-resolution optical image at 1.2 m/pixel. Fig. 8(b) provides the high-resolution DEM hillshade at 1.2 m/pixel generated in this study. Fig. 8(c) displays the high-resolution NAC DTM hillshade at 1.2 m/pixel, while Fig. 8(d) illustrates the SLDEM slope map at 60 m/pixel. Fig. 8(e) depicts the high-resolution DTM slope at 1.2 m/pixel generated in this study. Fig. 8(f) gives the high-resolution NAC DTM slope at 1.2 m/pixel. Figs. 8(g)-(i) are enlarged views of box 1 in Figs. 8(a)-(c), respectively. Similarly, Figs. 8(j), 8(k), and 8(l) correspond to enlarged views of box 2 in Figs. 8(a)-(c), respectively.

As shown in the optical image in Fig. 8(a), numerous discrete rock blocks with diameters less than 10 m exist in the region. The grayscale contrast between rock edges and crater wall bedrock is significant. The rock distribution exhibits clear heterogeneity, with dense rock stacking in local areas and isolated, discrete rocks in most areas.

A comparison between Figs. 8(a) and 8(b) shows that the rock distribution generated by the proposed method agrees with the real optical image in both shape and location. The shadow boundaries of discrete rocks are clear. The shadow transition at the rock-bedrock contact is natural. No obvious distortion or over-smoothing is observed. In contrast, the NAC DTM hillshade in Fig. 8(c) is strongly affected by artifact noise. As a result, the micro-relief information of discrete rocks is lost.

Box 1 in Fig. 8(g) shows 14 typical discrete rocks with diameters larger than 10 m outside the dense rock cluster. The generated result in Fig. 8(h) matches Fig. 8(g) in spatial location. The differences in contour size and position are within 10 pixels relative to Fig. 8(a). For the large rock with an area greater than 50 m² marked by the green arrow in Fig. 8(g), K-means

clustering identifies 103 shadow pixels. Under the same solar incidence azimuth, the corresponding generated result in Fig. 8(h) contains 115 shadow pixels. The difference is 11.7%. The large rock with an area greater than 50 m², indicated by the yellow arrow, has 108 shadow pixels, compared with 106 in the generated result, with a 1.9% difference.

Box 2 shows that the typical geomorphic feature of this region is a discretely distributed rock-stacking structure. Under the same solar incidence azimuth, the number of shadow pixels in the rock region indicated by the green arrow in the optical image is 170, while the corresponding number in the restored rock stack is 199, representing an increase of 17.1%. In contrast, Fig. 8(l) exhibits obvious positional deviation compared with Fig. 8(j), along with blurred rock contours.

Box 3 in Fig. 8(d) exhibits obvious thick vertical artifacts, and Fig. 8(f) also shows relatively thin vertical artifacts, these artifacts are eliminated in generated Fig. 8(e).

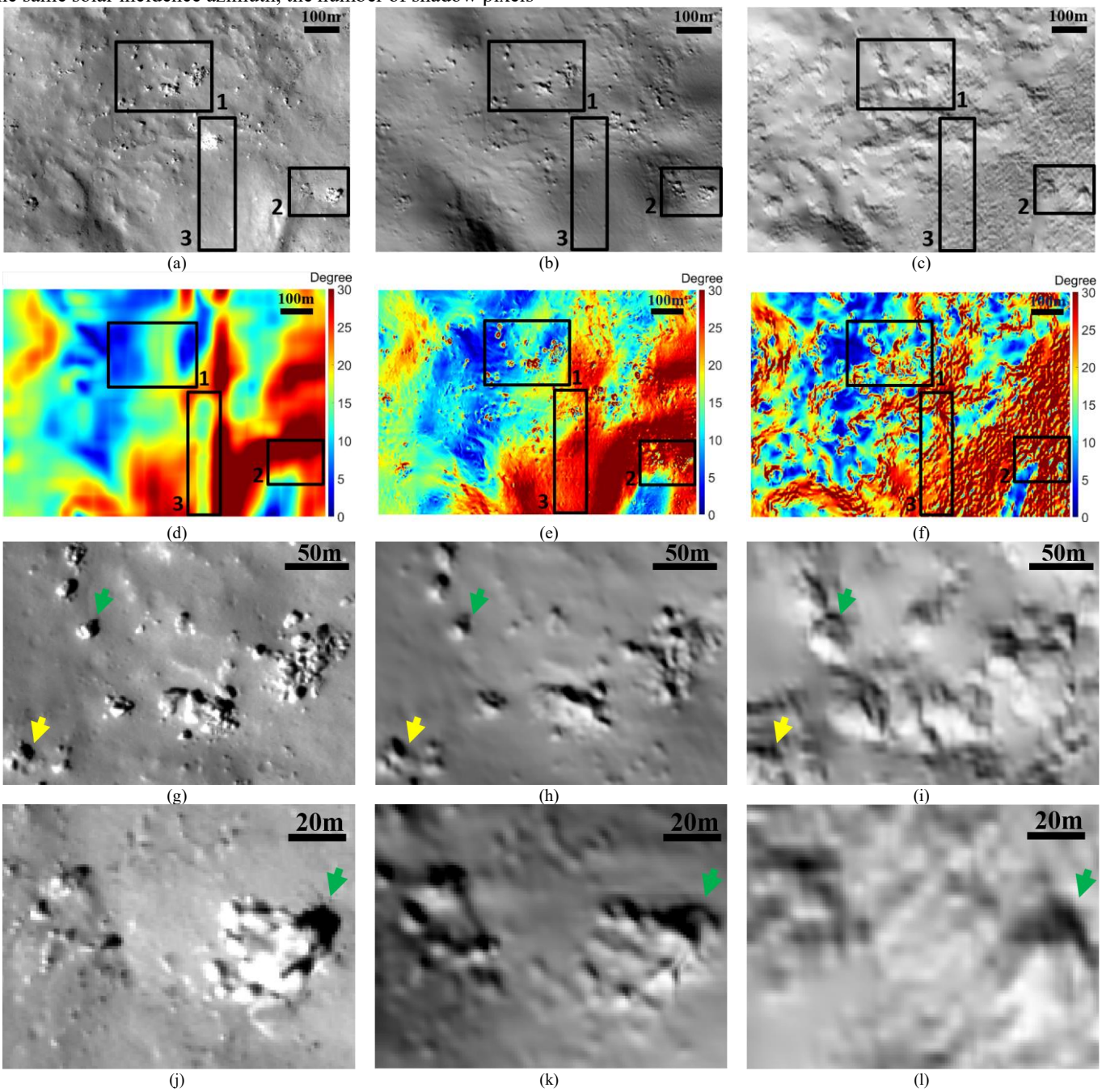


Fig. 8. Generation Results of Discrete Rock Regions. (a) Optical image at resolution of 1.2 m/pixel. (b) Our hillshade image at resolution of 1.2 m/pixel. (c) NAC DTM hillshade image at resolution of 1.2 m/pixel. (d) SLDEM slope image at resolution of 60 m/pixel. (e) Our slope image at resolution of 1.2 m/pixel. (f) NAC DTM slope image at resolution of 1.2 m/pixel. (g) Box 1 of optical image at resolution of 1.2 m/pixel. (h) Box 1 of our hillshade image at resolution of 1.2 m/pixel. (i) Box 1 of NAC DTM hillshade image at resolution of 1.2 m/pixel. (j) Box 2 of optical image at resolution of 1.2 m/pixel. (k) Box 2 of our hillshade image at resolution of 1.2 m/pixel. (l) Box 2 of NAC DTM hillshade image at resolution of 1.2 m/pixel.

E) Generation of Dense Rock Regions

To verify the applicability of the proposed method in dense rock regions, a crater region containing dense rocks at 19.695°

N, 30.818° E was selected, covering an area of 510 m × 360 m with an absolute elevation difference of 242 m. Fig. 9(a) presents the optical image at 0.6 m/pixel. Fig. 9(b) provides the high-resolution DTM hillshade at 0.6 m/pixel generated in this study. Fig. 9(c) displays the NAC DTM hillshade at 0.6 m/pixel. Fig. 9(d) illustrates the low-resolution SLDEM slope map at 60 m/pixel. Fig. 9(e) depicts the high-resolution DEM slope map at 0.6 m/pixel generated in this study. Fig. 9(f) gives the NAC DTM slope map at 0.6 m/pixel. Enlarged views of the boxed regions are shown in Figs. 9(g)-(i) for Figs. 9(a)-(c), respectively.

Because the central part of the optical image in Fig. 9(a) has very high albedo and cannot be adequately modeled, the left region was selected for discussion. A total of 29 rock blocks with diameters larger than 1 m are identified within box 1 in Fig. 9(g). The number of rock blocks at the same scale in Fig. 9(h) is 17. This represents a decrease of about 41.4%. Within box 2 in Fig. 9(g), 38 rock blocks with diameters larger than 1 m are identified. In Fig. 9(h), 23 rock blocks at the same scale are identified. This represents a decrease of about 39.5%. This difference in rock count results from limited fine-grained segmentation accuracy in the generation model. Some small rocks are not effectively separated and are merged into larger rock bodies. Nevertheless, the generated results remain highly consistent with the spatial distribution of rocks in the real scene. Further statistics of shadow pixels show that the rock marked by the arrow in Fig. 9(g) contains 111 shadow pixels. The corresponding rock in Fig. 9(h) contains 108 shadow pixels. The difference is about 2.7%. In contrast, Fig. 9(i) contains fewer than 100 shadow pixels. Its shadow morphology differs greatly from those in Figs. 9(g) and 9(h).

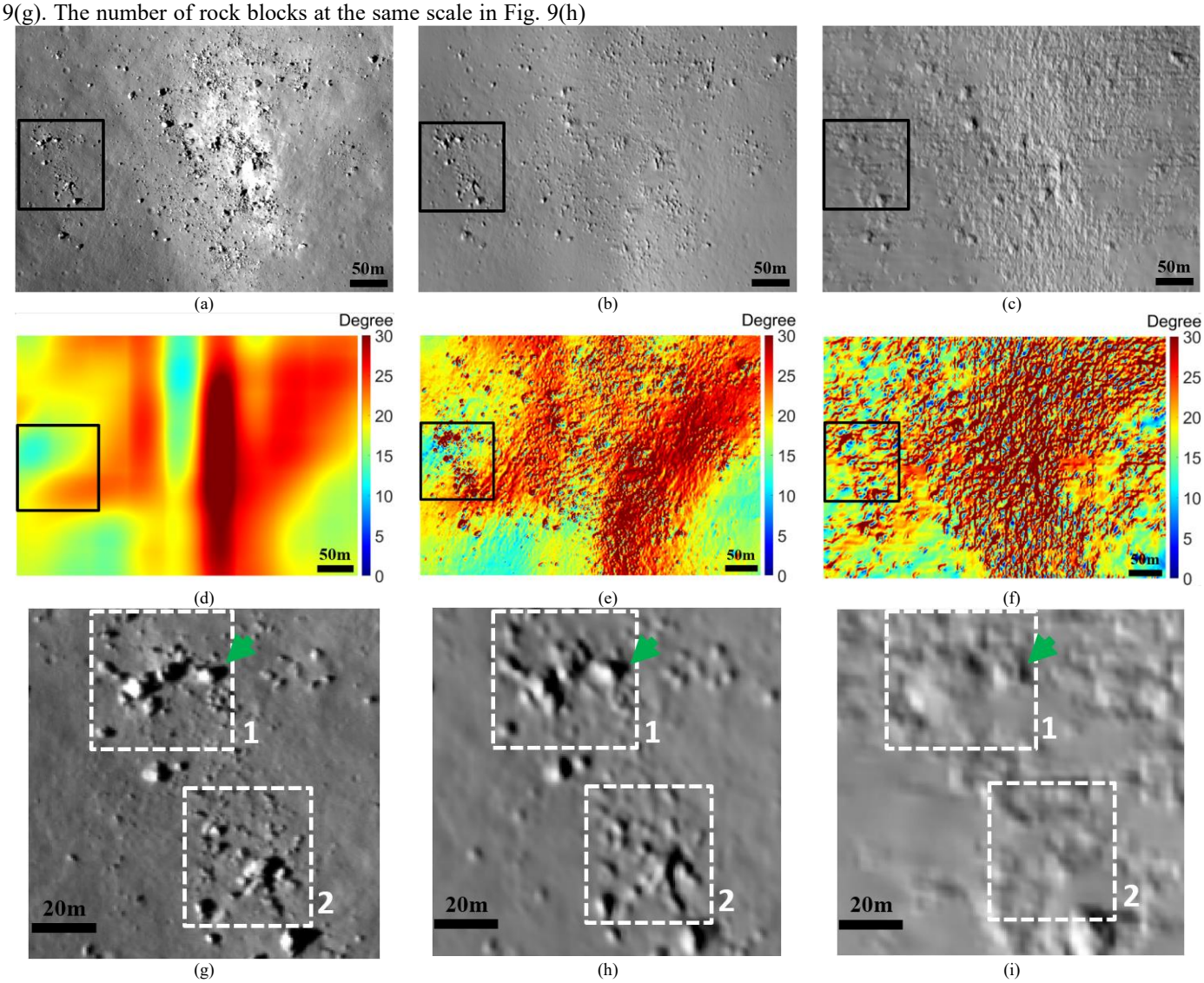


Fig. 9. Generation Results of Dense Rock Regions. (a) Optical image at resolution of 0.6 m/pixel. (b) Our hillshade image at resolution of 0.6 m/pixel. (c) NAC DTM hillshade image at resolution of 0.6 m/pixel. (d) SLDEM slope image at resolution of 60 m/pixel. (e) Our slope image at resolution of 0.6 m/pixel. (f) NAC DTM slope image at resolution of 0.6 m/pixel.

V. Application

A) Extraction and Validation of Rock Roughness

To verify the reliability of rock generation, the rock-rich Galilei-E region was reconstructed. DEM roughness parameters were then extracted and compared with SAR images.

Fig. 10 illustrates the high-resolution DTM generation result for the Galilei-E crater [17], located in the mid-low latitude

region at central coordinates of 13.9° N, 61.9° E. Fig. 10(a) shows the high-resolution optical image at 1 m/pixel. Fig. 10(b) is the 60 m/pixel low-resolution SLDEM slope map. Fig. 10(c) shows the high-resolution DTM slope map at 1 m/pixel resolution generated in this study. Fig. 10(d) displays the corresponding roughness map at 1 m/pixel. According to the definition [21], [22], roughness is calculated as the variance of slope values within a 5×5 sliding window centered on each grid cell in the 2D raster. The roughness map is used because a DTM slope map alone cannot describe texture caused by small-scale rocks. Fig. 10(e) shows the SAR image at 14 m/pixel.

Compared with Fig. 10(b), the generated DTM in Fig. 10(c) restores fine geomorphic units. These units include microtopographic undulations, terrace-like erosion traces, and small craters on the crater wall slope. The result also shows the texture levels and boundary features formed by material accumulation and erosion on the crater wall.

SAR images show strong backscatter in rock-dense regions. This response can support the identification of rock-rich areas. Multi-source overlay analysis shows that the boundaries of high-roughness regions in Fig. 10(d) are consistent with the grayscale anomaly boundaries of rock-rich areas in Fig. 10(a). They also match the spatial extent of bright high-backscatter spots in the SAR image in Fig. 10(e).

Figs. 10(f)–(j) show enlarged local details of the red-boxed regions in Figs. 10(a)–(e). These regions show strong topographic relief and many small rocks. Optical images are sensitive to illumination and shadow. SAR images are easily affected by speckle noise. In contrast, the generated DTM roughness map more clearly defines the boundary between individual rocks and surrounding slopes. It also shows the density and spatial aggregation of rock distribution. In the three circled regions in Fig. 10(i) and Fig. 10(j), the extracted DEM roughness values of small rock regions are greater than 100 degree². These regions correspond to Stokes parameter values greater than 2.5 dB in the SAR image.

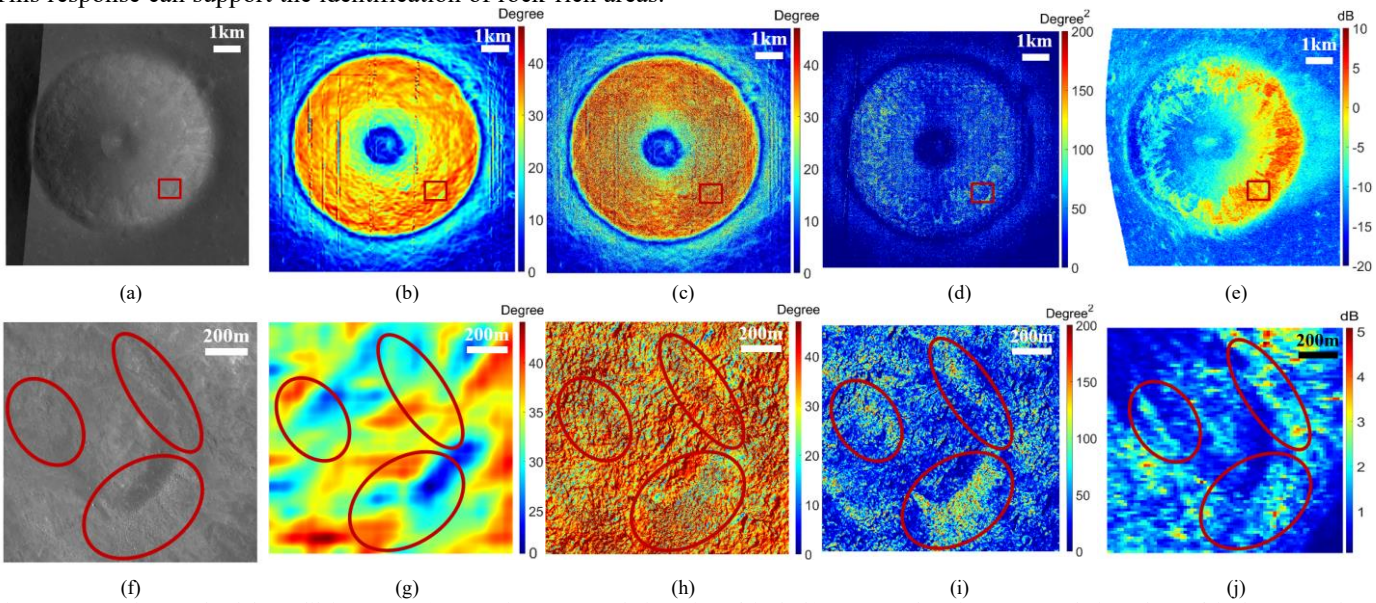


Fig. 10. Generation result of the Galilei-E crater. (a) NAC image at resolution of 1 m/pixel. (b) SLDEM slope image at resolution of 60 m/pixel. (c) Our slope image at resolution of 1 m/pixel. (d) Our roughness map at resolution of 1 m/pixel. (e) SAR image at resolution of 14 m/pixel. (f) Box of NAC image at resolution of 1 m/pixel. (g) Box of SLDEM slope image at resolution of 60 m/pixel. (h) Box of our slope image at resolution of 1 m/pixel. (i) Box of our roughness map at resolution of 1 m/pixel. (j) Box of SAR image at resolution of 14 m/pixel.

B) Generation of the PSR in Shackleton Crater

The lunar axial tilt is 1.54°. Some craters near the lunar poles receive no direct solar illumination throughout the year, and these regions are known as PSRs [23], [24], [25], [26], [27], [28].

Generating a DTM from high-resolution optical images under low illumination is a technical challenge, primarily constrained by two core limitations. First, insufficient illumination results in a significant decrease in the signal-to-noise ratio of optical images. It makes detailed image information more easily obscured by noise and directly affecting the accuracy and stability of feature point extraction during DTM generation. Second, under low-light conditions, the grayscale contrast of surface features is compressed and textural details are blurred, and the surface is susceptible to the inhomogeneity of weak reflected light. This prevents key algorithms, such as image matching and stereo intersection, from operating effectively, resulting in the spatial resolution and elevation accuracy of the generated DTM failing to meet high-level research requirements.

To address the above issues, several low-illumination optical images are specifically selected for model training in this study. For the generation of the PSR in Shackleton crater, the optical images of ShadowCam are adopted as input data. These images cover M047430847S, M047402566S, M045045299S, M045002872S, M047494456S, M047536878S, and M044974588S. The DEM product at 60 m/pixel resolution, LDEM_75N_60M_FLOAT, is used as the low-resolution DEM input. The DTM product at 5 m/pixel resolution, ldem_87s_5mpp, is employed as a reference to reconstruct the DTM of the Shackleton PSR at a resolution of approximately 1.88 m/pixel. Fig. 11(a) presents the 1.88 m/pixel ShadowCam image. Fig. 11(b) displays the slope map derived from the 60 m/pixel DEM. Fig. 11(c) shows the 5 m/pixel LDEM slope map,

while Fig. 11(d) illustrates the high-resolution 1.88 m/pixel DTM slope map produced in this work.

Figs. 11(e)–(h) show enlarged local details of the flat area in Box 1 of Figs. 11(a)–(d). Figs. 11(i)–(l) show enlarged local details of the crater wall region in Box 2 of Figs. 11(a)–(d). The visual results show that the reconstructed DTM recovers small crater structures better than the original 60 m/pixel resolution DEM. It also delineates fine topographic features, including crater rim textures, small protrusions, and micro-depressions.

For the flat terrain in Box 1, the corrected product [29] in Fig. 11(g) still contains residual stripe artifacts and lacks detailed crater information. By contrast, the high-resolution DTM in Fig. 11(h) removes these stripe artifacts and restores small craters.

For the crater region in Box 2, Fig. 11(k) is affected by severe artifacts. Detailed crater morphology is therefore difficult to identify. By contrast, the generated DTM in Fig. 11(l) suppresses stripe noise and reconstructs tiny craters on the crater wall. It shows richer topographic details than Fig. 11(j) and Fig. 11(k). The distribution and abundance of rocky features are also consistent with those observed in Fig. 11(i).

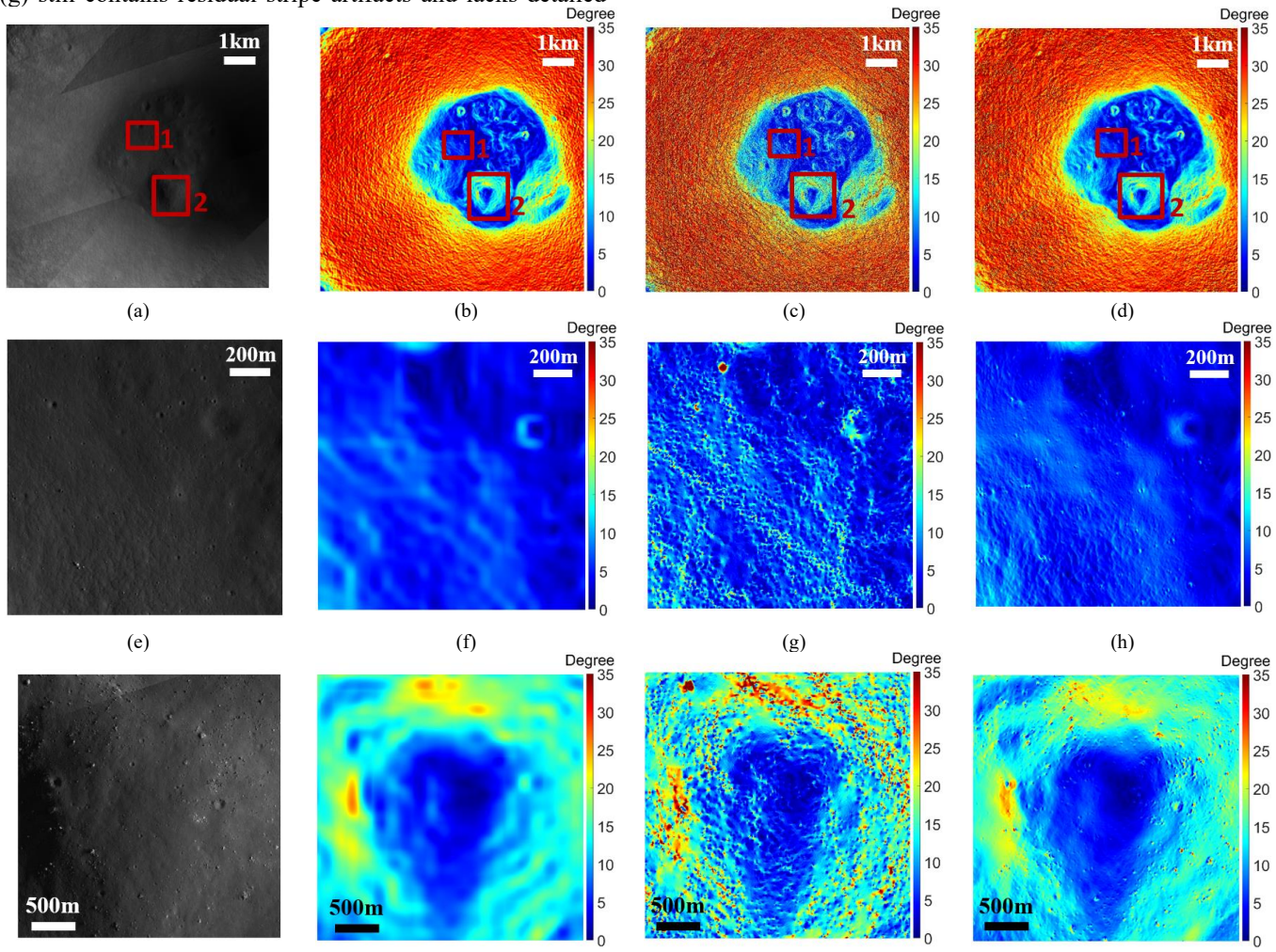


Fig. 11. Generation Results of Shackleton Regions. (a) ShadowCam image at resolution of 1.88 m/pixel. (b) LDEM slope image at resolution of 60 m/pixel. (c) ldem_87s_5mpp slope image at resolution of 5 m/pixel. (d) Our slope image at resolution of 1.88 m/pixel. (e) Box 1 of ShadowCam image at resolution of 1.88 m/pixel. (f) Box 1 of LDEM slope image at resolution of 60 m/pixel. (g) Box 1 of ldem_87s_5mpp slope image at resolution of 5 m/pixel. (h) Box 1 of our slope image at resolution of 1.88 m/pixel. (i) Box 2 of ShadowCam image at resolution of 1.88 m/pixel. (j) Box 2 of LDEM slope image at resolution of 60 m/pixel. (k) Box 2 of ldem_87s_5mpp slope image at resolution of 5 m/pixel. (l) Box 2 of our slope image at resolution of 1.88 m/pixel.

VI. Discussion

A) Influence of Solar Incidence Angle

During the generation process, changes in the illumination angle of the input high-resolution optical images would affect the generation performance. Therefore, the influence of the solar incidence angle on the model inference capability is analyzed in this section.

To examine the influence of solar incidence angle on DEM generation performance, only the solar incidence angle is set as a variable. NAC DTM data are downsampled by a factor of 9 to 18 m/pixel and used as the low-resolution DEM input. Optical images are downsampled to 2 m/pixel for inference. Fig. 12 shows slope maps of high-resolution DTMs generated from optical images under different solar incidence angles. The test region is located at low latitude, with central coordinates of 0.55° N, 23.47° E.

Figs. 12(a)–(d) correspond to optical image with solar incidence angles of 18.85°, 45°, 62.61°, and 73.64°, respectively. The spatial resolution is 2 m. The results show that

the DEM generated at a solar incidence angle of 62.61° provides the best detail restoration and texture representation. The result generated at 45° ranks second. The results generated at 18.85° and 73.64° are less effective. These results indicate that solar incidence angle strongly affects high-resolution DTM quality. The preferred range is approximately 40° to 60°.

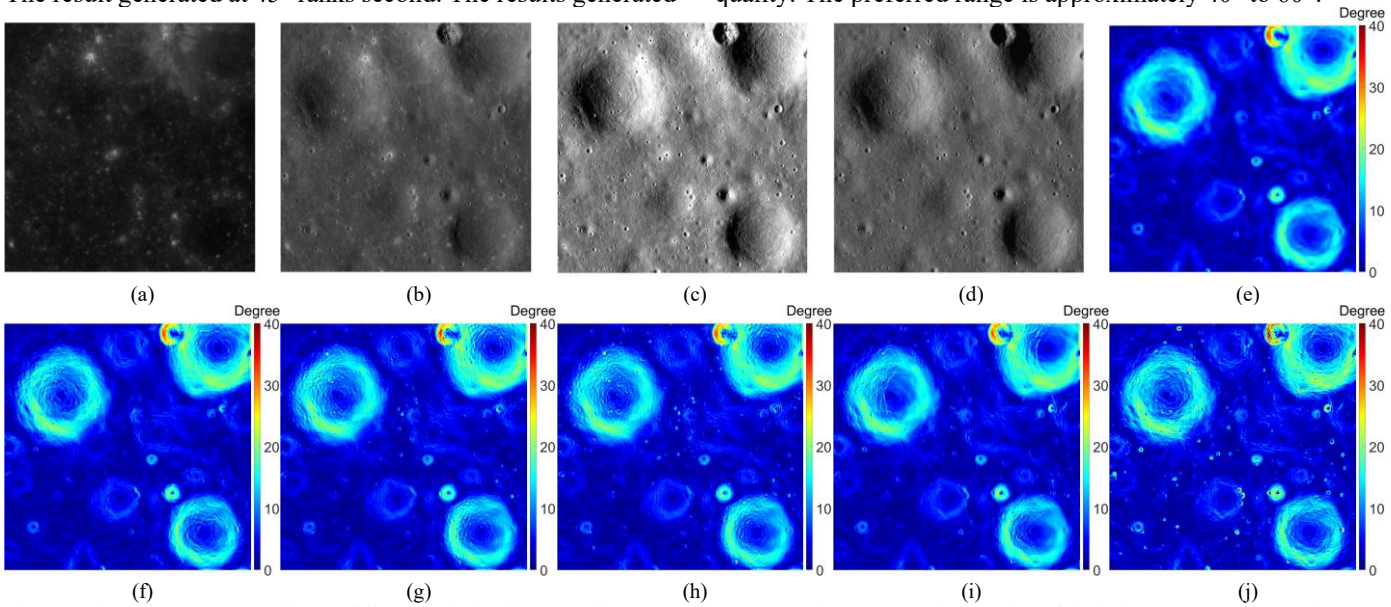


Fig. 12. Slope maps corresponding to different solar incidence angles. (a) M188085530RC image at resolution of 2 m/pixel. (b) M1190561570LC image at resolution of 2 m/pixel. (c) M150361817RC image at resolution of 2 m/pixel. (d) M1157600009RC image at resolution of 2 m/pixel. (e) Coarse DEM image at resolution of 18 m/pixel. (f) M188085530RC slope image at resolution of 2 m/pixel. (g) M1190561570LC slope image at resolution of 2 m/pixel. (h) M150361817RC slope image at resolution of 2 m/pixel. (i) M1157600009RC slope image at resolution of 2 m/pixel. (j) NAC DTM slope image at resolution of 2 m/pixel.

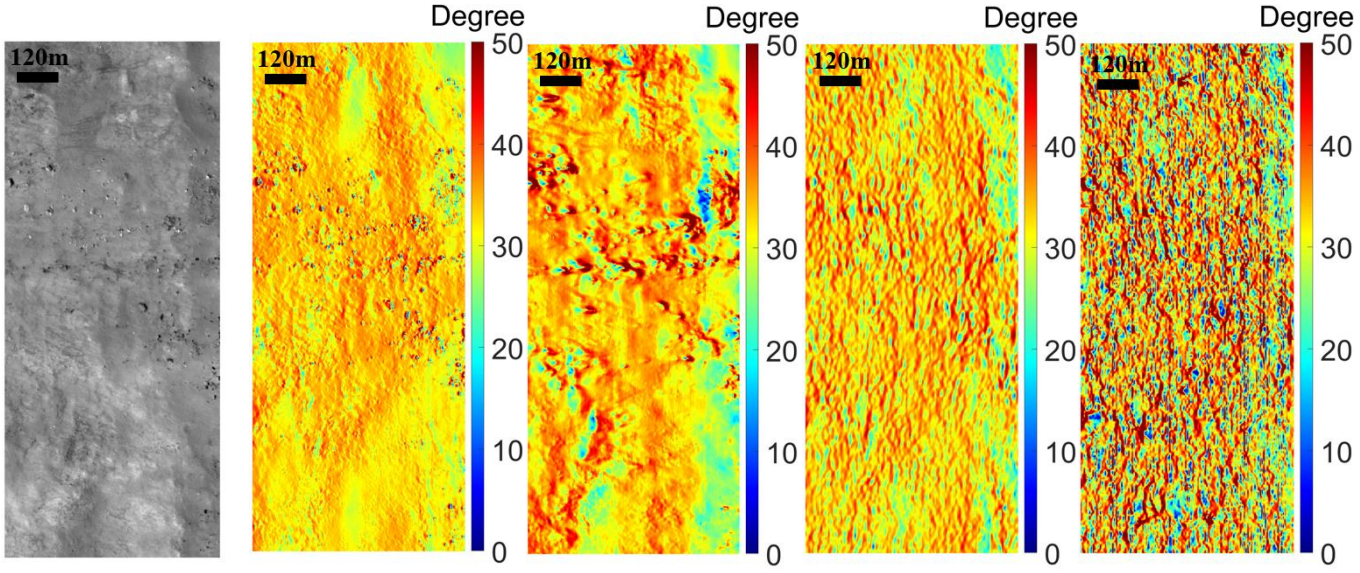
B) Artifacts Elimination

The slope expert model designed for lunar slope regions can generate slope data without vertical stripes. These stripes are present in the NAC DTM. This phenomenon is related to the proportion of training data from different terrain types. Fig. 13(a) shows the NAC optical image, with a spatial resolution of 1.2 m/pixel. Fig. 13(b) shows the slope map at 1.2 m/pixel predicted by the proposed slope expert model. Its training dataset contains 70% slope areas and 30% flat areas. Both subsets contain abundant topographic information. Fig. 13(c) shows the slope map at 1.2 m/pixel calculated by the flatland expert model. Its training data consist entirely of detailed flat-terrain samples. Fig. 13(d) shows the slope map at 1.2 m/pixel output by the model trained only on slope terrain datasets. Fig. 13(e) shows the NAC DTM slope map at 1.2 m/pixel, which serves as the reference.

As shown in Fig. 13(d), many clear vertical stripes remain when the model trained on a 100% crater dataset is used for crater regions. By contrast, the model trained on a 100% flat-region dataset shows no such stripes. This difference is caused by region-specific artifact noise in the NAC DTM slope used as ground truth. Artifact noise is absent or much weaker in flat regions. Models trained only on crater data can directly learn and reproduce these artifacts. To reduce this problem, about 30% flat-region data with rich topographic details were included in the training dataset of the slope expert model. As a result, this noise was effectively suppressed. Therefore, the crater results generated by the proposed model are free from stripe interference.

The flat expert model trained on a 100% flat-region dataset also has limitations. Severe height distortion occurs when this model is directly applied to slope regions. This is because slope terrain elevation characteristics are not sufficiently learned. As a result, the model cannot accurately describe relief features of slope landforms, such as crater walls.

The improved rock reconstruction performance of the proposed slope expert model is mainly related to its training dataset. This dataset contains 70% slope areas and 30% detailed flat terrain areas. This configuration combines complementary information from the two subsets. Flat terrain data support fine-grained detection and reconstruction of small surface targets, as shown in Fig. 13(c) and Fig. 4(c). By contrast, models trained only on slope terrain can model elevation well, but their ability to identify small geological objects remains limited.



(a) (b) (c) (d) (e)

Fig. 13. Generation results of slope regions using different training datasets. (a) NAC image at resolution of 1.2 m/pixel. (b) Slope image of our slope expert model at resolution of 1.2 m/pixel. (c) Hillshade image of our flat expert model at resolution of 1.2 m/pixel. (d) Slope image of our pure slope expert model at resolution of 1.2 m/pixel. (e) NAC DTM slope image at resolution of 1.2 m/pixel.

C) Ablation Study

To verify the effectiveness of each key component in the proposed framework, ablation experiments are conducted on the floor region of Shackleton crater, as reported in Table II. Performance is evaluated using the percentage of reconstruction results with RE below specific thresholds, as well as MAE and RMSE in meters.

First, the effectiveness of the proposed MMoE module is evaluated. Compared with the baseline BaseModel+Dense, BaseModel+MoE shows performance degradation on most metrics. In BaseModel+MoE, the FFN layer in the Swin Transformer is replaced by an MoE layer. The setting uses 4 experts, top-2 gating, and an auxiliary loss weight of 0.01. For example, RE < 2 m decreases from 38.47% to 38.42%, and RMSE increases from 4.53 m to 4.59 m. By contrast, BaseModel+MMoE improves the results. RE < 2 m and RE < 4 m increase to 39.41% and 66.96%, respectively. MAE and RMSE decrease to 3.27 m and 4.18 m, respectively.

Next, the role of the CA module is analyzed. When CA is added to BaseModel+MMoE, model performance is further improved. BaseModel+MMoE+CA increases the percentage of RE < 2 m to 43.86%. It also reduces MAE and RMSE to 3.13 m and 3.95 m, respectively. This result shows that the CA module helps preserve long-range contextual information. It also helps the model recover fine details in complex crater terrain.

Finally, the full model integrating MMoE, CA, and Lslope loss is evaluated. This model is denoted as BaseModel+MMoE+Lslope+CA. The experimental results show that this model achieves the best performance across all metrics. The percentages of RE < 2 m, RE < 4 m, and RE < 10 m reach 44.46%, 74.85%, and 98.83%, respectively. MAE and RMSE decrease to 2.84 m and 3.67 m, respectively. Compared with the baseline BaseModel+Dense, the full model achieves a 15.6% relative improvement in RE < 2 m and reduces RMSE by 19.0%. These results verify the complementary effects of all components.

It is worth noting that BaseModel+MMoE+Lslope leads to performance degradation. This result indicates that the slope loss alone cannot work effectively without other supporting modules. When combined with MMoE and CA, it provides additional supervision for terrain geometry and leads to the best final performance.

These ablation results show that MMoE, CA, and Lslope loss are all critical components of the model. Their effective integration is essential for high-precision DEM reconstruction of the lunar surface.

TABLE II
Comparations Of Results From Well-Trained Models For Each Component in Shackleton

Method	RE < 2m(%)	RE < 4m(%)	RE < 10m(%)	MAE(m)	RMSE(m)
BaseModel+Dense	38.47	64.47	95.35	3.62	4.53
BaseModel+MoE	38.42	64.21	95.00	3.66	4.59
BaseModel+MMoE	39.41	66.96	98.18	3.27	4.18
BaseModel+MMoE+Lslope	37.98	66.24	97.51	3.37	4.32
BaseModel+MMoE+CA	43.86	70.45	97.27	3.13	3.95
BaseModel+Dense+Lslope+CA	42.74	68.58	95.53	3.36	4.57
BaseModel+MoE+Lslope+CA	41.75	67.36	96.56	3.35	4.45
BaseModel+MMoE+Lslope+CA(ours)	**44.46**	**74.85**	**98.83**	**2.84**	**3.67**

The best results are in **blod**.

VII. Conclusion

This paper proposes MMoEDTM, a lightweight single-view lunar DTM reconstruction framework based on MRF-based MoE architecture. Unlike conventional dense models that suffer from limited generalization across diverse terrains, MMoEDTM achieves decoupled terrain modeling by using MRF to classify terrains and assign samples to dedicated expert networks. This design sustains robustness in crater and flat regions.

This study presents several designs to enhance generalization ability and network performance. First, MRF-based MoE architecture is proposed to improve the generalization performance of small-scale models. Second, CA module is introduced to embed low-resolution DEMs as terrain priors, bridging the representation gap between deep features and elevation information. Third, Lslope loss function, derived from the mathematical definition of slope maps, provides explicit constraints to preserve terrain details and contour fidelity in complex terrain reconstruction. It limited the rock shadow error to approximately 10%, which addresses the long-standing challenge of reconstructing rough rocky surfaces.

Validation experiments cover multiple lunar terrains, including mid-low latitude craters, flat regions, dense rock regions, and discrete rock regions. In the Galilei-E crater, quantitative roughness evaluation and SAR cross-validation demonstrate high consistency with SAR observations, confirming the model's generalization in complex terrain.

Furthermore, application on Shackleton crater using ShadowCam data demonstrate the framework's capability in PSRs. The resulting high-resolution DTMs provide valuable topographic support for future deep-space exploration missions, including landing site selection, path planning, and water-ice resource detection. In addition, high-resolution DTM generation under constraints of low-illumination conditions remains to be further improved in future research.

Acknowledgment

The Mini-RF data, LOLA DEM data, NAC data, ShadowCam data are available on NASA's Planetary Data System (http://pds-geosciences.wustl.edu). NAC DTM data are from https://wms.lroc.asu.edu/lroc/rdr_product_select. The authors would thank the date produce team for their great work in producing and maintaining the data used in this study.

References

[1] J. Liu, X. Zeng, C. Li *et al.*, "Landing Site Selection and Overview of China's Lunar Landing Missions," *Space Science Reviews,* vol. 217, no. 1, 2020.

[2] K. Di, B. Liu, Z. Liu *et al.*, "Review and prospect of lunar mapping using remote sensing data," *National Remote Sensing Bulletin,* vol. 20, no. 5, pp. 1230-1242, 2016.

[3] W. Bo, G. Jian, Z. Yunsheng *et al.*, "Integration of Chang'E-1 Imagery and Laser Altimeter Data for Precision Lunar Topographic Modeling," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 49, no. 12, pp. 4889-4903, 2011.

[4] B. Wu, H. Hu, and J. Guo, "Integration of Chang'E-2 imagery and LRO laser altimeter data with a combined block adjustment for precision lunar topographic modeling," *Earth and Planetary Science Letters,* vol. 391, pp. 1-15, 2014.

[5] B. Wu, W. C. Liu, A. Grumpe *et al.*, "SHAPE AND ALBEDO FROM SHADING (SAfS) FOR PIXEL-LEVEL DEM GENERATION FROM MONOCULAR IMAGES CONSTRAINED BY LOW-RESOLUTION DEM," *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences,* vol. XLI-B4, pp. 521-527, 2016.

[6] W. Chung Liu, B. Wu, and C. Wöhler, "Effects of illumination differences on photometric stereo shape-and-albedo-from-shading for precision lunar surface reconstruction," *ISPRS Journal of Photogrammetry and Remote Sensing,* vol. 136, pp. 58-72, 2018/02/01/, 2018.

[7] A. Grumpe, F. Belkhir, and C. Wöhler, "Construction of lunar DEMs based on reflectance modelling," *Advances in Space Research,* vol. 53, no. 12, pp. 1735-1767, Jun 15, 2014.

[8] H. Chen, X. Y. Hu, P. Glaser *et al.*, "CNN-Based Large Area Pixel-Resolution Topography Retrieval From Single-View LROC NAC Images Constrained With SLDEM," *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* vol. 15, pp. 9398-9416, 2022.

[9] O. Alexandrov, and R. A. Beyer, "Multiview Shape-From-Shading for Planetary Images," *Earth and Space Science,* vol. 5, no. 10, pp. 652-666, Oct, 2018.

[10] H. Chen, P. Gläser, X. Hu *et al.*, "ELunarDTMNet: Efficient Reconstruction of High-Resolution Lunar DTM From Single-View Orbiter Images," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 62, pp. 1-20, 2024.

[11] Y. Liu, Y. X. Wang, K. C. Di *et al.*, "A Generative Adversarial Network for Pixel-Scale Lunar DEM Generation from High-Resolution Monocular Imagery and Low-Resolution DEM," *Remote Sensing,* vol. 14, no. 21, Nov, 2022.

[12] T. H. Chen, Y. X. Wang, J. Nan *et al.*, "A Generative Adversarial Network for Pixel-Scale Lunar DEM Generation from Single High-Resolution Image and Low-Resolution DEM Based on Terrain Self-Similarity Constraint," *Remote Sensing,* vol. 17, no. 17, Sep 5, 2025.

[13] Y. Tao, J. P. Muller, S. J. Conway *et al.*, "Large Area High-Resolution 3D Mapping of the Von Karman Crater: Landing Site for the Chang'E-4 Lander and Yutu-2 Rover," *Remote Sensing,* vol. 15, no. 10, May 18, 2023.

[14] H. Chen, X. Y. Hu, and J. Oberst, "Pixel-Resolution Dtm Generation for the Lunar Surface Based on a Combined Deep Learning and Shape-from-Shading (Sfs) Approach," *Xxiv Isprs Congress: Imaging Today, Foreseeing Tomorrow, Commission Iii,* vol. 5-3, pp. 511-516, 2022.

[15] G. E. Hinton, "Products of experts," *Ninth International Conference on Artificial Neural Networks (Icann99), Vols 1 and 2*, no. 470, pp. 1-6, 1999.

[16] L. Bruzzone, and D. F. Prieto, "Automatic analysis of the difference image for unsupervised change detection," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 38, no. 3, pp. 1171-1182, 2000.

[17] N. Liu, and Y. Q. Jin, "Pol-SAR Image Simulation of the Lunar Surface With Data Analysis of Chandrayaan-2 and Mini-RF," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* vol. 16, pp. 10301-10310, 2023.

[18] M. S. Robinson, S. M. Brylow, M. Tschimmel *et al.*, "Lunar Reconnaissance Orbiter Camera (LROC) Instrument Overview," *Space Science Reviews,* vol. 150, no. 1, pp. 81-124, 2010/01/01, 2010.

[19] M. K. Barker, E. Mazarico, G. A. Neumann *et al.*, "A new lunar digital elevation model from the Lunar Orbiter Laser Altimeter and SELENE Terrain Camera," *Icarus,* vol. 273, pp. 346-355, 2016/07/15/, 2016.

[20] M. R. Henriksen, M. R. Manheim, K. N. Burns *et al.*, "Extracting accurate and precise topography from LROC narrow angle camera stereo observations," *Icarus,* vol. 283, pp. 122-137, 2017/02/01/, 2017.

[21] B. Li, X. Q. Wang, J. Zhang *et al.*, "Lunar textural analysis based on WAC-derived kilometer-scale roughness and entropy maps," *Planetary and Space Science,* vol. 125, pp. 62-71, Jun, 2016.

[22] Y. Gao, B. Wu, J. Hu *et al.*, "Centimeter-Scale Roughness Characteristics of Lunar Geological Structures and Permanently Shadowed Regions: Insights From Mini-RF SAR Data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* vol. 19, pp. 7726-7741, 2026.

[23] Z. Yin, N. Liu, and Y.-Q. Jin, "Simulation of the temperatures in the permanently shadowed region of the Moon's south pole and data validation," *Icarus,* vol. 411, pp. 115917, 2024/03/15/, 2024.

[24] F. Zhao, P. P. Lu, T. Y. Meng *et al.*, "Selection of Landing Sites for the Chang'E-7 Mission Using Multi-Source Remote Sensing Data," *Remote Sensing,* vol. 17, no. 7, Mar 21, 2025.

[25] J. Qian, Z. Ye, S. Qiu *et al.*, "Secondary scattering shape from shading for precise terrain reconstruction of the lunar permanently shadowed regions," *Icarus,* vol. 431, pp. 116494, 2025/05/01/, 2025.

[26] R. Jia, B. Wu, and P. Mahanti, "Shape-from-shading with secondary illumination for 3D reconstruction of permanently shadowed regions at

the lunar south pole," *Acta Astronautica,* vol. 238, pp. 889-903, 2026/01/01/, 2026.

[27] C. Chen, Z. Ye, Y. Xu *et al.*, "Large-Scale Block Bundle Adjustment of LROC NAC Images for Lunar South Pole Mapping Based on Topographic Constraint," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* vol. 17, pp. 2731-2746, 2024.

[28] T. Tao, Z. Ye, R. Huang *et al.*, "Lunar Terrain Modeling From Sparse Surface Points Using Signed Distance Fields and Dynamic Planar Features," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 63, pp. 1-15, 2025.

[29] M. K. Barker, E. Mazarico, G. A. Neumann *et al.*, "Improved LOLA elevation maps for south pole landing sites: Error estimates and their impact on illumination conditions," *Planetary and Space Science,* vol. 203, Sep, 2021.