Review on Deep Learning Techniques in Planetary Topographic Modeling

Hao Chen [1], Jürgen Oberst [1], Philipp Gläser [1], Konrad Willner [2]

[1] Institute of Geodesy and Geoinformation Science, Technische Universität Berlin, Berlin 10553, Germany - hao.chen.2@campus.tu-berlin.de, juergen.oberst@tu-berlin.de, philipp.glaeser@tu-berlin.de
[2] Institute of Planetary Research, German Aerospace Center (DLR), Berlin 12489, Germany - konrad.willner@dlr.de

Keywords: Deep learning, Orbital imagery, Planetary topographic modeling, Small bodies, The Moon.

Abstract

Topographic modeling using orbital imagery is a cornerstone of planetary photogrammetry and remote sensing, underpinning scientific exploration and analysis. While classical methods like stereo-photogrammetry (SPG) and (stereo)-photoclinometry (SPC) have long been developed, deep learning (DL) techniques have recently emerged as powerful alternatives, advancing rapidly in planetary topographic applications. This study briefly reviews the evolution of DL methods, contrasting their innovative approaches with the principles of traditional SPG and SPC techniques. We assess the efficacy of two representative DL models in reconstructing high-resolution topography for a large planetary body (the Moon) and a small asteroid (Itokawa), respectively. Our findings reveal that these DL methods successfully recover detailed terrain surfaces, even with limited input imagery, and produce results consistent with SPG- and SPC-derived models. These outcomes underscore the transformative potential of DL for efficient, robust topographic modeling across diverse planetary scales.

1. Introduction

While previously the primary techniques for planetary topographic modeling from orbital imagery were stereo-photogrammetry (SPG, Preusker et al., 2015; Henriksen et al., 2017) and (stereo)-photoclinometry (SPC, Gaskell et al., 2008; Alexandrov & Beyer, 2018), deep learning (DL) techniques have advanced rapidly in recent years (Chen et al., 2021; Liu et al., 2022). Since high-resolution images are acquired using different modes for large and small planetary bodies, tailored approaches are developed to accommodate their respective topographic modeling requirements, supporting both exploration missions and scientific research (Chen et al., 2022; Chen et al., 2024a).

For large planetary bodies, such as the Moon and Mars, high-resolution stereo images, e.g, by the Lunar Reconnaissance Orbiter (LRO) Narrow Angle Camera (NAC) or the High-Resolution Imaging Science Experiment (HiRISE), are available only for local areas due to the lack of build-in stereo imaging capability (McEwen et al., 2007; Robinson et al., 2010). DL methods typically use a single-view high-resolution image along with a corresponding coarse-resolution topographic model to predict a high-resolution model, potentially with large-area coverage (Liu et al., 2022; Tao et al., 2023). In contrast to the photoclinometry technique (Liu et al., 2021), which relies on illumination information and on the reflectance model for terrain reconstruction, the DL terrain estimation is performed by implicitly "learning" the relationship between 2D imagery and 3D topographic features through neural networks (Chen et al., 2021). Compared to the photoclinometry method, DL is robust when using images with different solar azimuth angles as input (Chen et al., 2022; Chen et al., 2024b). Besides, thanks to its high processing speed, it shows great promise for large-area terrain reconstruction (Tao et al., 2023). Recently, an efficient DL-based single-view method was proposed, demonstrating that it can achieve high-quality lunar topographic reconstruction (Chen et al., 2024b).

For small bodies such as Near-Earth Asteroids, shape modeling is typically carried out on a global scale using multi-view images acquired from spacecraft flybys or orbital observations (Gaskell et al., 2008; Preusker et al., 2015). Recently, the Neural Radiance Field (NeRF)-based Neural Implicit Method (NIM) has been introduced for modeling (Chen et al., 2024a; Chen et al., 2024c). In contrast to the SPG method, which uses discrete point clouds to represent the surface, the neural implicit method employs an implicit function to represent contiguous shape models (Chen et al., 2024c). SPG and SPC methods typically involve multiple intermediate steps (Preusker et al., 2015; Palmer et al., 2022). For example, SPC uses SPG to determine the 3D positions of maplets, then applies photoclinometry to estimate slopes and reconstruct the topography. Finally, it merges and optimizes the maplets to create a complete 3D surface model (Al Asad et al., 2021). In contrast, the neural implicit method can directly derive a complete model in an end-to-end manner. While its development in the field of small body shape modeling is still relatively recent, it has already demonstrated competitive results in accurate shape modeling (Chen et al., 2024a; Chen et al., 2024c).

In this paper, we will present two DL-based methods specifically designed for planetary topographic modeling – one for large planetary bodies, demonstrated with the Moon, and another for small bodies, exemplified by Itokawa. Through these case studies, we will evaluate the performance of DL methods in planetary topographic reconstruction.

2. Method

2.1 Single-view DL topographic modeling for the Moon

We proposed ELunarDTMNet, an efficient deep learning network designed for high-resolution lunar topographic modeling (Chen et al., 2024b). Built upon a hierarchical Transformer architecture, this network excels at capturing multi-scale terrain features from the single input image, enabling it to represent both fine details and broader topographic patterns effectively. To enhance reconstruction performance, ELunarDTMNet integrates a coarse-resolution topographic model as an initial elevation constraint, providing a foundational guide that improves the accuracy of the output. Tailored for high-resolution terrain reconstruction, ELunarDTMNet produces a pixel-level grid model to capture terrain details.

Elevation normalization is critical for ensuring robust network convergence. Narrow normalization ranges often diminish terrain contrast, flattening elevation undulations and compromising detail fidelity. To address this, we proposed an elevation-aware normalization strategy that leverages elevation statistics to determine an optimal normalization range. This approach better preserves elevation disparities across diverse lunar landscapes, significantly improving the faithfulness and richness of the reconstructed topography. Furthermore, given the varied elevation distributions inherent in lunar terrains, we introduced a mean-normalized loss function. This function penalizes vertical discrepancies based on a uniform mean scale, rather than fluctuating statistical measures, steering the network toward learning robust, generalizable features unaffected by topographic variability.

In practice, ELunarDTMNet typically generates small-scale topographic models from compact, small-sized input images, making it computationally efficient for predictions. To produce a comprehensive large-scale topographic mosaic, the network seamlessly merges adjacent small-scale models, stitching them into a unified, expansive representation that retains consistency and detail across broader regions.

2.2 Multi-view NIM for small body shaping modeling

The NIM requires images with precise interior and exterior parameters as its initial input, a critical step for accurately reconstructing 3D geometry from 2D observations (Mildenhall et al., 2021). Then, it generates a set of camera rays by sampling pixels across the image plane and computing the rays that originate at the camera center and pass through these pixels. From these derived rays, the system samples 3D points along each ray’s trajectory, capturing spatial information essential for modeling small body surfaces. The NIM leverages these points, paired with their corresponding viewing orientations, as input to the network, enabling it to infer both geometric and appearance properties of the scene. To implicitly represent the target surface, the signed distance function (SDF) is adopted, offering a continuous and differentiable representation of the body’s topography (Wang et al., 2021). The NIM employs multi-layer perceptrons (MLPs) as its network architecture to learn the 3D scene representation, encompassing the signed distance field for shape and the radiance field for view-dependent surface appearance, both of which are jointly optimized during training. This dual field allows the model to synthesize high-fidelity renderings and reconstruct detailed geometry.

To enhance its capability to capture fine-scale features, NIM incorporates various strategies. Positional encoding, for instance, transforms the 3D coordinates of input points into a higher-dimensional space, enabling the MLPs to better resolve high-frequency details often lost in low-dimensional inputs (Mildenhall et al., 2021). Additionally, to tackle the challenge of modeling small bodies with faint terrain variations, we proposed a mask-based sampling strategy (Chen et al., 2024a). This technique prioritizes ray sampling in regions of interest, guiding the network to focus computational resources on learning intricate topographic details. Moreover, to enrich the network’s understanding of local surface structure, neighboring points of the sampled 3D points are integrated into the input, providing contextual cues about spatial relationships and surface continuity (Chen et al., 2024a). This local context is particularly valuable for small body shape modeling, where irregular geometries and limited observational data pose significant challenges. By combining these enhancements, NIM achieves robust performance in reconstructing the complex, heterogeneous surfaces of small celestial bodies.

3. Datasets

This section describes the datasets required by ELunarDTMNet for lunar topographic modeling at the Chang’E-3 landing site and by NIM for shape modeling of the asteroid Itokawa.

3.1 Dataset for the Chang’E-3 landing site

ELunarDTMNet takes a high-resolution image and a coarse-resolution topographic model as inputs. In this study, we evaluate its performance by revisiting the Chang’E-3 landing site, in the northern Mare Imbrium. The NAC image (M1144936321) with a resolution of 1.6 m is used as the input image, while the SLDEM2015 model (Barker et al., 2016) at a resolution of 512 pixels per degree serves as the coarse-resolution model, as shown in Figure 1. It covers an area of 1120 km^2 and exhibits an elevation range of up to 380 meters, encompassing a variety of topographic features, including craters of diverse sizes and subtle mare undulations. The 5 m resolution SPG-derived topographic model (Henriksen et al., 2017) is used as the reference to assess the accuracy of our reconstructed 1.6 m resolution model.

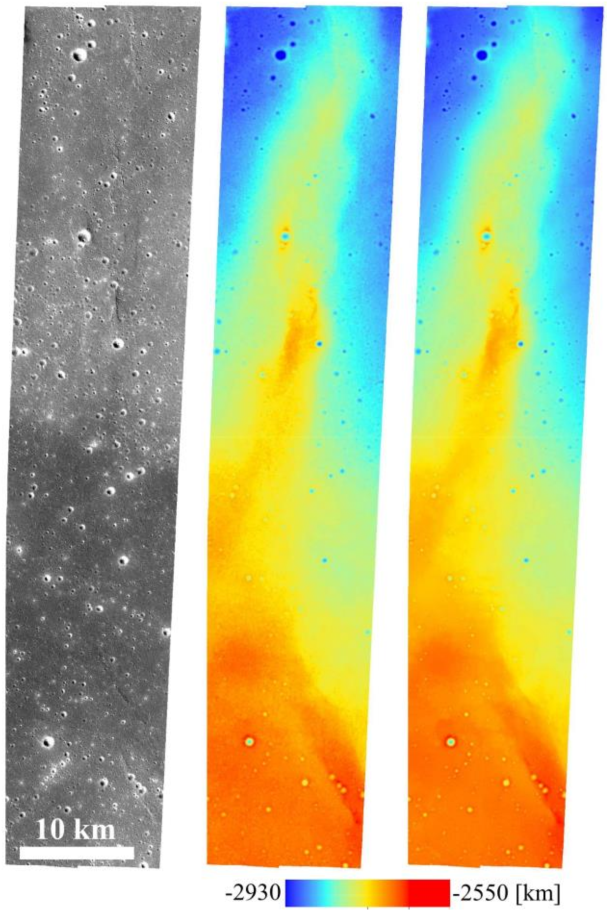

Figure 1. Input NAC image (left panel) and SLDEM2015 topographic model (middle panel) used for reconstructing the Chang’E-3 landing site topographic model. The right panel shows the SPG-derived reference model.

3.2 Dataset for Itokawa

The Hayabusa spacecraft, launched by JAXA in 2003, was the first mission to successfully return samples from an asteroid, specifically 25143 Itokawa, a near-Earth object with a highly irregular, bi-lobed shape. One of its key instruments, the Asteroid Multi-band Imaging Camera (AMICA) (Ishiguro et al., 2010), was designed to capture high-resolution images of the asteroid’s surface in multiple spectral bands, aiding in geomorphological and compositional analyses.

For this study, a subset of 52 images taken during the spacecraft’s proximity phase—at distances ranging from 7.5 to 8 km from the body—was selected for shape reconstruction. These images provide multi-angle views essential for accurately modeling the asteroid’s complex topography. Figure 2 presents some example images, showcasing Itokawa from four different perspectives.

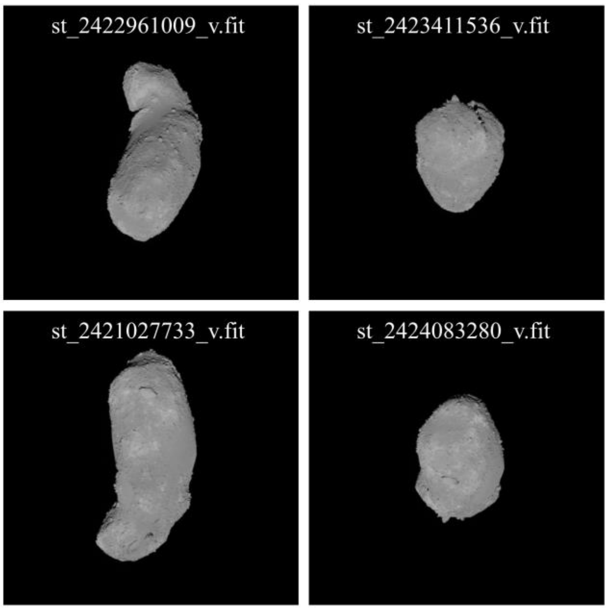

Figure 2. Example AMICA images used for the shape reconstruction of Itokawa.

4. Results

4.1 Reconstruction results for the Chang’E-3 landing site

Figure 3 presents the reconstructed topographic model of the Chang’E-3 landing site, achieved at a grid resolution of 1.6 m using ELunarDTMNet. The elevation trends in this model align closely with those of the SPG-derived topographic model presented in Figure 1, demonstrating the reliability of our reconstruction. Large-scale features, such as craters with diameters exceeding several hundred meters, are distinctly visible and well-defined in both the topographic model and its corresponding hill-shaded map, reflecting the model’s ability to capture prominent lunar structures. Additionally, the hill-shaded map clearly delineates the wrinkle ridge—a irregular yet significant tectonic feature—whose morphology corresponds closely with the NAC image shown in Figure 1, further validating the model’s fidelity to observed surface details.

The absolute elevation difference map, computed relative to the 5 m resolution SPG reference model, confirms the accuracy and reliability of our reconstructed topography. Approximately 48% of the reconstructed area exhibits elevation differences of less than 2 m, while 79% shows differences below 4 m, indicating a high degree of agreement with the reference. These metrics highlight ELunarDTMNet’s precision in reconstructing topography, particularly in a region characterized by diverse crater sizes and irregular elevation variations, offering valuable insights for geological analysis and future mission planning.

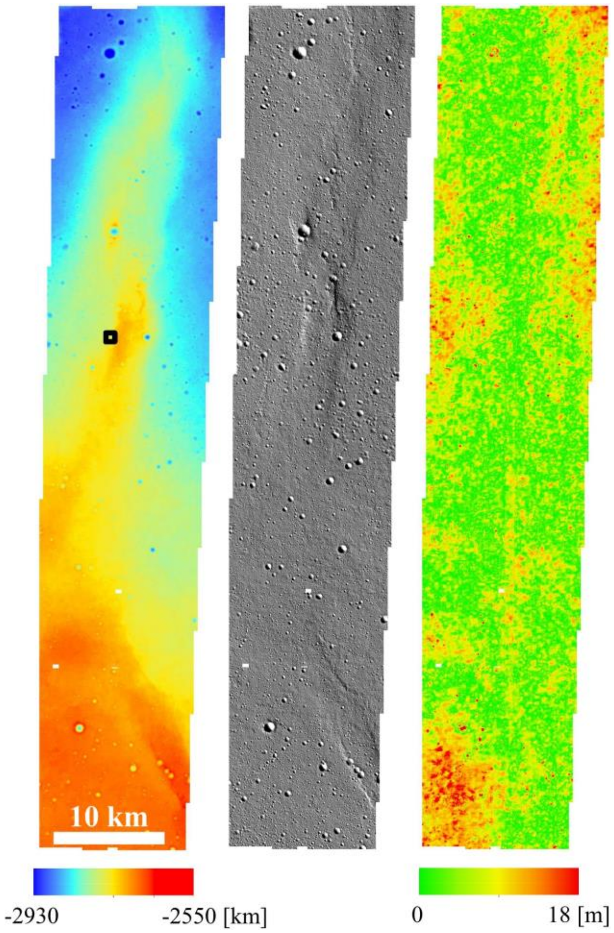

Figure 3. Reconstructed topographic model at 1.6 m resolution generated by ELunarDTMNet (left panel). Hill-shaded map derived from the reconstructed model (middle panel). Absolute elevation difference between the reconstructed model and the 5 m resolution SPG reference model (right panel). The black box indicates the location of the local area in Figure 4.

To further assess ELunarDTMNet’s capability in retrieving fine-scale terrain features, Figure 4 presents a detailed comparison within a local area, highlighting small-scale craters and an irregularly shaped wrinkle ridge as depicted in hill-shaded maps. Our 1.6 m resolution model excels at reconstructing these subtle features, accurately capturing the intricate outlines of small craters and the complex morphology of the wrinkle ridge. This high-fidelity underscores ELunarDTMNet’s effectiveness in resolving fine topographic details critical for geological interpretation. In contrast, the 5 m resolution SPG-derived model, while able to delineate the broader form of the wrinkle ridge, fails to preserve its detailed texture and omits the small craters, losing critical information about surface roughness and micro-topography. The SLDEM2015 model, with its coarser resolution of approximately 50 m, proves even less capable, entirely missing these fine-scale features in the local area. This resolution limitation renders it inadequate for capturing the nuanced terrain variations, such as subtle crater rims and ridge undulations, that are essential for understanding the geological evolution and surface processes in this region. The comparison in Figure 4 thus demonstrates ELunarDTMNet’s superior performance in enhancing topographic detail over existing models, offering significant advantages for high-resolution lunar studies.

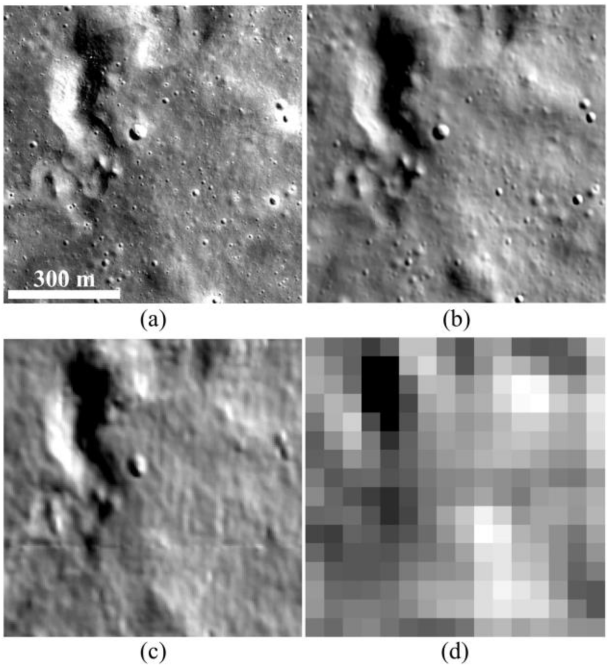

Figure 4. Comparison of fine-scale terrain reconstruction in a local area: (a) 1.6 m resolution NAC image, (b) 1.6 m resolution hill-shaded map derived from ELunarDTMNet, (c) 5 m resolution hill-shaded map from the SPG model, and (d) 50 m resolution hill-shaded map from SLDEM2015.

4.2 Reconstruction results for Itokawa

Figure 5 displays two distinct perspectives of the Itokawa shape model, reconstructed using the NIM method (Chen et al., 2024a) and the well-established SPC method (Gaskell et al., 2008). The comparison illustrates that the NIM method yields a shape model aligned with the SPC-derived model, affirming its robustness in capturing the asteroid’s irregular, rubble-pile geometry—a hallmark of Itokawa’s structure as observed by the Hayabusa mission. Remarkably, the NIM method achieves this level of consistency while utilizing a mere 6% of the image data required by the SPC method, underscoring its exceptional efficiency in scenarios where extensive imaging coverage is limited, such as during rapid flybys or resource-constrained missions. Furthermore, it is significant that the NIM method relies on input images of substantially lower resolution compared to some high-resolution images employed by the SPC method.

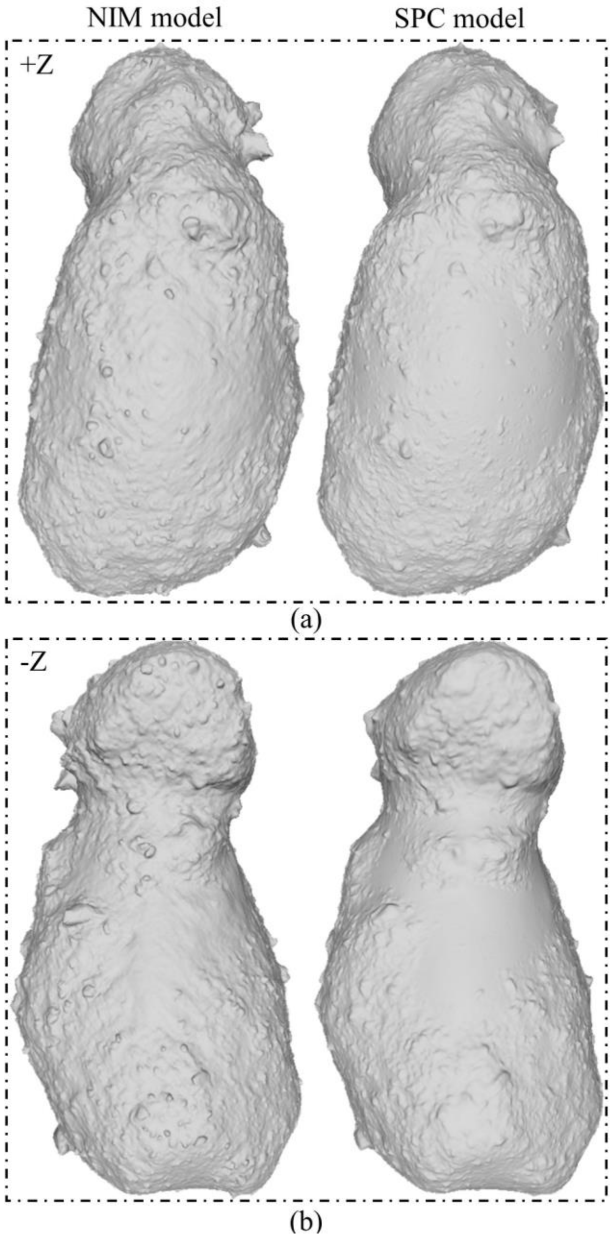

Figure 5. Two-views of the Itokawa shape models derived from the NIM method and the SPC method, revised from Chen et al., (2024a).

Furthermore, the NIM enhances its utility by generating rendered images from the learned 3D scene, providing a powerful means to validate its accuracy in reconstructing terrain features. Figure 6 indicates two real AMICA images from the Hayabusa mission with their corresponding synthetic renderings produced by the NIM (Chen et al., 2024b). The comparison reveals a striking consistency between the rendered and real images, with the NIM faithfully replicating intricate surface details—such as Itokawa’s boulder-strewn regolith and subtle topographic variations—demonstrating its proficiency in effectively learning and representing the asteroid’s complex surface geometry from limited input imagery. This high degree of visual fidelity not only confirms the NIM’s robustness in modeling 3D scenes but also highlights its potential for applications like virtual flyovers or surface analysis, where accurate rendering of fine-scale features is critical.

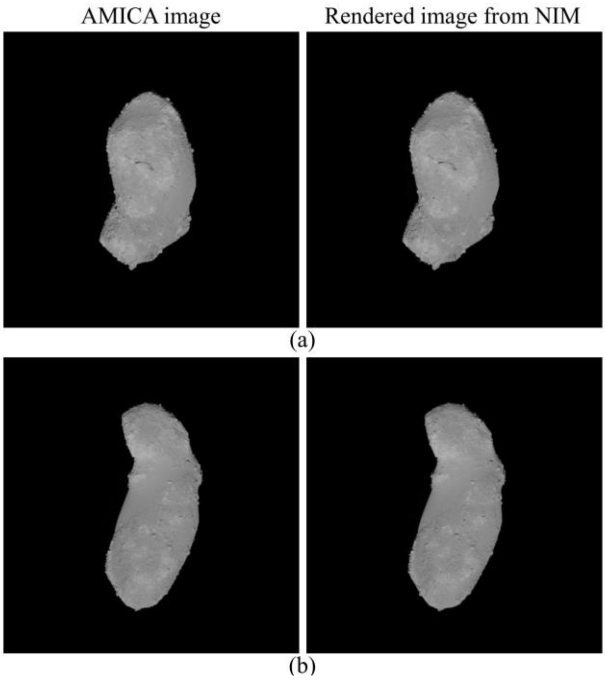

Figure 6. Comparison of the real AMICA images and the rendered images from the NIM method. (a) st_2421207934_v.fit, (b) st_2423141242_v.fit.

5. Conclusion

In this paper, we explored the application of DL techniques in planetary topographic modeling, addressing both large celestial bodies, such as planets, and smaller objects, including asteroids. Leveraging two representative DL approaches, we successfully reconstructed high-resolution topographic models for the Moon and the asteroid Itokawa, respectively. The results demonstrate that these DL methods offer significant promise for generating detailed, accurate topographic representations, effectively capturing fine-scale features like lunar craters and Itokawa's irregular surface geometry. These findings highlight the potential of DL to model terrain surfaces across diverse planetary scales, providing efficient and robust solutions even with limited input data. Looking forward, this work lays a foundation for further advancements in DL-based planetary science, potentially enhancing mission planning, surface analysis, and our understanding of extraterrestrial terrains.

Acknowledgements

The authors extend their gratitude to the LRO and Hayabusa missions for providing invaluable imagery, which has significantly advanced our understanding of the Moon and Itokawa.

References

Al Asad, M.M., Philpott, L.C., Johnson, C.L., Barnouin, O.S., Palmer, E., Weirich, J.R., Daly, M.G., Perry, M.E., Gaskell, R., Bierhaus, E.B. and Seabrook, J.A., 2021. Validation of stereophotoclinometric shape models of asteroid (101955) Bennu during the OSIRIS-REx mission. The Planetary Science Journal, 2(2), p.82.

Alexandrov, O. and Beyer, R.A., 2018. Multiview shape-from-shading for planetary images. Earth and Space Science, 5(10), pp.652-666.

Barker, M.K., Mazarico, E., Neumann, G.A., Zuber, M.T., Haruyama, J. and Smith, D.E., 2016. A new lunar digital elevation model from the Lunar Orbiter Laser Altimeter and SELENE Terrain Camera. Icarus, 273, pp.346-355.

Chen, H., Hu, X., Gläser, P., Xiao, H., Ye, Z., Zhang, H., Tong, X. and Oberst, J., 2022. 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, 15, pp.9398-9416.

Chen, H., Hu, X., Willner, K., Ye, Z., Damme, F., Gläser, P., Zheng, Y., Tong, X., Hußmann, H. and Oberst, J., 2024a. Neural implicit shape modeling for small planetary bodies from multi-view images using a mask-based classification sampling strategy. ISPRS Journal of Photogrammetry and Remote Sensing, 212, pp.122-145.

Chen, H., Gläser, P., Hu, X., Willner, K., Zheng, Y., Damme, F., Bruzzone, L. and Oberst, J., 2024b. ELunarDTMNet: Efficient reconstruction of high-resolution lunar DTM from single-view orbiter images. IEEE Transactions on Geoscience and Remote Sensing, 62, p. 4601820.

Chen, S., Wu, B., Li, H., Li, Z. and Liu, Y., 2024c. Asteroid-NeRF: A deep-learning method for 3D surface reconstruction of asteroids. Astronomy & Astrophysics, 687, p.A278.

Chen, Z., Wu, B. and Liu, W.C., 2021. Mars3DNet: CNN-based high-resolution 3D reconstruction of the Martian surface from single images. Remote Sensing, 13(5), p.839.

Gaskell, R.W., Barnouin‐Jha, O.S., Scheeres, D.J., Konopliv, A.S., Mukai, T., Abe, S., Saito, J., Ishiguro, M., Kubota, T., Hashimoto, T. and Kawaguchi, J.I., 2008. Characterizing and navigating small bodies with imaging data. Meteoritics & Planetary Science, 43(6), pp.1049-1061.

Henriksen, M.R., Manheim, M.R., Burns, K.N., Seymour, P., Speyerer, E.J., Deran, A., Boyd, A.K., Howington-Kraus, E., Rosiek, M.R., Archinal, B.A. and Robinson, M.S., 2017. Extracting accurate and precise topography from LROC narrow angle camera stereo observations. Icarus, 283, pp.122-137.

Ishiguro, M., Nakamura, R., Tholen, D.J., Hirata, N., Demura, H., Nemoto, E., Nakamura, A.M., Higuchi, Y., Sogame, A., Yamamoto, A. and Kitazato, K., 2010. The Hayabusa spacecraft asteroid multi-band imaging camera (AMICA). Icarus, 207(2), pp.714-731.

Liu, W.C. and Wu, B., 2021. Influence of solar incidence angle on single-image photoclinometry for precision lunar topographic mapping. ISPRS Journal of Photogrammetry and Remote Sensing, 182, pp.208-227.

Liu, Y., Wang, Y., Di, K., Peng, M., Wan, W. and Liu, Z., 2022. A generative adversarial network for pixel-scale lunar DEM generation from high-resolution monocular imagery and low-resolution DEM. Remote Sensing, 14(21), p.5420.

McEwen, A.S., Eliason, E.M., Bergstrom, J.W., Bridges, N.T., Hansen, C.J., Delamere, W.A., Grant, J.A., Gulick, V.C., Herkenhoff, K.E., Keszthelyi, L. and Kirk, R.L., 2007. Mars reconnaissance orbiter's high resolution imaging science experiment (HiRISE). Journal of Geophysical Research: Planets, 112(E5).

Mildenhall, B., Srinivasan, P.P., Tancik, M., Barron, J.T., Ramamoorthi, R. and Ng, R., 2021. Nerf: Representing scenes as neural radiance fields for view synthesis. *Communications of the ACM*, *65*(1), pp.99-106.

Palmer, E.E., Gaskell, R., Daly, M.G., Barnouin, O.S., Adam, C.D. and Lauretta, D.S., 2022. Practical stereophotoclinometry for modeling shape and topography on planetary missions. The Planetary Science Journal, 3(5), p.102.

Preusker, F., Scholten, F., Matz, K.D., Roatsch, T., Willner, K., Hviid, S.F., Knollenberg, J., Jorda, L., Gutiérrez, P.J., Kührt, E. and Mottola, S., 2015. Shape model, reference system definition, and cartographic mapping standards for comet 67P/Churyumov-Gerasimenko–Stereo-photogrammetric analysis of Rosetta/OSIRIS image data. Astronomy & Astrophysics, 583, p.A33.

Robinson, M.S., Brylow, S.M., Tschimmel, M.E., Humm, D., Lawrence, S.J., Thomas, P.C., Denevi, B.W., Bowman-Cisneros, E., Zerr, J., Ravine, M.A. and Caplinger, M.A., 2010. Lunar reconnaissance orbiter camera (LROC) instrument overview. Space science reviews, 150, pp.81-124.

Tao, Y., Muller, J.P., Conway, S.J., Xiong, S., Walter, S.H. and Liu, B., 2023. Large area high-resolution 3D mapping of the Von Kármán Crater: Landing site for the Chang’E-4 lander and Yutu-2 rover. Remote Sensing, 15(10), p.2643.

Wang, P., Liu, L., Liu, Y., Theobalt, C., Komura, T. and Wang, W., 2021, December. NeuS: learning neural implicit surfaces by volume rendering for multi-view reconstruction. In Proceedings of the 35th International Conference on Neural Information Processing Systems (pp. 27171-27183).