

Lunar Multitype Geological Structure Recognition Based on Cross-View Constraints

Chenya Li, Gaofeng Shu^{ID}, *Member, IEEE*, Xin Lu, Jianhui Zhao^{ID}, and Ning Li^{ID}, *Member, IEEE*

Abstract—The complex and diverse geological structures on the lunar surface serve as a direct record of its long evolutionary history. Comprehensively identifying and classifying these geological structures can not only deepen our understanding of lunar evolution but also support the planning of future lunar exploration missions and guide the detection of lunar energy resources. However, the diverse and intricate morphologies of lunar geological structures, coupled with certain similarities between different formations, make their automatic identification and classification a significant challenge. To address this issue, we propose a multitype lunar geological structure recognition network based on cross-view constraints, which mines differentiated feature information to enhance target identification and discrimination capabilities. This network effectively extracts heterogeneous and complementary features through cross-view feature extraction constraints. By employing a discrepancy-weighted loss function, the network focuses on regions where discrepancies arise in the recognition results across multiple views, thereby enhancing attention to divergent areas and learning feature representations in complex scenarios. Additionally, multiscale contextual information aggregation combines contextual features from different receptive fields, leveraging surrounding terrain to enhance discriminative information for lunar geological structures. Experimental results demonstrate that the proposed method exhibits significant superiority in the task of identifying and classifying multitype lunar geological structures, with a 1.9% improvement in mean intersection over union (mIoU).

Index Terms—Cross-view constraints, feature fusion, lunar geological structure identification, semantic segmentation.

I. INTRODUCTION

LUNAR geological structures can be broadly categorized into linear and circular formations based on their geometric characteristics [1]. Linear structures primarily include lunar rilles and wrinkle ridges, while circular structures are mainly composed of impact craters. As a distinct type of linear structure, wrinkle ridges are closely associated with compressional stresses within the Moon's interior [2]. Studying them can reveal variations in lunar internal stress and the history of tectonic activity [3]. The formation of lunar rilles, on the other hand, is linked to early volcanic activity or extensional stresses inside the Moon. Research on rilles can help trace early magmatic processes and geological movements while also guiding the exploration of lunar energy resources [4].

Received 28 July 2025; revised 10 September 2025; accepted 24 September 2025. Date of publication 29 September 2025; date of current version 9 October 2025. (Corresponding author: Gaofeng Shu.)

The authors are with the School of Computer and Information Engineering, Henan University, Kaifeng 475004, China, also with Henan Province Engineering Research Center of Spatial Information Processing, Kaifeng 475004, China, and also with Henan Key Laboratory of Big Data Analysis and Processing, Kaifeng 475004, China (e-mail: gaofeng.shu@henu.edu.cn).

Digital Object Identifier 10.1109/LGRS.2025.3615657

Impact craters record specific periods of bombardment, reflecting external impacts the Moon has endured throughout its history [5]. By comprehensively identifying and classifying these geological structures, we can systematically reconstruct the Moon's evolutionary trajectory from its formation to the present, providing critical and direct evidence for refining theories on solar system planetary evolution [6].

In previous studies, researchers have conducted systematic identification and classification of geological structures on the lunar surface. Harada et al. [7] employed a combination of edge detection and Hough transform based on digital elevation model (DEM) data to semiautomatically extract circular and linear structures. Jiang et al. [8] proposed an algorithm based on Chang'e-1 optical images. By extracting block-level image features and performing cluster analysis, they achieved the distinction between lunar highlands and maria. Nevertheless, it should be noted that the classification granularity of this approach was limited to image patches, leaving room for improvement in spatial resolution, while the classified types remained relatively simplistic. To enhance identification granularity, Wang et al. [9] made technical improvements by proposing an unsupervised classification method based on iterative self-organizing data analysis. While this study improved identification granularity and spatial resolution, the classification remained relatively coarse. To further refine geological structure classification, Ghosh et al. [10] proposed a supervised machine learning-based segmentation method for automatic landform recognition on Mars, aiming to transform remote sensing topographic data from orbital satellites into semantically labeled terrain maps. The study systematically evaluated two segmentation methods and three classification algorithms and was validated across six test sites on Mars.

However, these studies share a common area for improvement: none have fully explored and utilized the deep-level feature information of landforms, leading to a lack of discriminative information during the recognition and classification process. In current research on multitype geological structure identification on the lunar surface, mainstream methods primarily rely on clustering segmentation and traditional machine learning techniques. On the one hand, these methods can achieve satisfactory recognition results for geological structures with high distinguishability, such as plains, highlands, and impact craters. However, for complex and morphologically similar structures like wrinkle ridges and lunar rilles, traditional approaches often suffer from high rates of misclassification and missed detection. On the other hand, researchers have developed various highly customized algorithmic for different application scenarios, incorporating specific feature

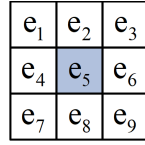


Fig. 1. Sliding window.

parameters for discriminative analysis. While these tailored methods can effectively locate geological structures in specific contexts, they still exhibit clear limitations in pixel-level recognition accuracy and algorithmic generalization capability. In contrast, deep learning methods demonstrate significant advantages. By mining deep semantic information such as contextual cues and multiscale features, they often achieve superior predictive performance and stronger generalization ability.

To achieve superior performance, this study employs a cross-view deep learning network to extract differentiated complementary information and contextual features, leveraging deep semantic representations, to provide richer discriminative information for multitype geological structure classification.

II. STUDY AREA AND DATA

In this section, this study constructs a comprehensive deep learning dataset for lunar geological structure using DEM data and aspect data for model training and evaluation.

A. Study Area

The lunar mare region contains abundant geological structures. Therefore, the study area of this research was specifically selected within the scope of the lunar mare. The dataset used covers the area from 90 °W to 45 °W in longitude and 0° to 60 °N in latitude. This region was chosen due to its rich distribution of geological structures as well as the availability of high-quality DEM and aspect data.

B. Data Source and Preprocessing

The DEM data utilized in this study are derived from a combination of LRO LOLA and the SELENE TC, with a resolution of 512 pixels/degree (59 m/pixel) [9]. The file IDs for these datasets are `sldem2015_512_00n_30n_270_315_float` and `sldem2015_512_30n_60n_270_315_float`. These data are archived in the planetary data system (PDS).

To capture the edges of lunar geological structures, we incorporated aspect data that are sensitive to variations in topographic slopes. The specific data processing procedure has been described in previous work [11]. In detail, the aspect data were generated from DEM data using the geospatial data abstraction library (GDAL) with a sliding window as illustrated in Fig. 1. The calculation process is as follows [12]:

$$\text{Slope}_x = \frac{(e_3 + 2e_6 + e_9) - (e_1 + 2e_4 + e_7)}{8} \quad (1)$$

$$\text{Slope}_y = \frac{(e_7 + 2e_8 + e_9) - (e_1 + 2e_2 + e_3)}{8} \quad (2)$$

$$\text{aspect} = \arctan\left(\frac{\text{Slope}_y}{\text{Slope}_x}\right). \quad (3)$$

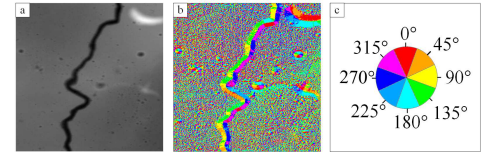


Fig. 2. Sample of DEM data and aspect data. (a) DEM data. (b) Aspect data. (c) Color-coding scheme for aspect.

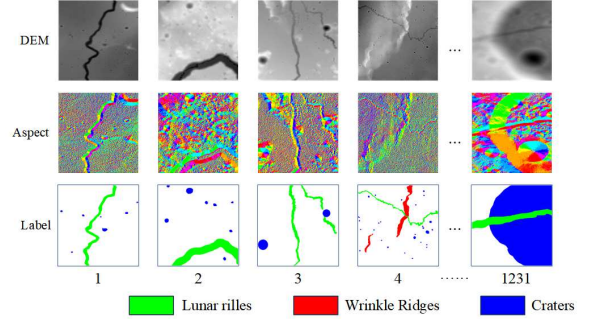


Fig. 3. Dataset samples.

The aspect data identify the downhill direction of maximum value change rate from each pixel to its neighboring pixels, providing critical information about the slope orientation of geological structures. The aspect values represent compass directions ranging from 0° to 360°. Fig. 2(a) displays the DEM data, Fig. 2(b) displays the aspect data visualization, while Fig. 2(c) presents a chromatic wheel diagram illustrating the color-coding scheme for these aspect values.

C. Dataset

The dataset was constructed by manually annotating lunar geological structures based on DEM data and aspect data, as shown in Fig. 3. To facilitate deep learning applications, the images were segmented into 1231 patches using a sliding window cropping method, with each patch measuring 512 × 512 pixels. Each data slice consists of three components: DEM data, aspect data, and visualized color-labeled data. Different categories of lunar geological structures are represented by distinct RGB colors for intuitive differentiation. The patches were then divided into training and testing sets at an 8:2 ratio.

III. METHOD

A. Overview

This method is designed to comprehensively explore and utilize differentiated deep semantic information. Liu et al. [13] demonstrated in semisupervised remote sensing image segmentation that multiview inference and cross-view mutual learning between weakly and strongly augmented samples can effectively alleviate model cognitive bias. Inspired by this work, we apply this methodology to lunar geological structure identification. To more effectively utilize differential information and contextual cues while enhancing the model's discrimination capability for various geological structure types, we propose a cross-view constrained lunar multitype geological structure recognition network (CVL-Net). The network architecture is shown in Fig. 4. Our approach employs two subviews with similar architectures but nonshared parameters for mutual learning. Each subview consists of an encoder

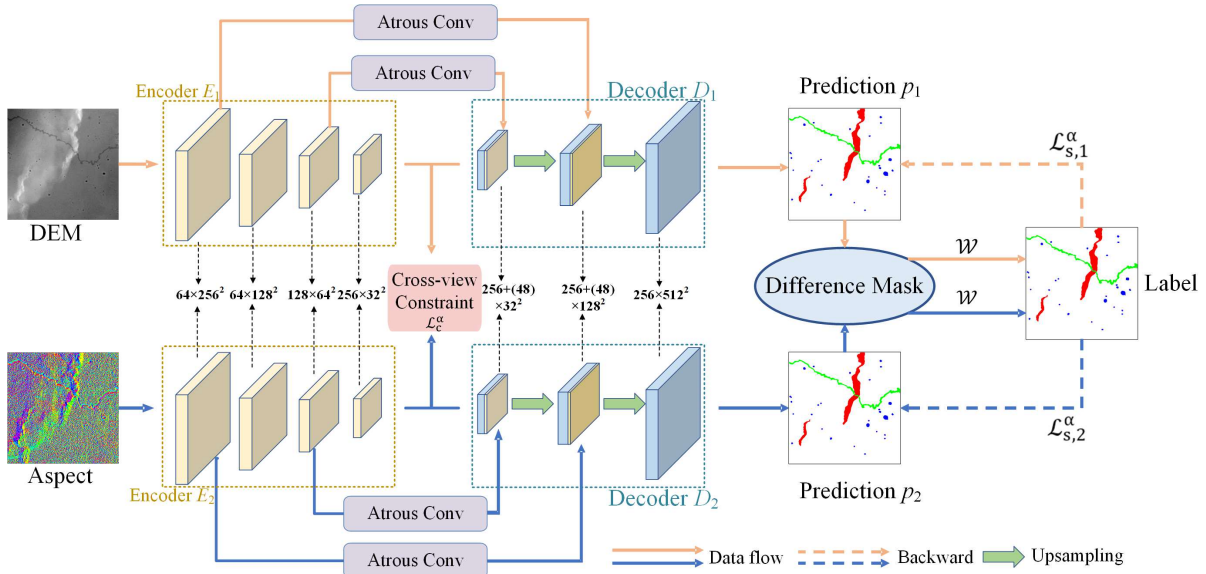


Fig. 4. Structure of CVL-Net.

E_i and a decoder D_i , where c represents the first and second subview, respectively.

B. Cross-View Constraints

The hypothesis that cross-view network models achieve better performance lies in the fact that the two views capture complementary information of the input data [13]. Considering that the goal of a cross-view network model is to extract distinct features from the two subviews, the outputs of the different encoders should differ; therefore, constraints are needed to ensure that the two views extract view-independent features. Aggregating information from two uncorrelated views leads to richer feature representations, enhancing the recognition of multitype lunar surface terrain structures.

In this study, the loss function is employed to constrain the features extracted from the two views, thereby promoting their divergence. Based on cosine similarity measurement, the feature map extracted from one view is treated as a negative sample for the feature map extracted from the other view. A strong constraint $\mathcal{L}_c^\alpha$ is imposed in the feature space to minimize the cosine similarity between the latent features f_i^α , forcing a significant difference between the features f_i of the i th view and the features f_{3-i} of the other view. This minimizes mutual information to avoid feature redundancy. The cross-view constraint is formulated as follows:

$$\mathcal{L}_{c,i} = 1 + \frac{f_i^\alpha \cdot f_{(3-i)}^{-\alpha}}{\|f_i^\alpha\| \times \|f_{(3-i)}^{-\alpha}\|}. \quad (4)$$

Here, $\|\cdot\|$ denotes the L2 norm (Euclidean norm) of a vector, $i \in 1, 2$ represents the first and second subview, respectively. $f_{(3-i)}^{-\alpha}$ is a gradient-free copy of the feature $f_{(3-i)}^\alpha$, preventing it from interfering with the parameter updates of the other view. The coefficient 1 ensures that the value of $\mathcal{L}_{c,i}$ remains nonnegative.

By adopting this loss function, the two views are encouraged to capture more diverse information from the input data. As the diversified features are effectively extracted, the model

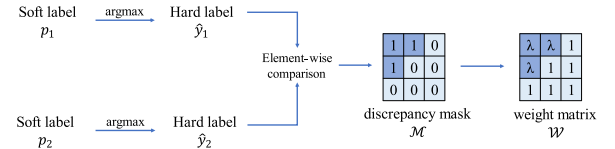


Fig. 5. Generation of the difference weight matrix.

gains a more comprehensive and in-depth understanding of the data, ultimately leading to significant improvements in overall segmentation and classification accuracy.

C. Divergence-Weighted Loss

In the loss function, applying uniform weights to all predicted locations ignores the divergence between the predictions of the two views. For easily identifiable regions, the predictions are typically consistent and reliable; whereas for challenging regions, the predictions from the two views may exhibit significant discrepancies due to feature complexity. Such discrepancies often reflect uncertainty in these regions or blind spots in the model's learning. Therefore, assigning higher weights to regions with prediction discrepancies is a more rational strategy.

Specifically, a discrepancy mask $\mathcal{M}$ and a weight matrix $\mathcal{W}$ are introduced for each image to increase the loss weighting on discrepant regions, compelling the model to focus more on areas where interview predictions diverge. The generation process of $\mathcal{M}$ and $\mathcal{W}$ is illustrated in Fig. 5.

First, transform the predicted probability of each pixel into per-image labels

$$\hat{y}_{i,n} = \text{argmax}(p_{i,n}). \quad (5)$$

Here, $p_{i,n}$ denotes the predicted soft labels. Perform an argmax operation on the predicted probability of each pixel. Then, the binary difference mask $\mathcal{M}$ for the n th input image across views is defined as follows:

$$\mathcal{M}(i, j) = \begin{cases} 1, & \text{if } \hat{y}_{1,n}(i, j) \neq \hat{y}_{2,n}(i, j) \\ 0, & \text{otherwise.} \end{cases} \quad (6)$$

Here, $\hat{y}_{1,n}(i, j)$ denotes the label of the (i, j) th pixel in the n th image obtained from the first view. $\mathcal{M}(i, j)$ indicates the presence of prediction discrepancy at the (i, j) th position. In this study, we assign higher weights to their cross-entropy losses. The weight matrix is defined as

$$\mathcal{W} = \lambda \cdot \mathcal{M} + (1 - \mathcal{M}). \quad (7)$$

Here, λ is a parameter greater than 1, which assigns larger weight values to discrepancy regions.

The cross-entropy loss can be rewritten as

$$\mathcal{L}_s = \frac{1}{N} \sum_{n=1}^N \frac{1}{2} \sum_{i=1}^2 \frac{1}{H \times W} \sum_{n=1}^{H \times W} \mathcal{W} \ell_{\text{CE}}(p_{i,n}, y_{i,n}) \quad (8)$$

where ℓ_{CE} denotes the cross-entropy loss function, N, H, W represent the number of channels, height, and width of a feature map, respectively.

In the new training loss function, we introduce a discrepancy-aware weighting mechanism that applies a larger weight parameter λ to regions with cross-view prediction inconsistencies. This design aims to enhance penalty on divergent regions, directing the model to focus more on areas with prediction discrepancies.

D. Multiscale Contextual Information Aggregation

This study performs aggregation of multiscale contextual information. This provides rich multiscale contextual information for identifying lunar surface geological structure types, thereby enhancing the model's capability to understand lunar geological formations. Specifically, the high-resolution feature maps in the encoder are first processed with dilated convolutions of different dilation rates to extract features at different scales while maintaining the same dimensions as the decoder feature maps. Then, these feature maps are fused with the decoder feature maps through skip connections. The mathematical expression of this process can be constructed as follows:

$$F_{\text{dili}} = \sigma(C_{r_i}(F_{\text{encoder}}; W_i)) \quad (9)$$

$$F_{\text{fused}} = \phi(\text{Conv}(F_{\text{dili}_1}, F_{\text{dili}_2}, \dots, F_{\text{dili}_n}), F_{\text{dec}}) \quad (10)$$

where F_{encoder} denotes the feature map from the encoder stage, F_{dili} represents the intermediate feature map, F_{dec} corresponds to the feature at the corresponding level in the decoder, C_{r_i} refers to the dilated convolution operation with a dilation rate of r_i , W_i represents the weight parameters of the convolution, and σ denotes a nonlinear activation function. F_{fused} denotes the resulting feature map after fusion.

IV. EXPERIMENTS

A. Evaluation Metrics

To comprehensively evaluate the performance of the proposed lunar multitype geological structure classification method, this study adopts commonly used semantic segmentation evaluation metrics: Precision, recall, $F1$ -score, and mean intersection over union (mIoU) as objective evaluation criteria. Their calculation formulas are as follows:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (11)$$

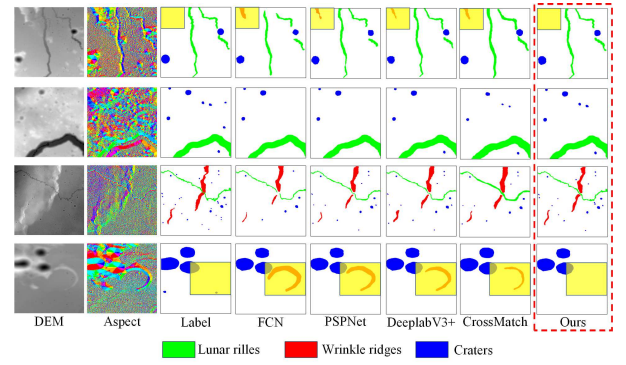


Fig. 6. Performance comparison of different methods.

TABLE I
LUNAR SURFACE GEOLOGICAL STRUCTURES RECOGNITION RESULTS

Method	Precision(%)	Recall(%)	F1-score(%)	mIoU(%)
FCN	79.5	77.2	78.3	77.6
PSPNet	81.3	79.5	80.3	79.4
DeeplabV3+	82.2	80.7	81.4	80.2
CrossMatch	83.1	83.5	82.6	80.8
Ours	84.3	86.8	85.5	82.7

TABLE II
IOU METRICS FOR DIFFERENT GEOLOGICAL STRUCTURE TYPES

Method	Background	Wrinkle Ridges	Lunar Rilles	Craters	mIoU
FCN	83.2	61.7	84.6	78.3	77.6
PSPNet	86.5	63.2	86.2	80.8	79.4
DeeplabV3+	87.2	64.5	86.6	81.2	80.2
CrossMatch	87.6	64.8	86.8	81.9	80.8
Ours	89.8	68.3	86.9	83.1	82.7

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (12)$$

$$F1\text{-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (13)$$

$$\text{mIoU} = \frac{1}{n} \sum_{i=1}^n \frac{\text{TP}_i}{\text{TP}_i + \text{FP}_i + \text{FN}_i}. \quad (14)$$

Among them, true positive (TP) represents the number of positive samples that are correctly predicted as positive by the model. True negative (TN) represents the number of negative samples that are correctly predicted as negative. False positive (FP) represents the number of negative samples that are incorrectly predicted as positive. False negative (FN) represents the number of positive samples that are incorrectly predicted as negative. Here, n represents the number of categories.

B. Comparison Experiments and Ablation Experiments

To validate the effectiveness and superiority of the proposed cross-view constrained lunar multitype geological structure identification model, this section conducts comparative experiments with several high-performance deep learning semantic segmentation models, including FCN [14], PSPNet [15], and DeeplabV3+ [16] and CrossMatch [13].

All comparison models were trained and tested using dual-channel data composed of DEM and aspect data concatenated together, while the proposed model separately inputs DEM data and aspect data into two subviews to fully leverage cross-view constraints for extracting differential information. Fig. 6

TABLE III
QUANTITATIVE EVALUATION OF ABLATION EXPERIMENTS

CVC	DWL	MCIA	Precision(%)	Recall(%)	F1-score(%)	mIoU(%)
			79.6	76.1	77.8	77.9
✓			82.7	84.3	83.4	81.3
✓	✓		83.1	84.7	83.8	81.8
✓	✓	✓	84.3	86.8	85.5	82.7

Note: CVC denotes Cross-View Constraints, DWL denotes Divergence-Weighted Loss, and MCIA denotes Multi-scale Contextual Information Aggregation.

qualitatively demonstrates the identification and classification results of different models for lunar multitype geological structures across various scenarios. Table I presents the comparative results of various performance metrics achieved by each model on the lunar multitype geological structure dataset. Table II presents the IoU metrics for each individual geological structure type.

1) *Qualitative Analysis*: Qualitative analysis of the prediction results from various models reveals the following findings.

1) All network models achieved satisfactory detection performance for lunar rilles, owing to their distinct elevation characteristics and consistent U-shaped morphology.

2) All five models demonstrate competent detection performance for fresh impact craters. Notably, the yellow bounding boxes in Fig. 6 highlight misclassification cases occurring during the identification process. When dealing with degraded craters, the fractured crater walls and degraded floor topography characteristics lead to disappearance of elevation differences between interior and exterior, causing the loss of typical crater features consequently. In these cases, the degraded crater walls become morphologically similar to wrinkle ridges and are difficult to distinguish. CVL-Net shows significant advantages in classification performance by extracting discriminative information from the two data modalities and incorporating multiscale contextual information.

3) Due to the undulating and complex elevation variations characteristic of lunar wrinkle ridges, significant discrepancies were observed in the prediction results across different models. In contrast, CVL-Net outperformed all comparative models by employing cross-view constraints to more effectively capture discriminative elevation semantics and slope orientation features, consequently achieving superior performance in the identification and classification of wrinkle ridges, lunar rilles, and impact craters.

2) *Quantitative Analysis*: Quantitative analysis of the prediction results of different geological structure types through different models reveals that in terms of overall performance, the four evaluation indicators of CVL-Net perform the best, as shown in Table I. In the lunar stream and fresh crater detection missions, due to their distinct features, the detection effects of these five models were roughly the same. In the detection of crease ridges, the advantage of CVL-Net is the most obvious. Its detection performance is improved by 5.1, 3.3, 2.5, and 1.9 percentage points, respectively, compared with the other four models as shown in Table II.

Ablation studies were conducted to validate the essential contribution of each proposed component, with the results summarized in Table III. The results conclusively demonstrate

that each proposed component contributes positively and indispensably to the final performance.

V. CONCLUSION

To address the current challenges in identifying and classifying multitype geological structures on the lunar surface, this study proposes a cross-view lunar geological structure recognition method. By leveraging cross-view constraints, the deep learning model significantly enhances its ability to interpret different types of data. A discrepancy-weighted loss function guides the model to focus more on the differences between multiview predictions. Additionally, multiscale context aggregation integrates contextual features at various scales, providing more comprehensive evidence for determining geological structure types. The proposed method achieves the identification of lunar wrinkle ridges, lunar rilles, and craters, improving F1-score and mIoU by 2.9% and 1.9%, respectively. In the future, a higher resolution data will be used to improve the accuracy and generalization of identifying lunar geographical features.

REFERENCES

- [1] T. Lu et al., "The 1: 2,500,000-scale global tectonic map of the Moon," *Sci. Bull.*, vol. 67, no. 19, pp. 1962–1966, Aug. 2022.
- [2] H. D. Tjia, "Lunar wrinkle ridges indicative of strike-slip faulting," *Geological Soc. Amer. Bull.*, vol. 81, no. 10, pp. 3095–3100, 1970.
- [3] R. A. Schultz, "Localization of bedding plane slip and backthrust faults above blind thrust faults: Keys to wrinkle ridge structure," *J. Geophys. Res., Planets*, vol. 105, no. E5, pp. 12035–12052, May 2000.
- [4] L. M. Pigue, K. A. Bennett, B. H. N. Horgan, and L. R. Gaddis, "Relationship between explosive and effusive volcanism in the montes apenninus region of the Moon," *J. Geophys. Res., Planets*, vol. 128, no. 11, pp. –1, Nov. 2023, Art. no. e2023JE007861.
- [5] D. Nesvorný et al., "Early bombardment of the Moon: Connecting the lunar Crater record to the terrestrial planet formation," *Icarus*, vol. 399, Jul. 2023, Art. no. 115545.
- [6] J. Liu et al., "Landing site selection and overview of China's lunar landing missions," *Space Sci. Rev.*, vol. 217, no. 1, pp. 1–25, Feb. 2021.
- [7] N. Harada et al., "Recognition algorithm for topographic features," in *Proc. 7th IEEE Int. Conf. Comput. Inf. Technol. (CIT)*, Oct. 2007, pp. 685–689.
- [8] H. Jiang, X. Tian, and A. Xu, "A new segmentation algorithm for lunar surface terrain based on CCD images," *Res. Astron. Astrophys.*, vol. 15, no. 19, p. 1604, 2015.
- [9] J. Wang, W. Cheng, C. Zhou, and X. Zheng, "Automatic mapping of lunar landforms using DEM-derived geomorphometric parameters," *J. Geographical Sci.*, vol. 27, no. 11, pp. 1413–1427, Nov. 2017.
- [10] S. Ghosh, T. F. Stepinski, and R. Vilalta, "Automatic annotation of planetary surfaces with geomorphic labels," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 1, pp. 175–185, Jan. 2010.
- [11] X. Lu, J. Sun, G. Shu, J. Zhao, and N. Li, "Detecting the lunar wrinkle ridges through deep learning based on DEM and aspect data," *Res. Astron. Astrophys.*, vol. 25, no. 8, Aug. 2025, Art. no. 085016.
- [12] P. A. Burrough and R. A. McDonnell, *Principles of Geographical Information Systems*. Oxford, U.K.: Oxford Univ. Press, 1998.
- [13] R. Liu et al., "CrossMatch: Cross-view matching for semi-supervised remote sensing image segmentation," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, 2024, Art. no. 5650515.
- [14] E. Shelhamer, J. Long, and T. Darrell, "Fully convolutional networks for semantic segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 4, pp. 640–651, Apr. 2017.
- [15] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, "Pyramid scene parsing network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 6230–6239.
- [16] L. C. Chen et al., "Encoder-decoder with atrous separable convolution for semantic image segmentation," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*. Berlin, Germany: Springer, 2018, pp. 833–851.