Deeply Seeking Boundary for Lunar Regolith Segmentation

Yifeng Wang[1,2*], Lingxin Wang[3*], Lu Zhang[3*], Yang Li[4], Chao Xu[5], Weiwei Zhang[3], Junyue Tang[3], Yanhong Zheng[6], Yong Pang[7], Shengyuan Jiang[3†], Yi Zhao[2†], Zongquan Deng[3]

[1]Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China
[2]School of Science, Harbin Institute of Technology, Shenzhen 518055, China
[3]School of Mechatronic Engineering, Harbin Institute of Technology, Harbin 150001, China
[4]Center for Lunar and Planetary Sciences, Chinese Academy of Sciences, Guiyang 550081, China
[5]School of Materials Science and Engineering, Harbin Institute of Technology, Harbin 150000, China
[6]Beijing Institute of Spacecraft System Engineering, Beijing 100094, China
[7]Beijing Spacecrafts, China Academy of Space Technology, Beijing 100089, China

Abstract

The sharp, intricate contours of lunar regolith particles hold critical clues to the Moon's geological evolution and inform engineering applications from habitat construction to spacecraft design, making their precise segmentation a task of significant scientific and engineering value. However, this task exposes a weakness in deep learning models known as spectral bias, an inherent tendency to learn smooth, low-frequency functions which causes them to systematically erase the very high-frequency boundary details that are of primary interest. To resolve this conflict, we propose a framework to deeply seek object boundaries. First, we propose High-Frequency Initialized LoRA (HiFi-LoRA) to counteract spectral bias. By initializing the LoRA adaptation matrices as the optimal low-rank approximation of a high-pass filter, it fundamentally enhances the model's high-frequency perception and injects a strong preference for edges. Second, we propose the Wavelet Energy Modulation (WEM) regularizer. It guides the model to learn the intrinsic correlation between contour complexity and mask area, forcing the model to build a geometric understanding of contour morphology upon its high-frequency perception, thereby enabling the generation of boundary details commensurate with the object's scale. Experimentally, we constructed the Lunar Regolith Segmentation Dataset (LRSD), the first large-scale benchmark with expert-annotated contours. Extensive experiments demonstrate that our method sets a new state of the art on this challenging benchmark, not only achieving top performance on regional metrics like mIoU and DSC but, more critically, drastically outperforming existing models on boundary accuracy. This work not only provides a powerful computational tool for lunar science but also offers a robust and synergistic design pattern for other fine-grained segmentation challenges.

Introduction

The analysis of lunar regolith is fundamental to advancing lunar science and enabling sustained human presence on the Moon [Zhang et al. 2021; Lu et al. 2023]. The morphology of regolith particles provides a direct window into the Moon's geological history, recording billions of years of micrometeoroid bombardment and space weathering in the absence of atmospheric and aqueous erosion [Zhang et al. 2019; Liu et al. 2012]. Furthermore, the engineering properties of the regolith, which are dictated by particle shape and texture, are critical for designing durable aerospace equipment, planning excavation and construction for future lunar bases, and developing efficient in-situ resource utilization techniques [Zhang, Fa, and Jia 2025]. Accurate, large-scale segmentation of regolith particles from high-resolution imagery is therefore a task of paramount scientific and engineering importance, providing the foundational data for these endeavors.

A unique challenge in this domain is the critical importance of particle boundaries [Ding et al. 2024]. Unlike terrestrial soils, whose particles are typically rounded by weathering, lunar regolith particles are characterized by sharp, angular, and highly irregular contours [Cambioni et al. 2021]. These fine-grained, high-frequency details are not noise but are instead the primary carriers of geological information [Hsu et al. 2022]. Consequently, for lunar regolith analysis, segmentation is not merely about identifying the general area of a particle. It is about precisely delineating its complex boundary. Standard segmentation models, even when they achieve high regional overlap scores, often fail at this crucial task by producing overly smooth predictions that erase the very details of scientific interest, rendering them unsuitable for this application [Yang et al. 2025a].

This failure is not an incidental flaw of specific architectures but a manifestation of a deep-seated and well-documented limitation of neural networks known as spectral bias, or the frequency principle [Rahaman et al. 2019; Chen et al. 2021]. Numerous studies have shown that deep neural networks trained via gradient descent are inherently biased towards learning low-frequency functions before they can capture high-frequency details [Fridovich-Keil, Gontijo Lopes, and Roelofs 2022; Kiessling and Thor 2022]. This phenomenon, rigorously analyzed through tools like

*These authors contributed equally.
†Corresponding authors.

the Neural Tangent Kernel (NTK), reveals that the convergence rate for different frequency components is unequal, with lower frequencies being learned faster [Jacot, Gabriel, and Hongler 2018; Bietti and Mairal 2019]. For image segmentation, this bias translates into a natural tendency to produce smooth outputs that capture the low-frequency bulk of objects while neglecting the high-frequency information present at sharp edges and fine textures [Shi et al. 2022; Fang et al. 2023; Zhang et al. 2025b]. This poses a fundamental challenge for standard deep learning approaches when applied to problems where boundary fidelity is the primary objective [Luo et al. 2022; Yang et al. 2025b].

To resolve the fundamental contradiction between the critical importance of Lunar regolith particle contours and the inherent tendency of deep learning models to erase high-frequency details, we propose a framework designed to deeply seek object boundaries in fine-grained segmentation tasks. Our main contributions are as follows:

- We constructed the Lunar Regolith Segmentation Dataset (LRSD), the first large-scale benchmark with expert-annotated, high-frequency contours. This provides the community with the essential data needed to drive and validate research on this critical task.
- To counteract the model's inherent bias towards smoothness, we propose High-Frequency Initialized LoRA (HiFi-LoRA), which initializes the LoRA adaptation matrices to be the best low-rank approximation of a high-pass filter, thereby fundamentally enhancing its high-frequency perception capability and injecting a strong preference for edges.
- To ensure the generated contours are physically plausible, we introduce the Wavelet Energy Modulation (WEM) regularizer, which guides the model to learn the intrinsic correlation between contour complexity and mask area, thereby adaptively regulating the high-frequency perception capability endowed by HiFi-LoRA to ensure the model generates commensurate boundary details for objects of different scales.
- By guiding the model to deeply explore object boundaries, our work achieves the first high-fidelity segmentation of lunar regolith particles, providing a powerful computational tool to advance lunar geology and engineering.

Methodology

HiFi-LoRA

The accurate segmentation of fine-grained details presents a persistent challenge in computer vision, particularly within scientific domains such as the analysis of lunar regolith, where sharp, angular particle boundaries carry critical geological information. While large foundation models like the Segment Anything Model (SAM) have demonstrated remarkable semantic understanding, they often falter in these specialized tasks [Kirillov et al. 2023; Mazurowski et al. 2023]. A primary reason for this deficiency is their tendency to over-smooth fine structures, erasing details crucial for analysis [Zheng et al. 2025]. This behavior is not an incidental flaw but a manifestation of a fundamental property known as spectral bias, where deep neural networks trained via gradient descent inherently prioritize learning low-frequency functions over high-frequency details [Xu et al. 2020; Chen et al. 2024]. Low-Rank Adaptation (LoRA) has become a standard for adapting such models [Fu et al. 2023; Liu et al. 2022; Hu et al. 2022]. However, the conventional LoRA framework, which relies on a task-agnostic random initialization, fails to counteract this inherent spectral bias [Cao et al. 2021]. The model is still forced to learn high-frequency sensitivity entirely from the data, an inefficient process that often converges to an overly smooth solution. This gap highlights the need for an adaptation strategy that is not merely parameter-efficient but also frequency-aware from its inception. To address this fundamental challenge, we introduce High-Frequency initialized Low-Rank Adaptation, HiFi-LoRA, which enhances the capability of pretrained models by embedding a high-frequency prior directly into the adaptation mechanism. For a pretrained weight matrix $W_0 \in \mathbb{R}^{d\times k}$, LoRA freezes W_0 and introduces a trainable low-rank update $\Delta W = BA$, where $B \in \mathbb{R}^{d\times r}$ and $A \in \mathbb{R}^{r\times k}$, with the rank r being substantially smaller than d and k. The modified forward pass is expressed as

$$h = W_0 x + BAx, \tag{1}$$

where x is the input and h is the output hidden state. Standard LoRA initializes A from a random Gaussian distribution while setting B to zeros, an uninformed strategy that makes capturing subtle boundary features inefficient.

Our central idea is to replace the standard random initialization of LoRA with a meticulously designed deterministic scheme. We hypothesize that by initializing the weight update ΔW to approximate a high-pass filter, we can predispose the model to be sensitive to edges and textures from the very first iteration of training. This approach guides the optimization process from a neutral starting point to a more promising region of the parameter space, one that inherently favors the preservation of boundary details. This design philosophy forms a direct correspondence with the physical properties of our target problem. The angular and coarse contours of lunar regolith particles, formed due to constant micrometeoroid bombardment and the absence of atmospheric weathering, manifest directly as high-frequency signals in the image domain . Therefore, our proposed method is not merely an algorithmic improvement but a solution deeply motivated by the physical nature of the problem itself. The entire architectural design of our HiFi-LoRA initialization is illustrated in Fig. 1.

To transform the concept of a high-frequency prior into a concrete mathematical entity, we first construct an ideal high-pass filter. We employ the two-dimensional Discrete Cosine Transform (DCT) as the tool for this construction due to its excellent energy compaction properties for real-valued image data, a characteristic famously leveraged in standards like JPEG . An $N \times N$ DCT matrix C contains elements C_{ij} given by the formula

$$C_{ij} = \alpha_i \cos\left(\frac{\pi(2j+1)i}{2N}\right), \tag{2}$$

where the indices i and j range from 0 to $N-1$. The scaling

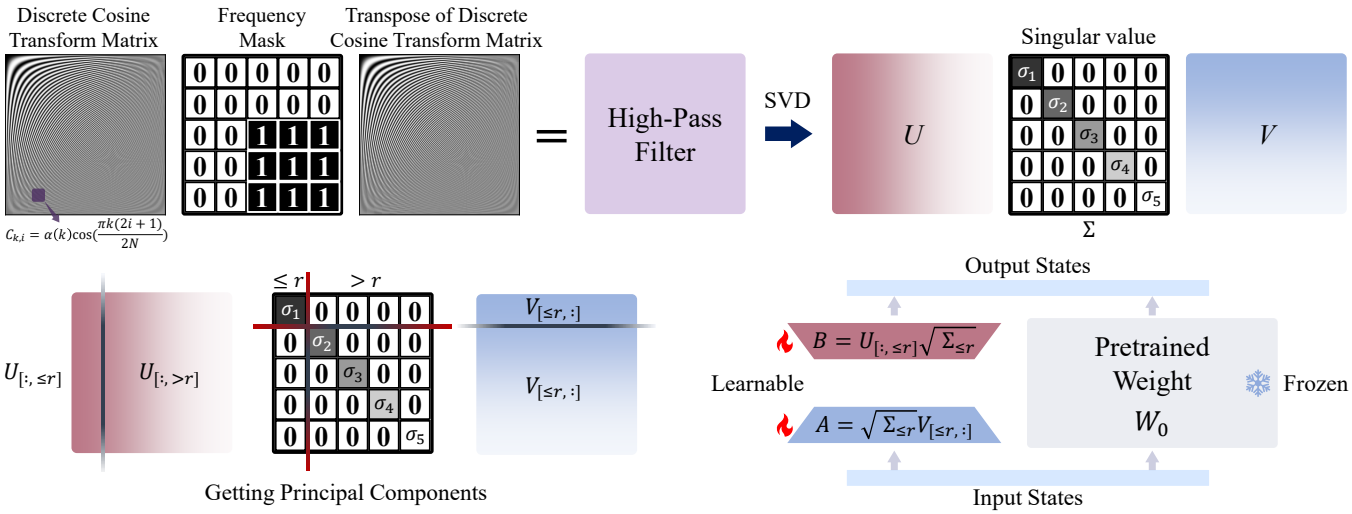

Figure 1: Diagram of HiFi-LoRA initialization. HiFi-LoRA begins by constructing a high-pass filter from a DCT matrix. This filter undergoes SVD to extract its principal components. The most significant singular values and their corresponding vectors are then used to initialize the LoRA matrices A and B, embedding a high-frequency prior into the model before fine-tuning.

factor α_i is defined as

$$\alpha_i = \begin{cases} \sqrt{1/N} & \text{if } i = 0 \\ \sqrt{2/N} & \text{if } i > 0 \end{cases}. \tag{3}$$

This DCT matrix C is orthogonal. For a target weight matrix of size $d \times d$, we construct a corresponding $d \times d$ spatial-domain high-pass filter matrix H. This process begins with a binary frequency mask matrix $M \in \{0,1\}^{d\times d}$. The elements M_{uv} of this mask are set to 0 for indices (u, v) corresponding to low frequencies and 1 for indices corresponding to high frequencies. By transforming this frequency mask M back to the spatial domain using the inverse DCT, we synthesize the high-pass filter H. Because C is orthogonal, its inverse transform is equivalent to its transpose

$$H = C^T M C, \tag{4}$$

where H now represents a linear operator that effectively enhances the high-frequency components of an input signal.

The next critical challenge is that our constructed filter H is a full-rank matrix, whereas the LoRA update ΔW must be low-rank. A direct assignment of $\Delta W = H$ is infeasible. To resolve this rank mismatch, we must find the best low-rank approximation of H. According to the Eckart-Young-Mirsky theorem [Golub, Hoffman, and Stewart 1987], the optimal rank-r approximation of a matrix in the Frobenius norm is obtained through its Singular Value Decomposition (SVD). We therefore decompose the high-pass filter H as

$$H = U\Sigma V^T. \tag{5}$$

Here, $U \in \mathbb{R}^{d\times d}$ and $V \in \mathbb{R}^{d\times d}$ are orthogonal matrices containing the left and right singular vectors of H respectively. $\Sigma \in \mathbb{R}^{d\times d}$ is a diagonal matrix whose diagonal elements are the singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_d \geq 0$, sorted in descending order. These singular values quantify the importance of each principal component of the filter, with larger values corresponding to the most dominant transformational directions of the high-pass filtering operation.

Based on the SVD result, we define the initialization rule for HiFi-LoRA. We select the top r largest singular values and their corresponding left and right singular vectors to construct the rank-r approximation of H. This approximation is $H_r = U_r\Sigma_r V_r^T$, where $U_r \in \mathbb{R}^{d\times r}$ is the matrix formed by the first r columns of U, $V_r \in \mathbb{R}^{d\times r}$ is the matrix formed by the first r columns of V, and $\Sigma_r \in \mathbb{R}^{r\times r}$ is the diagonal matrix containing the top r singular values. We then initialize the LoRA matrices A and B as

$$B_{init} = U_r\sqrt{\Sigma_r}, \tag{6}$$

$$A_{init} = \sqrt{\Sigma_r}V_r^T. \tag{7}$$

In this initialization scheme, we distribute the singular values evenly between A and B by using the square root. This practice balances the norms of the two initial weight matrices, thereby promoting more stable gradient propagation during the early stages of training and avoiding optimization issues that could arise if one matrix had a much larger magnitude than the other. Ultimately, in a network layer adapted with HiFi-LoRA, the forward pass is modified to

$$h = W_0 x + s \cdot BAx, \tag{8}$$

where the pretrained weights W_0 remain frozen, while only the matrices A and B, now imbued with a high-frequency prior, participate in training and updates. The scalar s is a tunable scaling factor that controls the magnitude of the adaptation. In this manner, we transform LoRA from a generic adaptation technique into a specialized tool tailored to solve segmentation tasks with complex, high-frequency boundaries.

We establish the theoretical efficacy of HiFi-LoRA by analyzing its reshaping of learning dynamics within the Neural Tangent Kernel (NTK) framework. The convergence rate of different frequency components in the learning dynamics is determined by the eigenvalues of the NTK. Spectral bias arises because the NTK's eigenvalues corresponding

to high-frequency functions are significantly smaller than those for low frequency, leading to slower learning of high-frequency details. In LoRA fine-tuning, the learning process is governed by the NTK of the trainable LoRA branch, Θ_{lora}. We derive that this kernel's structure is determined by the initialization of matrices A and B:

$$\Theta_{\text{lora}}(x, x') = s^2(BB^T)(x^T x') + s^2(x^T A^T A x')I_d. \quad (9)$$

For standard LoRA, the initial kernel Θ_{std} degenerates due to $B_{\text{init}} = 0$ and relies on a random matrix A_{init}, exhibiting an inherent low-pass characteristic that leads to spectral bias. In contrast, HiFi-LoRA's deterministic initialization scheme constructs a fundamentally different initial kernel, Θ_{hifi}. Its critical component is structured by the principal components (singular vectors V_r and values Σ_r) of a high-pass filter:

$$K_H(x, x') = x^T(V_r \Sigma_r V_r^T)x'. \quad (10)$$

This kernel operator, K_H, effectively measures the alignment of inputs along key high-frequency directions. Consequently, the spectral structure of Θ_{hifi} is biased towards high-frequency functions from the outset, ensuring that the learning dynamics corresponding to high-frequency components have large effective eigenvalues. In essence, our analysis reveals that HiFi-LoRA reshapes the learning dynamics by directly embedding a high-frequency preference into the spectral structure of the learning operator (the NTK) at the beginning of training, thereby mitigating spectral bias. The full derivation is provided in the Appendix.

Wavelet Energy Modulation

While HiFi-LoRA primes the model to be sensitive to high-frequency information, we introduce a further mechanism to explicitly guide the output generation towards detailed and physically plausible contours. A naive approach of simply maximizing boundary details is suboptimal because the expected complexity of a mask's contour is intrinsically linked to its overall size (we provide a rigorous statistical proof in the Appendix demonstrating that the expected mask boundary length is maximized when the area ratio is 50%). Therefore, we propose the Wavelet Energy Modulation (WEM) regularizer, a novel loss term that dynamically adjusts the target boundary complexity based on the predicted mask's area. First, we quantify the boundary detail of the predicted probability mask $P \in^{H \times W}$ using its normalized wavelet energy E. We employ a 2D Discrete Wavelet Transform (DWT) to decompose the mask P into its approximation sub-band P_{LL} and detail sub-bands P_{LH}, P_{HL}, and P_{HH}. The normalized wavelet energy E is the ratio of the energy in the high-frequency detail sub-bands to the total energy of the mask. This is a fully differentiable operation defined as

$$E = \frac{\sum_{i,j}(P_{LH}[i,j]^2 + P_{HL}[i,j]^2 + P_{HH}[i,j]^2)}{\sum_{i,j} P[i,j]^2}. \quad (11)$$

Next, we define the mask area ratio γ as the mean pixel intensity of the predicted mask, $\gamma = \frac{1}{HW}\sum_{i,j} P[i,j]$. Our WEM regularizer R_{WEM} is then formulated to modulate the target energy E based on this area ratio γ. The regularizer is constructed by

$$R_{WEM} = \gamma(\gamma - 1)\log(E) - |1 - 2\gamma|\log(1 - E). \quad (12)$$

The brilliance of this formulation lies in how it implicitly defines an optimal target energy E_{opt} that minimizes R_{WEM} for any given γ. By setting the derivative of R_{WEM} with respect to E to zero, we find this optimal energy to be $E_{opt}(\gamma) = \frac{\gamma(1-\gamma)}{\gamma(1-\gamma)+|1-2\gamma|}$. This function's behavior perfectly captures our physical intuition. For masks that are very small where $\gamma \to 0$ or very large where $\gamma \to 1$, the optimal energy E_{opt} approaches zero, encouraging simpler boundaries. Conversely, when the mask occupies half the image area where $\gamma = 0.5$, E_{opt} reaches its maximum value of one, encouraging the most complex and detailed contours. This adaptive relationship is illustrated in Fig. 2. The final training objective combines the primary segmentation loss L_{seg} with our WEM regularizer, weighted by a hyperparameter λ. The total loss is $L_{total} = L_{seg} + \lambda R_{WEM}$. This completes our framework, which not only initializes the model with a high-frequency bias but also actively sculpts the final output to produce exceptionally detailed segmentations.

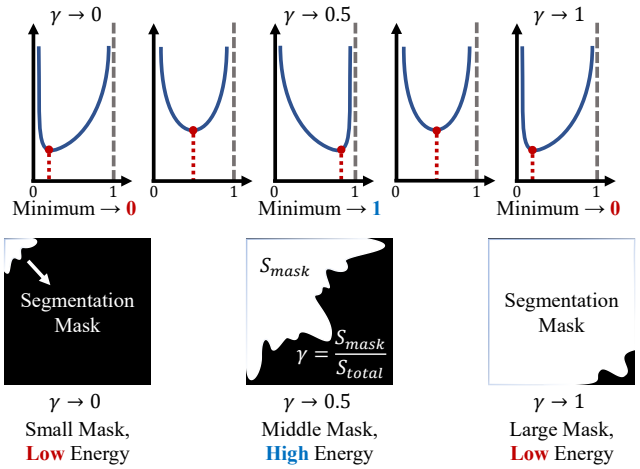

Figure 2: Conceptual illustration of the WEM regularizer. The loss function (blue curve) dynamically changes based on the mask area ratio γ. For small or large masks (low γ or high γ), the loss is minimized at a low wavelet energy E. For masks covering approximately half the image ($\gamma \approx 0.5$), the loss is minimized at a high wavelet energy, encouraging the model to produce more detailed contours.

Experiments and Results

Experiment Dataset

We constructed a bespoke benchmark dataset designed specifically for the fine-grained segmentation of lunar regolith particles. The lunar samples were provided by the China National Space Administration (CNSA). We used the ZEISS Xradia 520 Versa model X-ray micro-CT instrument to scan lunar regolith samples. This device can achieve a spatial resolution of up to 0.7 microns, thereby obtaining high-precision imaging data for accurately capturing the complex contours and textures of particles. Given the critical importance of particle contours, we established a stringent annotation quality control protocol to ensure label precision. Each image was manually annotated independently by

two trained specialists to generate initial binary segmentation masks. We quantified the agreement between these two independent annotations using the Intersection over Union (IoU). If the IoU between the two annotations exceeded the threshold of 0.95, we considered the annotators to be in high agreement on the particle boundaries. In such cases, the final ground truth was defined as the pixel-wise intersection of the two masks. This conservative strategy ensures that the labels for high-confidence samples contain only the pixels of utmost certainty, creating a reliable gold standard for evaluating boundary accuracy. Conversely, an IoU at or below 0.95 indicated disagreement on the contour details. These disputed samples were escalated to a senior planetary geologist for a third, decisive annotation, which served as the final ground truth. This expert adjudication process ensures that even the most challenging samples with complex or faint boundaries possess the highest degree of scientific accuracy. Through this process, we built the first-of-its-kind Lunar Regolith Segmentation Dataset (LRSD) with finely annotated contours. To provide a comprehensive evaluation, we employed a complementary suite of metrics. We use the Dice Similarity Coefficient (DSC) and mean Intersection over Union (mIoU) to assess the accuracy of regional overlap. More importantly, to directly quantify model performance on the core challenge of boundary delineation, we adopted the 95th percentile Hausdorff Distance (HD95). HD95 specifically measures the distance between the predicted and ground truth boundaries, where a lower value signifies more precise contour matching [Li et al. 2025].

Comparative Results

We evaluated our method with a comprehensive suite of contemporary segmentation models, with quantitative results in Table 1 and qualitative comparisons in Fig. 3. For this task, the HD95 is the most critical metric, as it measures the precision of boundary delineation, which holds the key geological information [Junjie et al. 2025; Zhang et al. 2025c].

Method	**mIoU ↑**	**DSC ↑**	**HD95 ↓**
nnU-Net	80.7	89.3	13.55
SegFormer	85.1	92.0	9.97
MaskFormer	85.8	92.4	8.92
Mask2Former	86.5	92.8	7.65
VWFormer	86.8	93.0	7.08
EDAFormer	87.2	93.2	6.54
U-Mamba	86.9	93.1	6.97
SegMAN	87.5	93.4	<u>6.39</u>
SAM (LoRA Tuned)	84.3	91.5	11.18
MedSAM (LoRA Tuned)	<u>87.9</u>	<u>93.6</u>	6.74
Ours	**89.1**	**94.2**	**4.53**

Table 1: Comparison with latest segmentation methods.

Our evaluation begins with convolutional architectures, where nnU-Net serves as a particularly strong baseline. As a framework that automatically configures and optimizes a U-Net pipeline for any given dataset, nnU-Net has become a dominant force in medical segmentation challenges and represents the peak of what a systematically optimized convolutional architecture can achieve [Isensee et al. 2021]. Yet, despite its sophisticated self-adaptation, it fails on our task. Its high HD95 score, visually confirmed in Fig. 3, shows it produces excessively smooth predictions that erase the sharp, angular contours of the regolith particles. Advanced Transformer and state-space models, including SegFormer, MaskFormer, Mask2Former, VWFormer, EDAFormer, U-Mamba and the state-of-the-art SegMAN [Xie et al. 2021; Cheng, Schwing, and Kirillov 2021; Cheng et al. 2022; Yan, Wu, and Zhang 2024; Yu et al. 2024; Ma, Li, and Wang 2024; Fu, Lou, and Yu 2025], show improved regional accuracy (mIoU, DSC). However, their boundary precision remains limited, as evidenced by their HD95 scores. This persistent failure across diverse, powerful architectures points to a systemic weakness known as spectral bias, where networks inherently favor learning smooth, low-frequency functions. The fractal-like edges of lunar regolith particles are high-frequency signals that standard training struggles to learn.

The analysis of large foundation models further clarifies this challenge. The Segment Anything Model, when fine-tuned with standard LoRA , yields a poor HD95 of 11.18. This empirically validates our motivation that a generic, task-agnostic fine-tuning strategy may be incapable of overcoming spectral bias and instead converges to an overly smooth solution. The case of MedSAM is particularly insightful. MedSAM inherits SAM's powerful architecture and is further specialized on a massive dataset of over one million medical images. This pre-training on biological structures such as organs and tumors, which are characterized by macroscopically continuous and relatively smooth boundaries, endows MedSAM with an exceptionally strong low-frequency prior [Zhao et al. 2025; Zhang et al. 2024; Wang et al. 2024]. Consequently, MedSAM excels at identifying the main particle bodies, achieving a strong DSC score. The model's expertise on smooth biological objects becomes a powerful but misleading bias when confronted with the sharp, irregular, non-biological contours of lunar regolith [Zhang et al. 2025a]. Fig. 3 clearly shows MedSAM capturing the particle core accurately but failing at its jagged edges. This proves that even massive pre-training can be counterproductive if the frequency characteristics of the source and target domains are misaligned, powerfully arguing for a mechanism that explicitly injects a high-frequency prior during adaptation.

Our full method, combining HiFi-LoRA and the WEM regularizer, sets a new state of the art. It achieves the highest mIoU and DSC scores, but most importantly, it delivers a breakthrough in boundary precision, reducing the HD95 to 4.53. The visual results in Fig. 3 are definitive. While all other models produce smoothed edges and blunted corners, our prediction is nearly indistinguishable from the ground truth, faithfully capturing the finest details. This precision is a direct result of our framework's design, which both counters the model's inherent spectral bias and explicitly guides it to generate detailed, physically plausible contours.

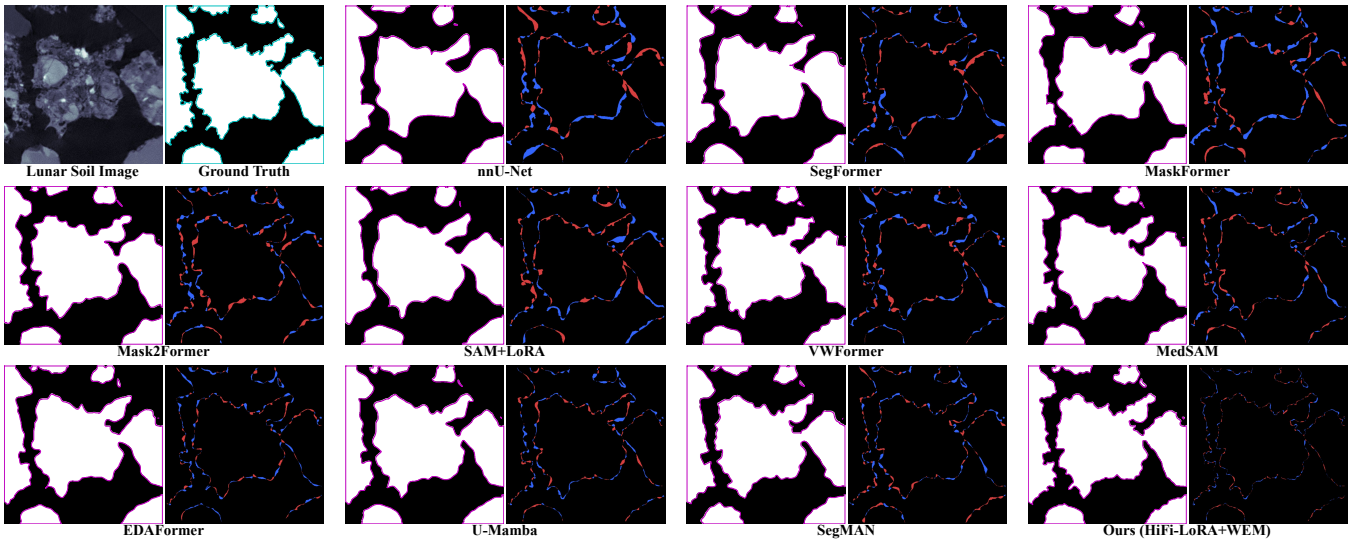

Figure 3: Visual comparison of different methods. For each method, the predicted segmentation mask is shown alongside a error map that visualizes its discrepancy from the ground truth. blue pixels denote false negatives (missed predictions), while red pixels denote false positives (incorrect predictions). A lower density of both colors indicates a more accurate segmentation.

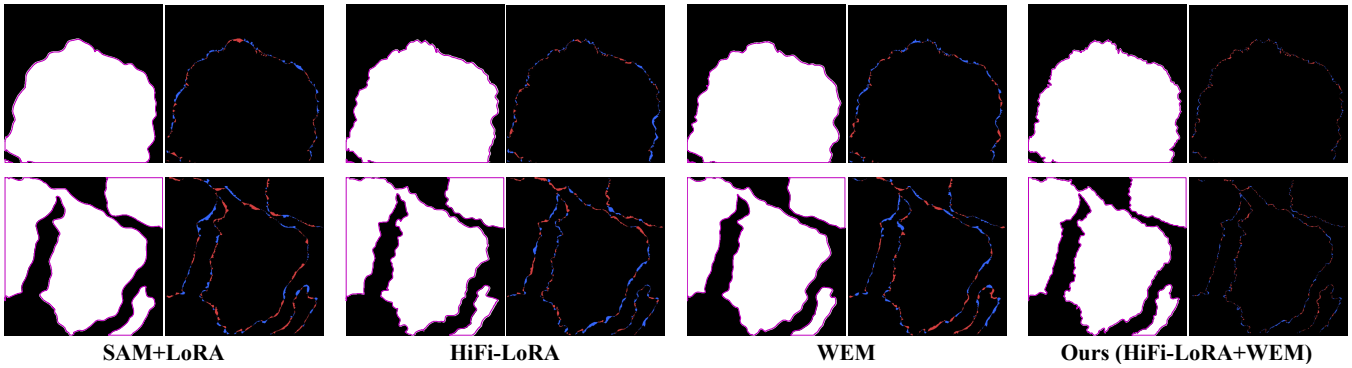

Figure 4: Visual results of the ablation study. On the error maps, blue pixels denote false negatives (missed predictions), while red pixels denote false positives (incorrect predictions).

Ablation Study

To validate the contribution of each component, we conducted a systematic ablation study, with quantitative results in Table 2 and visual comparisons in Fig. 4. Our baseline, a standard LoRA-tuned SAM, performs poorly on boundary metrics (HD95), and its mask in Fig. 4 is visibly oversmoothed, clearly defining the problem we aim to solve.

Method	HiFi-LoRA	WEM	mIoU	DSC	HD95
Baseline	✗	✗	84.3	91.5	11.18
Ablation 1	✓	✗	87.0	93.0	6.15
Ablation 2	✗	✓	86.8	92.9	7.31
Ours (Full)	✓	✓	**89.1**	**94.2**	**4.53**

Table 2: Ablation experiments on the proposed modules.

Introducing only the HiFi-LoRA initialization causes a dramatic improvement, slashing the HD95 score by nearly 45% to 6.15. This quantitative leap is mirrored in Fig. 4, where the mask contour becomes immediately sharper and more detailed. This provides powerful evidence for our core hypothesis. By initializing the LoRA matrices to approximate a high-pass filter, HiFi-LoRA instills a high-frequency prior that directly counters the model's spectral bias, forcing it to be sensitive to edge information from the start of training. Applying only the WEM regularizer also yields a significant improvement, lowering the HD95 to 7.31. This demonstrates that WEM is an effective supervision, explicitly guiding the model to generate contours with a level of complexity appropriate for the mask's size, thereby penalizing the overly simplistic predictions of the baseline.

The true strength of our framework is revealed when HiFi-LoRA and WEM are combined. Our full model achieves the state-of-the-art HD95 score of 4.53, a result far superior to what either component can achieve alone. The visual evidence in Fig. 4 is unequivocal, showing a level of detail that surpasses all other configurations. This highlights a powerful synergy. HiFi-LoRA acts as an internal prior, modifying the model's parameter space to give it the inherent capability to perceive high-frequency features. WEM acts as an external supervisor, modifying the learning objective to provide

targeted guidance on how to shape the final output. A model with only HiFi-LoRA has the capability but lacks specific guidance, while a model with only WEM has the guidance but lacks the specialized internal capability to execute it perfectly. The combination provides both, enabling the model to not only perceive the finest details but also to sculpt them into a final mask that is both complex and physically plausible. This synergy is the key to our framework's success.

Comparative Analysis of SAM and MedSAM

By pre-training on vast datasets, foundation models acquire powerful, general-purpose visual priors, offering a strong starting point for specialized downstream tasks. To assess the influence of different foundational priors on our method, we compared the generalist foundation model SAM with the specialist foundation model MedSAM [Ma et al. 2024; Zhao et al. 2024]. SAM possesses a broad visual prior from diverse natural images, whereas MedSAM has a specialized prior from CT data. We compare the segmentation accuracy of SAM and MedSAM when fine-tuned with our framework using varying amounts of lunar regolith training data. With limited training data, specifically at 3000 samples, MedSAM demonstrates an advantage. Its HD95 score is lower than SAM's with higher regional accuracy on mIoU and DSC. This initial superiority is attributable to MedSAM's extensive pre-training on over one million medical images. The CT images within its training data share a similar imaging style with our lunar regolith CT data, providing it with a powerful and directly relevant prior that allows it to rapidly adapt to the general morphology of the regolith particles even with minimal examples.

As the training set expands, SAM's accuracy improves at a steeper rate, surpassing MedSAM's HD95 score at approximately the 5000 samples and establishing a definitive lead with the full 7000-sample dataset. This trend suggests that while MedSAM's medical imaging prior is advantageous for low-data scenarios, it also imposes an implicit "smoothness" constraint. The model is biased towards the relatively smooth contours typical of biological structures, which ultimately becomes a bottleneck, limiting its ability to fully capture the sharp, high-frequency, non-biological nature of lunar regolith particles. In contrast, SAM, as a more generalist foundation, possesses fewer restrictive domain-specific priors. While this results in weaker initial performance, it also makes the model more adaptable. When supplied with sufficient domain-specific data, SAM can more effectively leverage the explicit high-frequency guidance from our HiFi-LoRA and WEM framework. It learns the unique and complex boundary characteristics of the target domain without being constrained by pre-existing biases, ultimately achieving a superior level of contour fidelity. In summary, a powerful, domain-agnostic prior for high-frequency detail, as provided by our methods, can be more critical than a domain-specific but potentially mismatched semantic prior from the foundation model itself.

Discussion

This study may provide a inspiring paradigm for the Parameter-Efficient Fine-Tuning (PEFT). We believe that PEFT should not be a task-agnostic parameter compression tool, but rather a structured injectors of domain-specific priors. Traditional PEFT methods like LoRA rely on random initialization, forcing the model to passively learn task characteristics from data during optimization [Jiang et al. 2025]. This work, conversely, adopts an active and deliberate design. We first establish a mathematical representation of the downstream task's core feature: identifying high-frequency boundary details as critical for lunar regolith segmentation and precisely representing this requirement as a high-pass filter operator using the Discrete Cosine Transform. Subsequently, we extract the operator's core low-rank components via Singular Value Decomposition and use them as the deterministic initial state for the HiFi-LoRA matrices. This approach encodes abstract domain knowledge into concrete network parameters, reshaping the starting point of fine-tuning. The model no longer needs to combat spectral bias from scratch; it is instead endowed with targeted high-frequency perceptual capabilities from the onset of training. This prior-injection strategy transitions PEFT from data-driven adaptation to knowledge-driven guidance.

The WEM regularizer further completes this framework by imposing geometric constraints in the loss function, providing correct physical guidance for the model's high-frequency capabilities. The synergy between HiFi-LoRA and WEM, injecting priors into the parameter space and loss space respectively, demonstrates a sophisticated design where "capability endowment" and "behavioral constraint" are complementary. The pathway from task characterization to prior injection revealed in this work offers a novel methodology for designing more intelligent and effective specialized adapters for large models, marking an evolution from parameter efficiency to knowledge efficiency.

Conclusion

This paper introduces the task of high-fidelity lunar regolith particle segmentation, a challenge of critical value for revealing lunar geological history and guiding engineering applications like base construction and in-situ resource utilization. This task is fundamentally challenged by the spectral bias of deep networks, which conflicts with the scientific necessity of capturing fine-grained contours. To this end, we propose HiFi-LoRA, which counteracts spectral bias by initializing adaptation matrices to approximate a high-pass filter, thereby reconfiguring the initial Neural Tangent Kernel to favor high-frequency functions. This mechanism is complemented by the devised Wavelet Energy Modulation regularizer, which provides adaptive supervision to ensure the geometric plausibility of the generated contours. We constructed the world's first lunar regolith particle segmentation dataset with fine-grained annotations from domain experts, providing a critical benchmark for research in this area. This work provides a computational tool for lunar science and engineering, while also offering a theoretically grounded design pattern for overcoming spectral bias in other domains where precise high-frequency modeling is essential.

Acknowledgments

The lunar samples (No. CE5C0600YJFM00305H, CE5C0600YJFM00404H, CE5Z0806YJFM003H) were provided by the China National Space Administration (CNSA). This work was supported in part by the National Nature Science Foundation Committee (NSFC) of China (No. 62473115), the University Innovative Team Project of Guangdong Province (No. 2022KCXTD039), the Key R&D Program of Heilongjiang Province (No. 2024ZXDXB52), and the Open Fund of National Key Laboratory of Deep Space Exploration (No. NKDSEL2024003). We also thank the Analytical and Testing Center at Harbin Institute of Technology for the instrument support.

References

Bietti, A.; and Mairal, J. 2019. On the inductive bias of neural tangent kernels. *Advances in Neural Information Processing Systems*, 32.

Cambioni, S.; Delbo, M.; Poggiali, G.; Avdellidou, C.; Ryan, A.; Deshapriya, J.; Asphaug, E.; Ballouz, R.-L.; Barucci, M.; Bennett, C.; et al. 2021. Fine-regolith production on asteroids controlled by rock porosity. *Nature*, 598(7879): 49–52.

Cao, Y.; Fang, Z.; Wu, Y.; Zhou, D.-X.; and Gu, Q. 2021. Towards Understanding the Spectral Bias of Deep Learning. In *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI-21*, 2205–2211.

Chen, G.-Y.; Gan, M.; Chen, C. P.; Zhu, H.-T.; and Chen, L. 2021. Frequency principle in broad learning system. *IEEE Transactions on Neural Networks and Learning Systems*, 33(11): 6983–6989.

Chen, L.; Fu, Y.; Gu, L.; Yan, C.; Harada, T.; and Huang, G. 2024. Frequency-aware feature fusion for dense image prediction. *IEEE transactions on pattern analysis and machine intelligence*.

Cheng, B.; Misra, I.; Schwing, A. G.; Kirillov, A.; and Girdhar, R. 2022. Masked-attention mask transformer for universal image segmentation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 1290–1299.

Cheng, B.; Schwing, A.; and Kirillov, A. 2021. Per-pixel classification is not all you need for semantic segmentation. *Advances in neural information processing systems*, 34: 17864–17875.

Ding, L.; Zhou, R.; Yu, T.; Yang, H.; He, X.; Gao, H.; Wang, J.; Yuan, Y.; Wang, J.; Wang, Z.; et al. 2024. Lunar rock investigation and tri-aspect characterization of lunar farside regolith by a digital twin. *Nature Communications*, 15(1): 2098.

Fang, Z.; Chen, Y.; Wang, Y.; Wang, Z.; Ji, X.; and Zhang, Y. 2023. Weakly-supervised semantic segmentation for histopathology images based on dataset synthesis and feature consistency constraint. In *Proceedings of the AAAI conference on artificial intelligence*, volume 37, 606–613.

Fridovich-Keil, S.; Gontijo Lopes, R.; and Roelofs, R. 2022. Spectral bias in practice: The role of function frequency in generalization. *Advances in Neural Information Processing Systems*, 35: 7368–7382.

Fu, Y.; Lou, M.; and Yu, Y. 2025. SegMAN: Omni-scale context modeling with state space models and local attention for semantic segmentation. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, 19077–19087.

Fu, Z.; Yang, H.; So, A. M.-C.; Lam, W.; Bing, L.; and Collier, N. 2023. On the effectiveness of parameter-efficient fine-tuning. In *Proceedings of the AAAI conference on artificial intelligence*, volume 37, 12799–12807.

Golub, G. H.; Hoffman, A.; and Stewart, G. W. 1987. A generalization of the Eckart-Young-Mirsky matrix approximation theorem. *Linear Algebra and its applications*, 88: 317–327.

Hsu, H.-W.; Wang, X.; Carroll, A.; Hood, N.; and Horányi, M. 2022. Fine-grained regolith loss on sub-km asteroids. *Nature Astronomy*, 6(9): 1043–1050.

Hu, E. J.; yelong shen; Wallis, P.; Allen-Zhu, Z.; Li, Y.; Wang, S.; Wang, L.; and Chen, W. 2022. LoRA: Low-Rank Adaptation of Large Language Models. In *International Conference on Learning Representations*.

Isensee, F.; Jaeger, P. F.; Kohl, S. A.; Petersen, J.; and Maier-Hein, K. H. 2021. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nature methods*, 18(2): 203–211.

Jacot, A.; Gabriel, F.; and Hongler, C. 2018. Neural tangent kernel: Convergence and generalization in neural networks. *Advances in neural information processing systems*, 31.

Jiang, C.; Zhu, F.; Chen, X.; Zhu, J.; Zheng, B.; Wang, Y.; and Zhang, Z. 2025. DMLoRA: Dynamic Multi-Subspace Low-Rank Adaptation. In *Companion Proceedings of the ACM on Web Conference 2025*, 1038–1042.

Junjie, H.; Zheng, M.; Yuzhu, W.; Yujian, B.; Yizhe, W.; Zhongbin, S.; and Lifeng, G. 2025. YOLOv8-DDS: A lightweight model based on pruning and distillation for early detection of root mold in barley seedling. *Information Processing in Agriculture*.

Kiessling, J.; and Thor, F. 2022. A computable definition of the spectral bias. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, 7168–7175.

Kirillov, A.; Mintun, E.; Ravi, N.; Mao, H.; Rolland, C.; Gustafson, L.; Xiao, T.; Whitehead, S.; Berg, A. C.; Lo, W.-Y.; et al. 2023. Segment anything. In *Proceedings of the IEEE/CVF international conference on computer vision*, 4015–4026.

Li, S.; Dan, M.; Chu, Y.; Yu, J.; Zhao, Y.; and Zhao, P. 2025. RetiDiff: Diffusion-Based Synthesis of Retinal OCT Images for Enhanced Segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 516–525. Springer.

Liu, H.; Tam, D.; Muqeeth, M.; Mohta, J.; Huang, T.; Bansal, M.; and Raffel, C. A. 2022. Few-shot parameter-efficient fine-tuning is better and cheaper than in-context learning. *Advances in Neural Information Processing Systems*, 35: 1950–1965.

Liu, Y.; Guan, Y.; Zhang, Y.; Rossman, G. R.; Eiler, J. M.; and Taylor, L. A. 2012. Direct measurement of hydroxyl in the lunar regolith and the origin of lunar surface water. *Nature Geoscience*, 5(11): 779–782.

Lu, X.; Chen, J.; Ling, Z.; Liu, C.; Fu, X.; Qiao, L.; Zhang, J.; Cao, H.; Liu, J.; He, Z.; et al. 2023. Mature lunar soils from Fe-rich and young mare basalts in the Chang'e-5 regolith samples. *Nature Astronomy*, 7(2): 142–151.

Luo, T.; Ma, Z.; Xu, Z.-Q. J.; and Zhang, Y. 2022. On the exact computation of linear frequency principle dynamics and its generalization. *SIAM Journal on Mathematics of Data Science*, 4(4): 1272–1292.

Ma, J.; He, Y.; Li, F.; Han, L.; You, C.; and Wang, B. 2024. Segment anything in medical images. *Nature Communications*, 15(1): 654.

Ma, J.; Li, F.; and Wang, B. 2024. U-mamba: Enhancing long-range dependency for biomedical image segmentation. *arXiv preprint arXiv:2401.04722*.

Mazurowski, M. A.; Dong, H.; Gu, H.; Yang, J.; Konz, N.; and Zhang, Y. 2023. Segment anything model for medical image analysis: an experimental study. *Medical Image Analysis*, 89: 102918.

Rahaman, N.; Baratin, A.; Arpit, D.; Draxler, F.; Lin, M.; Hamprecht, F.; Bengio, Y.; and Courville, A. 2019. On the spectral bias of neural networks. In *International conference on machine learning*, 5301–5310. PMLR.

Shi, Z.; Mettes, P.; Maji, S.; and Snoek, C. G. 2022. On measuring and controlling the spectral bias of the deep image prior. *International Journal of Computer Vision*, 130(4): 885–908.

Wang, Z.; Zhang, Y.; Wang, Y.; Cai, L.; and Zhang, Y. 2024. Dynamic pseudo label optimization in point-supervised nuclei segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 220–230. Springer.

Xie, E.; Wang, W.; Yu, Z.; Anandkumar, A.; Alvarez, J. M.; and Luo, P. 2021. SegFormer: Simple and efficient design for semantic segmentation with transformers. *Advances in neural information processing systems*, 34: 12077–12090.

Xu, K.; Qin, M.; Sun, F.; Wang, Y.; Chen, Y.-K.; and Ren, F. 2020. Learning in the frequency domain. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 1740–1749.

Yan, H.; Wu, M.; and Zhang, C. 2024. Multi-Scale Representations by Varying Window Attention for Semantic Segmentation. In *International Conference on Representation Learning*, volume 2024, 9895–9911.

Yang, S.; Chen, Y.; Tian, Z.; Wang, C.; Li, J.; Yu, B.; and Jia, J. 2025a. VisionZip: Longer is Better but Not Necessary in Vision Language Models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 19792–19802.

Yang, S.; Liu, J.; Zhang, R.; Pan, M.; Guo, Z.; Li, X.; Chen, Z.; Gao, P.; Li, H.; Guo, Y.; et al. 2025b. Lidar-llm: Exploring the potential of large language models for 3d lidar understanding. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, 9247–9255.

Yu, H.; Cho, Y.; Kang, B.; Moon, S.; Kong, K.; and Kang, S.-J. 2024. Embedding-free transformer with inference spatial reduction for efficient semantic segmentation. In *European Conference on Computer Vision*, 92–110. Springer.

Zhang, J.; Zhou, B.; Lin, Y.; Zhu, M.-H.; Song, H.; Dong, Z.; Gao, Y.; Di, K.; Yang, W.; Lin, H.; et al. 2021. Lunar regolith and substructure at Chang'E-4 landing site in South Pole–Aitken basin. *Nature Astronomy*, 5(1): 25–30.

Zhang, L.; Sun, B.; Cai, L.; Wang, Y.; Zhang, Y.; Jiang, S.; Zhang, K.; and Zhang, Y. 2025a. Counting by Points: Density-Guided Weakly-Supervised Nuclei Segmentation in Histopathological Images. In *Proceedings of the 33rd ACM International Conference on Multimedia*, 2900–2908.

Zhang, M.; Fa, W.; and Jia, B. 2025. Provenance and evolution of lunar regolith at the Chang'e-6 sampling site. *Nature Astronomy*, 1–11.

Zhang, T.; Xu, K.; Yao, Z.; Ding, X.; Zhao, Z.; Hou, X.; Pang, Y.; Lai, X.; Zhang, W.; Liu, S.; et al. 2019. The progress of extraterrestrial regolith-sampling robots. *Nature Astronomy*, 3(6): 487–497.

Zhang, Y.; Fang, Z.; Wang, Y.; Zhang, L.; Guan, X.; and Zhang, Y. 2025b. Category prompt mamba network for nuclei segmentation and classification. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, 10284–10292.

Zhang, Y.; Wang, Y.; Fang, Z.; Bian, H.; Cai, L.; Wang, Z.; and Zhang, Y. 2024. Dawn: Domain-adaptive weakly supervised nuclei segmentation via cross-task interactions. *IEEE Transactions on Circuits and Systems for Video Technology*.

Zhang, Y.; Zhou, Y.; Wang, Y.; Xiao, J.; Wang, Z.; Zhang, Y.; and Chen, J. 2025c. The Four Color Theorem for Cell Instance Segmentation. In *Forty-second International Conference on Machine Learning*.

Zhao, H.; Meng, H.; Yang, D.; Xie, X.; Wu, X.; Li, Q.; and Niu, J. 2024. GuidedNet: Semi-Supervised Multi-Organ Segmentation via Labeled Data Guide Unlabeled Data. In *Proceedings of the 32nd ACM International Conference on Multimedia*, MM '24, 886–895.

Zhao, H.; Niu, J.; Liu, X.; Xie, X.; Kuang, L.; Yang, H.; Dai, B.; Meng, H.; and Wang, Y. 2025. Keep Your Friends Close, and Your Enemies Farther: Distance-aware Voxel-wise Contrastive Learning for Semi-supervised Multi-organ Segmentation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 21832–21842.

Zheng, Y.; Zhong, B.; Liang, Q.; Li, N.; and Song, S. 2025. Decoupled Spatio-Temporal Consistency Learning for Self-Supervised Tracking. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, 10635–10643.