National Aeronautics and Space Administration

FY24 Strategic University Research Partnership (SURP)

Lowering the Barriers to Planetary Science Studies with a Large Mars Model

Principal Investigator: Umaa Rebbapragada (398)
JPL Co-Investigators: Steven Lu (398), Serina Diniega (322)
Arizona State University: Hannah Kerner (Assistant Professor), Mirali Purohit (Ph.D. Student)

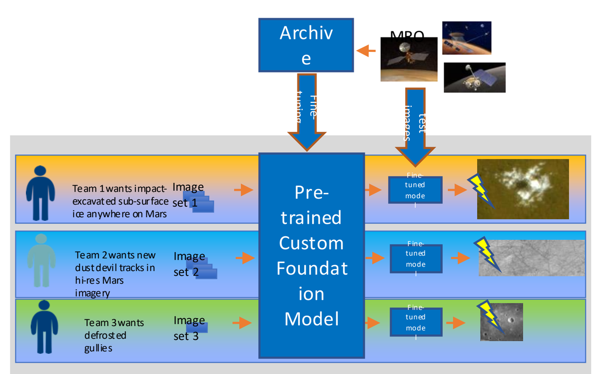
A Foundation Model for Mars

What are Geospatial Foundation Models Useful For?

The deep learning revolution brought forth neural network architectures that were trained from large corpuses of labeled digital photography (e.g., cats, dogs, chairs).

Geospatial Foundation Models are neural networks that are trained through self-supervision [1,2] using large archives of unlabeled geospatial data. They promise:

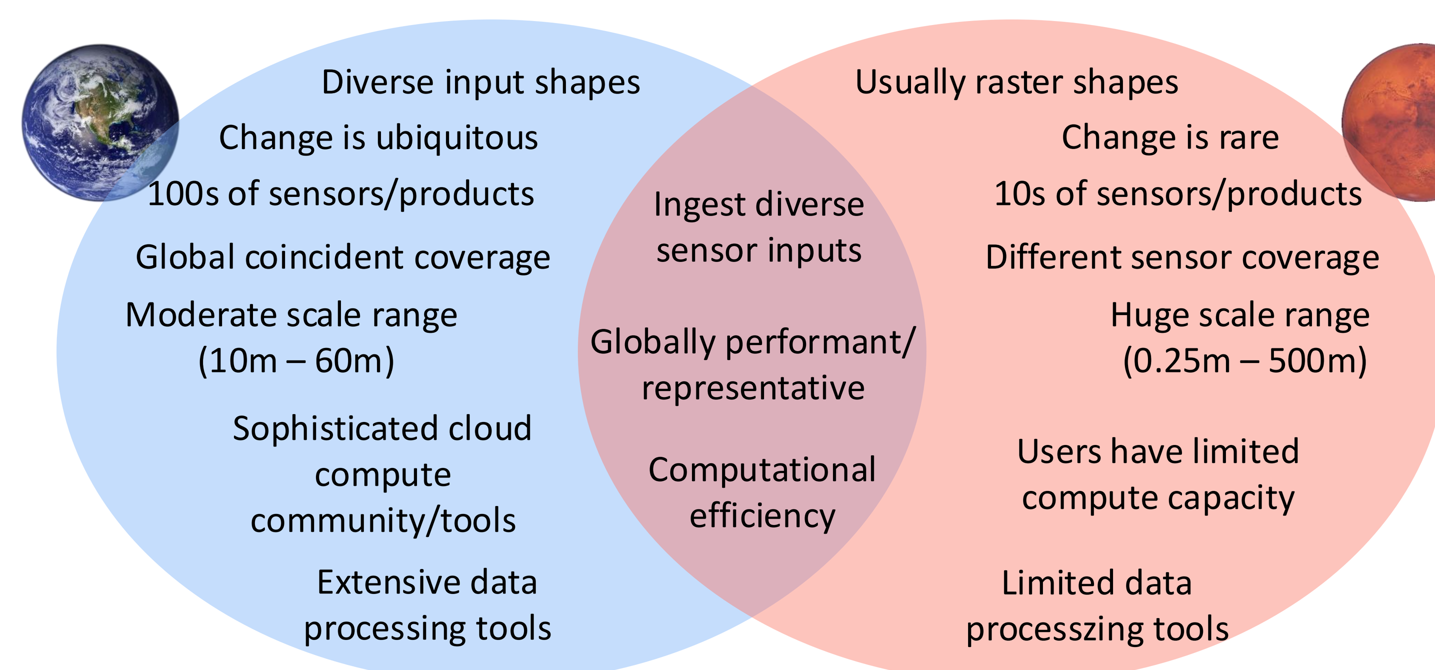
- Better performance
- Large reduction in number of training images needed for fine-tuning
- Increased number of ”downstream” tasks supported => increased science



There have been hundreds of geospatial FMs created from Earth-observed data. See [3] for a comprehensive list.

Mars has different sensor coverage and science infrastructure support. FM strategies for Earth may not apply.

Building Foundation Models for Earth vs. Mars



Question: Should we invest in custom FMs for Mars science?

Year 2 Progress

We continued attempts to train a competent Mars FM with promising results. In Year 2 we matured experiments utilizing a **model-editing approach** [4] that strategically combines single-sensor models.

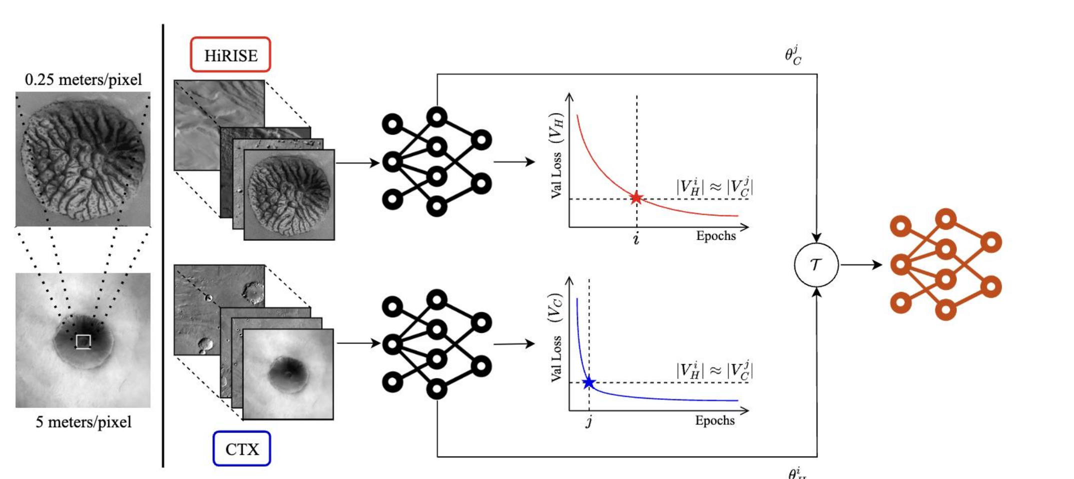


Figure 1. Overview of proposed approach for constructing a Mars orbital FM using model editing with EVL. Models are trained separately on HiRISE (0.25 m/pixel) and CTX (5 m/pixel) data. Then, checkpoints are selected where the validation losses align, ensuring compatibility for merging. The models on selected checkpoints are then merged using task arithmetic operation ($\mathcal{T}$) [23].

Results show that a model editing approach (**τ**) beats models built from either HiRISE (H) and CTX (C) alone, or a model trained simultaneously on both (H+C) on the HiRISENet downstream task.

	HiRISENet			DoMars16		
	ResNet	EfficientNet	ViT	ResNet	EfficientNet	ViT
Zero	0.63 ± 0.02	0.60 ± 0.06	0.85 ± 0.01	0.80 ± 0.01	0.79 ± 0.01	0.77 ± 0.01
ImageNet	0.60 ± 0.01	0.65 ± 0.00	$\mathbf{0.92 \pm 0.01}$	0.66 ± 0.04	0.65 ± 0.01	$\mathbf{0.93 \pm 0.00}$
H	0.75 ± 0.04	0.88 ± 0.00	0.91 ± 0.01	0.87 ± 0.00	$\mathbf{0.92 \pm 0.01}$	0.90 ± 0.00
C	0.73 ± 0.03	0.85 ± 0.02	0.91 ± 0.01	0.83 ± 0.02	$\mathbf{0.92 \pm 0.00}$	0.91 ± 0.00
H + C	0.72 ± 0.03	0.81 ± 0.02	0.91 ± 0.01	0.83 ± 0.01	0.90 ± 0.01	0.90 ± 0.00
$\mathcal{T}(H_{EVL}, C_{EVL})$	$\mathbf{0.82 \pm 0.04}$	$\mathbf{0.90 \pm 0.02}$	$\mathbf{0.92 \pm 0.01}$	$\mathbf{0.91 \pm 0.00}$	$\mathbf{0.92 \pm 0.01}$	0.92 ± 0.00

Table 1. Comparison of the proposed approach with different baselines. The proposed approach ($\mathcal{T}(H_{EVL}, C_{EVL})$) outperforms or performs on par with the other baselines, demonstrating its effectiveness in handling multi-resolution data. Results are reported as *mean ± standard deviation* of the F1-score on HiRISENet and DoMars16.

We also created the first Mars benchmark designed to systematically evaluate models across a broad range of Mars-related tasks using both orbital and surface imagery. **Mars-Bench** comprises 20 datasets spanning classification, segmentation, and object detection, focused on key geologic features such as craters, cones, boulders, and frost.

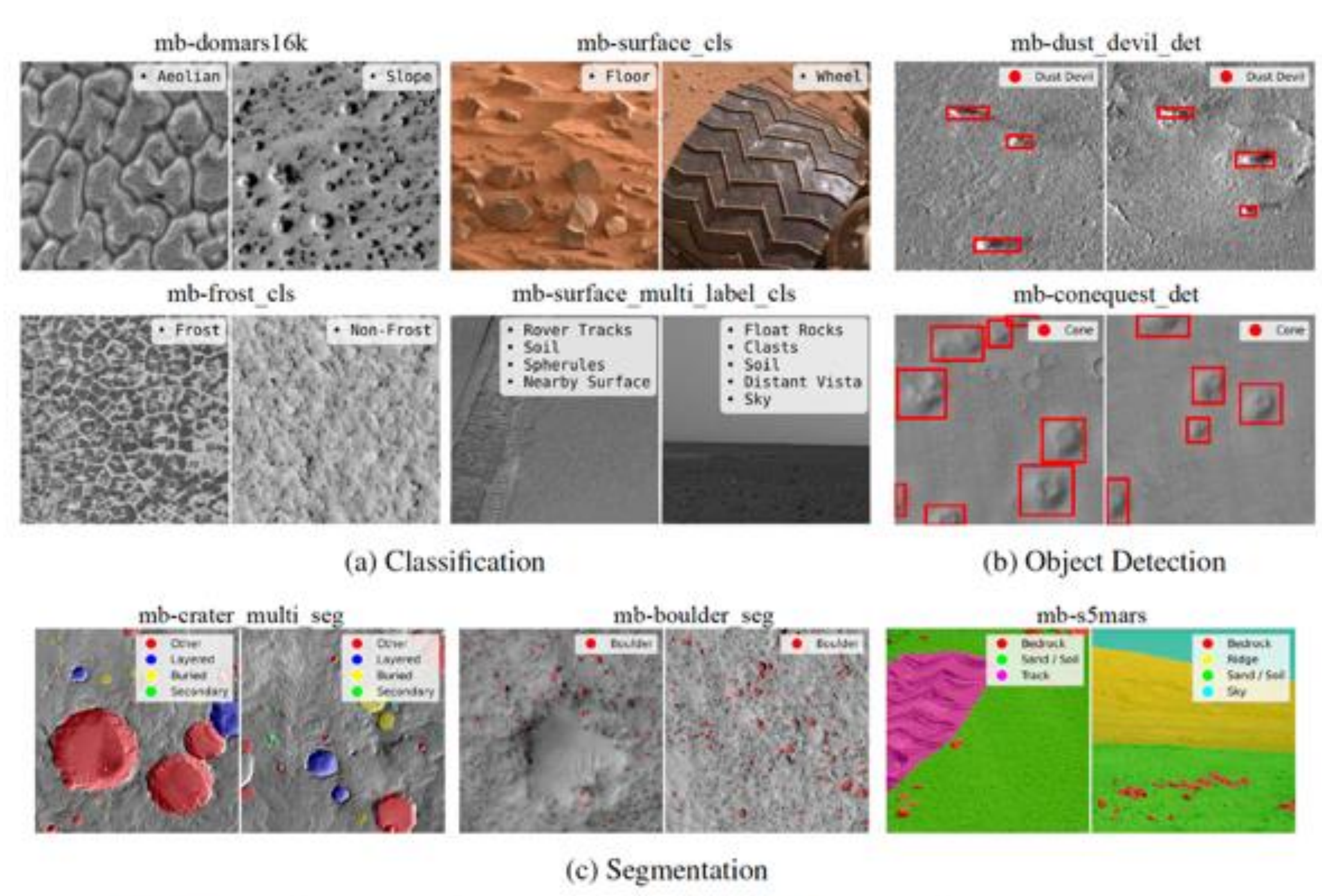


Figure 1: Representative samples from selected Mars-Bench datasets, from all three task categories.

Year 3 Focus

- Experiment with larger model sizes and pre-training samples
- Expand validation to include all Mars-Bench use cases
- Adding pre-training data from THEMIS and CRISM
- In-depth science analysis on select use cases showing science benefits reaped from the use of the best-performing Mars FM
- Deploying LMM models onto the PDS Imaging Node

National Aeronautics and Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

www.nasa.gov

RPD-005 Clearance Number: CL#25-4519

Publications & Presentations

- M. Purohit, Y. Lu, U. Rebbapragada, H. Kerner, “Model-Editing for Mars: A Foundation Model for Multi-Resolution Orbital Science”, Submitted to CVPR 2026.
- M. Purohit, B. Gajera, V. Malaviya., I. Mehta, K. Kasodekar, J. Adler, Y. Lu, U. Rebbapragada, H. Kerner, “Mars-Bench: A Benchmark for Evaluating Foundation Models for Mars Science Tasks”, In Proc. of NeurIPS Datasets and Benchmarks Track and ML for Physical Sciences Workshop, 2025
- U. Rebbapragada, M. Purohit, K. H. R. Kerner, and S. Lu (2024), Investigating the Benefits of a Large Mars Model, Abstract IN51B-04 presented at AGU Fall Meeting 2024, Washington, D.C., 9-13 Dec.
- M. Purohit, S. Lu, S. Diniega, U. Rebbapragada, H. Kerner, “Investigating the Benefits of Foundation Models for Mars Science, In 10th Int. Conf on Mars, 2024, Pasadena, CA

References:

1. He, K., Chen, X., Xie, S., Li, Y., Dollár, P., & Girshick, R. (2022). Masked Autoencoders Are Scalable Vision Learners. In Proc. of CVPR 2022
2. Chen, T., Kornblith, S., Norouzi, M., & Hinton, G. (2020). A Simple Framework for Contrastive Learning of Visual Representations. In Proc. of ICML 2020.
3. https://github.com/Jack-bo1220/Awesome-Remote-Sensing-Foundation-Models?tab=readme-ov-file
4. G. Ilharco, M. T. Ribeiro, M. Wortsman, S. Gururangan, S. Schmidt, H. Hajishirzi, and A. Farhadi. Editing models with task arithmetic. preprint arXiv:2212.04089, 2022