

Mosam Dabhi

PH.D. STUDENT

212 Elliot Dunlap Smith Hall, Carnegie Mellon University, Pittsburgh, PA, USA

☎ (+1) 412-726-4454 | ✉ mosam@cmu.edu | 🏠 mosamdabhi.github.io | 📄 mosam-dabhi-9395b09a/ | 🐦 @mosamdabhi | 🏠 Mosam Dabhi

Research Interests

My PhD research is dedicated to advancing 3D vision and multimodal applications by developing geometric foundation models. These efforts aim to enhance scene understanding and human-computer interaction, pushing the boundaries of general AI, out-of-distribution (OOD) reasoning, and 3D vision graphics. This research lays the foundation for novel approaches in generative AI, robotics – enabling robust applications across a wide range of disciplines.

Machine Learning Self-supervised learning, Neural Priors, Generative modeling
Computer vision 3D Vision, Multi-view geometry, Autolabeling
Robotics SLAM, Scene Graphs

Education

Carnegie Mellon University

PH.D. IN ROBOTICS

• **Advisor:** Simon Lucey and Laszlo Attila Jeni

Pittsburgh, PA, USA

Aug. 2021 - Present

Carnegie Mellon University

M.S. IN ROBOTICS

• **Advisor:** Simon Lucey
• Thesis: Multi-view NRSfM: Affordable Setup for High-Fidelity 3D Reconstruction

Pittsburgh, PA, USA

Aug. 2019 - May 2021

National Institute of Technology

B.TECH. IN ELECTRONICS AND COMMUNICATION ENGINEERING

Surat, Gujarat, India

Aug. 2013 - May 2017

Research Experiences

Carnegie Mellon University

GRADUATE RESEARCH ASSISTANT

• **Advisors:** Simon Lucey, Laszlo Attila Jeni
• Pioneering the concept of Geometric Foundation Models to enhance 3D vision for AI, reasoning, and action capabilities, focusing on integrating geometry with foundational models.
• Developed and introduced models including MV-NRSfM [1], MBW [2], 3D-LFM [3], and S-GAM [4], advancing scalable 3D label generation and ground truth in natural settings.
• MBW enables auto-labeling at scale with minimal supervision, using temporal and spatial neural priors for landmark labeling across dimensions.
• 3D-LFM presents a universal lifting model from 2D to 3D, employing graph-based transformers to ensure consistency across diverse object categories without requiring specific knowledge.
• S-GAM, offers a novel approach to creating dynamic 3D models of deformable objects from monocular footage, using inverse kinematics and Gaussian splatting for template-free deformable rigs.

Pittsburgh, PA, USA

Aug. 2019 - Present

Apple AI Research

RESEARCH SCIENTIST INTERN

• Multimodal time-series modeling, shaping the foundation for lightweight AI foundation models.

Cupertino, CA, USA

June 2023 - Present

Apple, Inc.

RESEARCH SCIENTIST INTERN

• Developed few-shot learning techniques and Out-Of-Distribution detection algorithms.
• Auto-labeling in the wild, driving substantial financial savings by auto-generating 3D labels for computer vision applications.

Sunnyvale, CA, USA

May 2022 - Aug. 2022

Apple, Inc.

RESEARCH INTERN

• Active learning and self-supervised learning strategies.

Sunnyvale, CA, USA

May 2021 - Aug. 2021

Apple, Inc.

RESEARCH INTERN

Sunnyvale, CA, USA

May 2020 - Aug. 2020

- Foundational work on machine learning based multi-view 3D geometry, paving the way for affordable setups to generate 3D groundtruth for computer vision applications. [1]

Carnegie Mellon University

RESEARCH ASSISTANT

Pittsburgh, PA, USA

May 2017 - May. 2019

- **Advisor:** Nathan Michael
- Robotic exploration and mapping in real-time for search and rescue operations, enabling superior robot-to-robot communication on extra-terrestrial and sub-terrestrial surfaces [5].
- Achieved flights in challenging, GPS-denied terrains, hitting accelerations over 12 m/s^2 [6]
- Experience-driven Model Predictive Control (EPC) tailored for platforms with computational constraints [7, 8].
- Motion and path planning in cluttered environments through mixed-integer programming [9].

Indian Institute of Science

RESEARCH INTERN

Bangalore, India

May 2015 - Jul. 2015

- **Advisor:** Prasanta Kumar Ghosh
- Home automation through an HMM model harnessing MFCC features and speaking rate analysis.
- Speech recognition with a focus on digit identification utilizing the aforementioned model.

Publications

- [1] **Mosam Dabhi**, Chaoyang Wang, Kunal Saluja, Laszlo Jeni, Ian Fasel, and Simon Lucey. High Fidelity 3D Reconstructions with Limited Physical Views. In *2021 International Conference on 3D Vision (3DV)*. IEEE, 2021.
- [2] **Mosam Dabhi**, Chaoyang Wang, Tim Clifford, Laszlo Jeni, Ian Fasel, and Simon Lucey. Multi-view Bootstrapping in the Wild. In *Thirty-sixth Conference on Neural Information Processing Systems Datasets and Benchmarks Track*. NeurIPS, 2022.
- [3] **Mosam Dabhi**, Laszlo A Jeni, and Simon Lucey. 3d-lfm: Lifting foundation model. *2024 Computer Vision and Pattern Recognition (CVPR)*, 2024.
- [4] **Mosam Dabhi**, Laszlo A Jeni, and Simon Lucey. S-gam: Skeletal-guided gaussian articulated models. *In submission*, 2024.
- [5] Wennie Tabib, Kshitij Goel, John Yao, **Mosam Dabhi**, Curtis Boirum, and Nathan Michael. Real-Time Information-Theoretic Exploration with Gaussian Mixture Model Maps. In *Robotics: Science and Systems*, 2019.
- [6] Alex Spitzer, Xuning Yang, John Yao, Aditya Dhawale, Kshitij Goel, **Mosam Dabhi**, Matt Collins, Curtis Boirum, and Nathan Michael. Fast and agile vision-based flight with teleoperation and collision avoidance on a multirotor. In *International Symposium on Experimental Robotics*, pages 524–535. Springer, 2018.
- [7] **Mosam Dabhi**, Alexander Spitzer, and Nathan Michael. Aggressive Flight Performance using Robust Experience-driven Predictive Control Strategies: Experimentation and Analysis. Technical Report CMU-RI-TR-19-08, Carnegie Mellon University, Pittsburgh, PA, June 2019.
- [8] **Mosam Dabhi**, Vishnu R Desaraju, and Nathan Michael. Evaluation of Explicit Experience-driven Predictive Control on a Computationally Constrained Platform. Technical report, Carnegie Mellon University, Pittsburgh, PA, June 2017.
- [9] **Mosam Dabhi**, Vishnu Desaraju, and Nathan Michael. Planning Aggressive, Dynamically Feasible and Optimal Trajectories for Autonomous Vehicles in Cluttered Environments using Mixed Integer Programming. Technical report, Carnegie Mellon University, Pittsburgh, PA, 2016.

Honors & Awards

2019 - Now	Apple Research Grant , Apple Inc.	<i>Sunnyvale, CA, USA</i>
2017	Research Scholarship , Federation of Indian Chambers of Commerce & Industry	<i>Pittsburgh, PA, USA</i>
2016 - 17	Summer Scholar , Robotics Institute Summer Scholar	<i>Pittsburgh, PA, USA</i>
2016	Undergraduate thesis funding , TEQIP Award, MHRD, Government of India	<i>Surat, India</i>

Academic Services

2022 - 2024	Conference Paper Reviewer , NeurIPS; CVPR; ICCV; ECCV; WACV	
2021	Conference Paper Reviewer , IROS	
2020	Conference Paper Reviewer , International Conference on Humanoid Robots	
2021 - 2024	Admissions Committee , Master of Science, Robotics, CMU (MSR)	<i>Pittsburgh, PA, USA</i>
2024	Admissions Committee , Master of Science, Computer Vision, CMU (MSCV)	<i>Pittsburgh, PA, USA</i>
2023	M.S. in Robotics Thesis Committee , Examinee: Heng Yu, Aarush Gupta	<i>Pittsburgh, PA, USA</i>
2017 - 2019	Admissions & Admin. Committee , Robotics Institute Summer Scholars, CMU (RISS)	<i>Pittsburgh, PA, USA</i>

Teaching Experiences

Carnegie Mellon University

Pittsburgh, PA, USA

TEACHING ASSISTANT

- **Spring 2022** : Robot Localization and Mapping with Prof. Michael Kaess
- **Fall 2022** : Geometry-Based Methods in Vision with Prof. Shubham Tulsiani

Selected coursework

Carnegie Mellon University

Pittsburgh, PA, USA

LEARNING FOR 3D VISION (**A+**), COMPUTER VISION (**A**), ADVANCED MACHINE LEARNING (**A**), CONVEX OPTIMIZATION

(**A**), ROBOT LOCALIZATION AND MAPPING (**A+**), MATHEMATICS FUNDAMENTALS FOR ROBOTICS (**A**), KINEMATICS,

DYNAMICS, AND CONTROLS (**A+**), ETHICS IN ROBOTICS (**A**)

Aug. 2019 - Present

Proficient Skills

Programming languages

PRIMARY: PYTHON, C/C++, \LaTeX , MATLAB

SECONDARY: CUDA, LUA, HTML, JAVASCRIPT

Software libraries

PRIMARY: PYTORCH, TENSORFLOW, BLENDER, COLMAP

SECONDARY: OPENAIGYM, TORCH, CAFFE, OPENCV, VLFEAT, PTHREAD