Hardik Shah

hardik01shah@gmail.com
hardik01shah.github.io
linkedin.com/in/hardik01shah

Education

2023–Present **MSc in Computer Science**, *ETH Zurich*, **[5.70/6]** (Transcript).

2019–2023 **B.E. in Computer Science Engineering**, *BITS Pilani, Goa*, **[9.64/10]** (Transcript).
w/ Minor in Data Science

- **Institute Rank 6** in a batch of 900 students
- Recipient of BITS Goa **Merit Scholarship** for all 8 semesters awarded to **top 10** students across all departments– 100% tuition fee waiver.

Research Experience

Institution **Scandit AG, Zurich** [Jul'24-Present]
Computer Vision Intern

Project Title Lightweight Interest Point Detection and Matching for SLAM-Based AR Visualization

Description Exploring the feasibility of replacing traditional keypoint detectors with **learned detection and matching methods** in the tracking pipeline of Scandit's MatrixScan product, which leverages SLAM for AR-based inventory management on resource-constrained devices. Optimizing training paradigms, and exploring lightweight architectures for efficient on-device inference along with custom evaluation benchmarks.

Institution **Robotics and Perception Group, ETH Zurich** [Feb'24-Present]
Graduate Student Researcher, Supervisor: Prof. Dr. Davide Scaramuzza

Project Title Vision-Language Grounded Semantic Exploration using CLIP Features

Description Developed a unified CLIP-based representation **combining geometry and semantics** for **Object Goal Navigation** in unseen environments. Improved navigation efficiency by predicting unexplored regions of the occupancy map accompanied with a closed-loop benchmark for occupancy prediction. Demonstrated that integrating CLIP feature based representation along with **frontier-based exploration** outperforms recent segmentation-based methods, highlighting its robust generalization and zero-shot capabilities.

Institution **Google Research** [Jan'23-Jun'23]
Student Researcher, Machine Learning and Optimization Team

Project Title End-to-End Neural Network Compression via $\frac{l_1}{l_2}$ latency surrogates [**CVPRW '24**]

Description Developed a versatile **neural network compression** toolbox that optimizes for the model's FLOPs via a novel $\frac{l_1}{l_2}$ latency surrogate in various compression methods, including **pruning** and **low-rank factorization**. Extended the FLOPs regularizer to optimize over actual on-device latency using a latency look-up table of the target device. Achieved 11% reduction in latency on Pixel-6, and 15% reduction in FLOPs in compressing MobileNetV3 on ImageNet-1K without drop in accuracy, while still requiring $3\times$ less training compute than SOTA NAS techniques.

Institution **Robot Vision Lab, Karlsruhe University of Applied Sciences** [May'22-Aug'22]
Summer Research Intern, DAAD WISE Scholarship, Supervisor: ***Prof. Niclas Zeller***

Project Title Camera based 3D Dense Reconstruction for Mobile Robots [Website] [Code]

Description Designed an end to end pipeline for multi-view stereo dense 3D reconstruction from a handheld stereo-camera(**Intel RealSense**) that outputs stable dense pointclouds. In particular, extracted and tracked keyframe poses and keypoints from BASALT-VIO. Encoded information from multiple camera views in a cost volume used for self-supervised training of a **U-Net** adapted architecture design-MonoRec. Benchmarking of trajectory estimation done on rectified TUM-VI dataset before deployment.

Research Projects

Title **POLD2: Unified Point and Line Feature Detection and Description** [Spring '24]
Supervisor - Prof. Dr. Marc Pollefeys, Computer Vision and Geometry Group, ETH Zurich

Description Developed POLD2, a deep learning-based pipeline that jointly detects and describes both **point and line features** in images, optimizing feature extraction for 3D vision tasks like SLAM and pose estimation. By sharing computations between points and lines, **POLD2 achieves a significant 9.5x speedup** in inference time compared to traditional methods, while maintaining comparable accuracy. The solution integrates a **novel line extraction module** and demonstrates robust performance across multiple challenging datasets in tasks such as homography estimation.

Title **A Monocular Visual Odometry Pipeline** [Report] [Code] [Demo] [Fall '23]
For the course Vision Algorithms for Mobile Robotics, ETH Zurich

Description Implemented a comprehensive Monocular VO pipeline in Python. Landmark and camera pose initialization, alongside a continuous pipeline for **key point tracking** and **landmark triangulation**. Additional functionalities include **bootstrapping using 2D↔2D and 2D↔3D estimates** to mitigate tracking losses and **local bundle adjustment** for trajectory refinement. Evaluated on KITTI, Malaga, Parking, and a custom dataset covering a 400m path in Zurich captured with a Realsense D435i camera.

Title **Project Kratos, A Mars Rover** [Website] [Code] [Demo Video] [2020 - 2022]
Autonomous Subsystem lead

Description Development of a mars rover as part of the University Rover Challenge(URC). Team lead of the Autonomous Subsystem, responsible for autonomous traversal. Program design, implementation and deployment of mapping, planning and control nodes on Jetson Xavier for obstacle avoidance and object tracking(arrows, ARTags).
Path planning and Perception-Implemented A*, RRT*, Dijkstra's on a 4-adjacency grid graph obtained from binary occupancy grid generated by ZED2i camera;
Tracking-Employed transfer learning on **YOLOv3**, **Mask R-CNN** algorithms for arrow detection. Achieved ROS integration using darknet_ROS(20 fps).
Control-Wrote a custom P-controller based **visual servoing** algorithm for following arrows and ARTags.

Technical Strengths

Languages Python, C++, C, JAVA, C#, MATLAB, Latex, HTML, CSS

Softwares Pytorch, Tensorflow, Keras, JAX, Numpy, OpenCV, Unity, Gazebo, Verilog, Robot Operating System (ROS), AutoCAD, Android Studio, Git

Relevant Coursework

ETH Zurich Probabilistic Artificial Intelligence, Information Security, Computer Vision, Planning and Decision Making for Autonomous Robots, Vision Algorithms for Mobile Robotics, 3DVision, Machine Perception, *Big Data, *Design of Parallel and High Performance Computing **in progress*

Awards and Achievements

2022 University Rover Challenge, Utah: Project Kratos secured **1st** position in India
2022 Anatolian Rover Challenge, Turkey: Project Kratos secured **2nd** position globally
2022 Recipient of **DAAD WISE** research scholarship(Germany)
2022 Recipient of **MITACS Globalink** research scholarship(Canada)
2022 Recipient of the Singapore International Pre-Graduate Award (**SIPGA**)

Other Projects and Experiences

Title **Robotics Summer School [Website]** [Jul'24]
RobotX Initiative, ETH Zurich

Description Attended the Robotics Summer School, organized by RobotX, with 50 students from around the world. Deploying autonomous software on wheeled robots for search and rescue missions. Hands-on tutorials in key robot planning modules, including state estimation, SLAM, exploration path planning, motion planning, and object detection.

Institution **Google Research** [Aug'22-Dec'22]
Student Researcher, Machine Learning and Optimization Team
Undergraduate Thesis, Supervisor: ***Dr. Prateek Jain*** *(Sr. Staff Research Scientist, Google)*

Project Title Machine Learning Optimization for object detection on low-end smartphones.

Description Optimized on-device latency of large **character recognition models** used for OCR tasks in Google products like Lens, for **faster on-device inference** while maintaining accuracy. Improved parameter efficiency for OCR tasks by extending Singular Value Decomposition(SVD) techniques and Orthogonal Matching Pursuit(OMP) on 1x1 convolution kernels. Experimentally observed constant performance with 33% less parameters and 10% reduction in latency.

Title **RGB Guided Sparse Depth Completion** [Jun'21 - Dec'23]
Prof. Sravan Danda, Prof. Aditya Challa, BITS Goa

Description Existing methods for **depth completion and estimation** tend to overfit with very less generalization across datasets. Focused on developing methods to identify statistical patterns in coupled RGB-depth maps. Redefined depth completion as interpolation problem on a grid graph with sparse-depth seed values. Using empirical results from hypotheses testing on LiDAR depth data for seed selection and context-aware **spatial seed propagation**. Achieved comparable results against computationally heavy deep learning based methods on **KITTI dataset**.