

DHRUV PATEL

dhruv.r.patel14@gmail.com | (404)-203-7003 | github.com/dhruv2012 | [linkedin.com/in/dhruv2012](https://www.linkedin.com/in/dhruv2012) | github.com/dhruv2012

EDUCATION

Georgia Institute of Technology <i>Master of Science, Robotics</i>	Aug 2023 - May 2025 GPA: 4.0/4.0
National Institute of Technology (NIT) Surat, India <i>Bachelor of Technology, Electronics and Communication Engineering</i>	July 2016 - July 2020 GPA: 8.4/10.0

WORK EXPERIENCE

Honda Research Institute USA <i>Research Associate Intern</i>	San Jose, CA, USA May 2024 – Aug 2024
---	---

- Curated a custom dataset of driving scenes from California cities for HRI's Autonomous Vehicle (AV) platform and developed perception algorithms for intersection detection and navigation.
- Benchmarked different vision backbones for intersection detection – CNNs, Vision Transformer, YoloV8, multi-modal foundation models such as Grounding DINO, with visual cue-guided YoloV8 achieving the best 86% accuracy and 90% recall on a custom-HRI dataset (10k intersections, 500k images) from distances over 100m.
- Evaluated map-based transformer architectures for 3D virtual lane and trajectory prediction at intersections in Birds-Eye View (BEV) space on NuScenes, OpenLane, and a custom HRI dataset.

Robot Learning and Reasoning Lab (RL2), Georgia Tech <i>Graduate Student Researcher Advisor: Dr. Danfei Xu, Assistant Prof. and Research Scientist - NVIDIA AI</i>	Atlanta, GA, USA Jan 2024 – Present
--	---

- Developing generalizable egocentric bimanual robot manipulation on the ALOHA platform by leveraging scalable human play data from Meta's ARIA glasses, aiming to minimize reliance on resource-intensive robot teleoperation data.
- Implemented transformer-based vision encoder for trajectory prediction based on Action Chunking with Transformers (ACT), and validated a 60% performance increase over base ACT and an 83% increase on toy-bowl place using human play data. Currently validating on complex long-horizon tasks like cloth folding, object sorting, and packing objects into containers.
- Achieved embodiment-agnostic visual representations using techniques such as auxiliary KL divergence loss and masking manipulator.

Computer Vision, Georgia Institute of Technology [Webpage] <i>Graduate Teaching Assistant</i>	Atlanta, GA, USA Aug 2023 – May 2024 Aug 2024 - Present
---	---

- Designing assignments, and offering mentorship in topics encompassing Feature Matching, Geometry, Object Detection, Segmentation, Transformers, GANs, Diffusion Models, and more, to ~200 graduate students.

Google Summer of Code [Webpage] <i>Open-source Software Developer – Unify AI</i>	Mountain View, CA, USA June 2023 - Aug 2023
--	---

- Enhanced GradSLAM, a differentiable SLAM system, by developing multi-framework support across PyTorch, TensorFlow, JAX, and NumPy, expanding its versatility and usability.
- Optimized deployment using graph compiler, enabling seamless integration with JAX for enhanced performance.

Robotics Research Center, IIT Hyderabad <i>Project Associate</i>	Hyderabad, India July 2021 - July 2023
--	--

- **Autonomous Driving Scene Understanding - ZF Friedrichshafen and QUT Robotics, Australia** [[Webpage](#)]
 - Researched ADAS perception (object detection, tracking, and semantic segmentation) for adverse weather conditions.
 - Proposed Gated Differentiable Image Processing (GDIP) framework for object detection in adverse weather conditions.
 - Improved performances by 5.84 and 16 mAPs in foggy and low-lighting weather, over state-of-the-art with a ~3x speedup.
- **UAV-based Assessment of Civil Structures** [[Website](#)]
 - Led a team of 5 in estimating critical structural parameters using UAV-based Visual Remote Sensing, leveraging Structure-from-Motion, state estimation, combined with classical Computer Vision and Deep Learning algorithms.
 - Constructed 3D models, estimated ROIs (storey/window) with an error of 2.3% by fusing UAV's odometry information; quantified occupancy of roof-top objects, and devised a strategy to estimate plan-shape/roof area with an avg. error of 4.7%.
 - Achieved 99.04% accuracy in estimating pounding effect using RANSAC-based plane fitting.
 - Developed an open-source software library (UVRSABI), soon to be adopted by the Govt. of India.

Amdocs <i>Software Engineer</i>	Pune, India Aug 2020 - June 2021
---	--

- Developed cross-functional telecommunication software solutions for Comcast's Orion project (USA).
- Collaborated with global product owners, ensuring end-to-end feature development, integration, and validation with the testing team.
- Technical Stack: Java, ReactJS, SQL, Spring Boot, Maven, and Jenkins.

Swaayatt Robots <i>Research Intern</i>	Bhopal, India April 2020 - July 2020
--	--

- Improved Visual Odometry (VO) and SLAM pipelines for Level-5 Autonomous Vehicles.
- Devised a semantic variant of the Iterative Closest Point (ICP) algorithm outperforming vanilla ICP by 97% (matching loss) and 50% (convergence time), respectively, on the Semantic KITTI dataset.

- Implemented FaceNet, a Deep Learning-based Face Recognition system, using an NN4 variant of the inception network, and validated the system on a custom-made facial image dataset of 25 students.

PROJECTS AND EXTRA-CURRICULAR

Behavior Cloning and Learning Dynamics Models for Robot Manipulation

Jan 2024 - March 2024

- Implemented MLP, RNNs and Diffusion policy variants for manipulation tasks (lift, square and tool hang) in Robomimic. Diffusion Policy achieved faster convergence over RNNs and improved success rates by 24% over MLPs.
- Implemented different learning dynamics models for robot manipulation. Improved Random Shooting success rates by 30% in a pushing task simulation through gradient-based trajectory optimization.

Using Vision-Language Models (VLMs) for Natural Language Human Feedback [WebPage]

Feb 2024 - May 2024

- Developed an interactive robot learning algorithm leveraging Vision-Language Models (VLMs) for natural language human feedback, eliminating the need for designing reward functions for diverse tasks in the meta-world environment.
- Achieved a 1300x improvement in sample efficiency for door/drawer close and button press tasks compared to baseline methods such as RoboCLIP, while maintaining performance comparable to traditional evaluative feedback approaches like TAMER.

UG Project - Autonomous Agricultural Robot (AGRIBOT) [Webpage]

Oct 2019 - June 2020

- Developed autonomy stack for a 4-wheel skid-steer drive, and simulated it in Gazebo using RGB camera, GPS, and IMU.
- Implemented a lightweight encoder-decoder architecture for crop-weed classification task, having 100x lesser parameters than SOTA like UNet.
- Achieved 96.48% accuracy on CWFID and 99.471% mean accuracy, 98.035% mean IoU on the Bonn dataset for crop-weed classification task. Runs with low latency of <2.5 fps (on Nvidia 940MX).

Asia-Pacific Robot Contest – RoboCon 2018, 2019 [Webpage]

June 2017 - June 2019

- Led a 15-person team in RoboCon 2019, building a 4-wheel Holonomic Drive and Quadruped Robot.
- Developed autonomous holonomic drives using line following and odometry through feedback from line sensor, Gyroscope, IMU, and Encoders on Atmel AVR and ARM microcontroller hardware.
- Organized INSIGHT 1.0, a technical symposium with a footfall of 500-plus people.
- Mentored Embedded Systems and Robotics projects such as RFID-based Identification and Wireless control of mobile robots.

SKILLS

Languages and Tools: Python, C/C++, Java, JavaScript, Embedded C/C++, SQL, Bash, CMake, Git, Spring, Jenkins, Docker

Libraries and Frameworks: PyTorch, TensorFlow, Keras, MATLAB, Pandas, NumPy, SciPy, Matplotlib, OpenCV, Point Cloud Library (PCL), Robot Operating System (ROS) 1/2, Gazebo.

AWARDS AND ACHIEVEMENTS

- JNTE Scholarship:** Prestigious abroad education scholarship, joining an elite group of 5600 scholars since 1892, including esteemed Indian Presidents and scientists.
- UAV-based Visual Remote Sensing for Automated Building Inspection (UVRSABI)**
 - Spotlight presentation at the CVCIE Workshop, ECCV 2022.
 - Inaugurated by Dr. S. Velmurugan (Chief Scientist, CRRRI) to deploy in Telangana, India (Sept 2022)
 - A high-impact project (25 selected out of 300+ research projects) at IIIT-H's RnD Showcase 2023.
- Top-contributing employee** award at Amdocs (May 2021).
- Secured **36k INR funding** from the TEQIP-III program (Govt. of India) for **AGRIBOT** and presented it at **ROS Agriculture**.
- Secured **12th** and **13th rank** in **RoboCon 2018 and 2019** respectively, among 100-plus universities.
- Best Working Model - Stirling Engine** at the National Science Day Celebrations, Physical Research Laboratory, India.

PUBLICATIONS

- GDIP: Gated Differentiable Image Processing for Object Detection in Adverse Conditions**
S. Kalwar*, D. Patel*, A. Aanegola, K. R. Konda, S. Garg, and K. M. Krishna, 2023 IEEE International Conference on Robotics and Automation (ICRA), London, United Kingdom, 2023. [Webpage] [Code] [Paper]
- UAV-Based Visual Remote Sensing for Automated Building Inspection**
Srivastava K*, Patel D*, Jha AK, Jha MK, Singh J, Sarvadevabhatla RK, Ramancharla PK, Kandath H, Krishna KM, European Conference on Computer Vision 2022. Cham: Springer Nature Switzerland. [Webpage] [Code] [Paper]
- SRTGAN: Triplet Loss based Generative Adversarial Network for Real-World Super-Resolution**
Patel D*, Jain A*, Bawkar S, Khorasiya M, Prajapati K, Upla K, Raja K, Ramachandra R, Busch C, 7th International Conference on Computer Vision and Image Processing (CVIP) 2022. [Webpage] [Code] [Paper]
- Design of an Autonomous Agriculture Robot for Real Time Weed Detection using CNN**
Patel D*, ShankaraNarayanan H.*, Gandhi M* and Darji A, AVES 2021 conference. [Code] [Paper]