Tushar Vaidya

Leadership, Technology, Research

Hyderabad, India
(+91) 9975862868
tushar5610@gmail.com
www.linkedin.com/in/tushar-vaidya-065aa342

Positions Held

Director - Robotics, May'26 - Present
DreamVu Inc., Philadelphia, PA, USA, Leading the Physical Intelligence group at DreamVu, driving the Physical AI data pipeline and Vision–Language–Action (VLA) model integration for Humanoid Robots, in close collaboration with strategic clients including Google and Genesis.

- Strategic Charter: Motion Retargeting for humanoid robots from human action video data; building World Models (WM) and World Action Models (WAM) via Gaussian Splats pipelines with physics-based rendering on Gaussians; Autocalibration pipeline for egocentric manipulation videos fusing wrist, exo, and ego cameras with object-of-focus tracking
- Key Technologies: Physical AI Data Pipelines, VLA Foundation Models, Motion Retargeting, World Models (WM), World Action Models (WAM), Gaussian Splatting, Physics-Based Gaussian Rendering, Multi-Camera Autocalibration (Wrist/Exo/Ego), Humanoid Robot Integration, Client Engineering with Google & Genesis

Head, - Embodied AI, July'25 - May'26
General Manager, Meril Group, Robotics R&D, Leading the Embodied AI(Physical Intelligence) Research efforts across verticals including Humanoid Robots(Unitree G1 Edu), Aerial Robotics, Autonomous Cars, and Warehouse Robots, with core focus areas in Reinforcement Learning, Vision–Language–Action (VLA) Foundation Models, Sim2Sim and Sim2Real Transfer Learning, Perception, and SLAM.

- Key Technologies: Groot 1.6 , π0-fast, OpenVLA, Diffusion policy learning, Gaussian Splats SLAM, Dense 3D Reconstruction, NeRF, Sensor Fusion, Lidar + IMU + Camera Calibration, VLA models, Robotics Foundation Models, ISSAC Sim Data pipeline

Sr. Research Scientist & Tech Lead, Oct'21–May'25
AI Center Of Excellence (AICoE), Jio, Dual role combining applied research and technical leadership within the 3D Vision group. As **Tech Lead**, managed the Perception team and delivered production-grade SLAM and Dense Mapping microservices; generated novel IPs including GPU-enabled Bundle Adjustment, Multi-Sensor Calibration, and Dense Implicit Mapping pipelines. As **Sr. Research Scientist**, drove original research adapting Robotics Vision and SLAM to Activity Recognition for indoor and outdoor environments.

- Key Technologies: Implicit Learning, Dense 3D Reconstruction, NeRF, Gaussian Splatting, Multi-View Reconstruction, Optimization, SLAM, SfM, Sensor Fusion, Lidar + IMU + Camera Calibration

Consultant, Feb'19-Oct'20
Standard Cognition (standard.ai), *(YC and Soft Bank Funded Uniocorn)*, Adaptation of Robotics Vision and Simultaneous Localization and Mapping techniques to Autonomous Checkout at retails stores of standard cognition, Extrinsic Calibration of Indoor cameras

- Keywords: Multi-View Reconstruction, Optimization, SLAM, SfM

Consultant Nov'17-May'18
ExplorerAI (explorer.ai), Consulting for Inception of Simultaneous Localization and Mapping(SLAM) to generate and optimize city-scale maps for explorer.ai, Also contributing to product pipeline of explorer.ai at openbasemaps.explorer.ai and flux.explorer.ai for automatic calibration methods for multimodal SLAM pipeline.

- Keywords: City-Scale Mapping, SLAM, Loop Closure Detection, Scene Recognition, Siamese Networks, Bi-Directional LSTM, SLAM, Mapping, LIDAR, G2O, GTSAM, MAPLab, RTABMAP

Research Scholar Dec 16-Present

Robotics Research Center, IIIT Hyderabad, Mainly working on novel techniques on robotics vision and application to navigational planning

- Keywords: MDP, POMDP, Dec-POMDP, LSTM, CNN, Visual SLAM

Research Assistant Jan'15-Dec'16

Collaborated work, Center for AI and Robotics (CAIR), Ministry of Defence, Govt. Of India, Project Performed in Collaboration from IIIT Hyderabad
Multi-Robot SLAM estimates for multi-Robot Environment

- Keywords: Loop Closure Detection, Dloop, GTSAM, isam, Optimization

Education

2009–2013 **B Tech**, *Computer Science and Engineering, University Of Pune, Pune, India*
Major Thesis: Content Based Video Retrieval

Technical Skills

Programming: Python(expert), C,C++(expert),Haskell(medium),Cuda

Robotics Software: Robot Operating System(ROS1,ROS2),OpenCV, Point Cloud Library(PCL),GTSAM, G2O, OpenMVG, MapLab, RtabMAP

AI software Stack: PyTorch, Tensorflow, ReinforceJS

Scientific Media

Patents

Europe, Thailand, Australia: Vaidya, Tushar S., et al. "Teeth Wear Monitoring system for Bucket Wheel Excavator" EU: EP22720454, Australia: AU22441613 , Thailand: 2401005248

Conference Proceedings

ICRA'20: Vaidya, Tushar S., et al. "Topological Mapping for Manhattan-like Repetitive Environments" International Conference on Robotics and Automation 2020 (ICRA'20)

ICCSP'13: Vaidya, Tushar S., et al. "Performance gain of content based video retrieval technique using intermediate block truncation coding on different color spaces." 2013 International Conference on Communication and Signal Processing. IEEE, 2013.

ICRA'16: Vaidya, Tushar S., et al. "On-line solution for de-mining in a stochastic field" 2016 International Conference on Robotics and Automation, Humanitarian Robotics and Automation Challange, ICRA'16

Journal Articles

IJCA'13: Vaidya, Tushar S., et al. "Color content based video retrieval using block truncation coding with different color spaces." International Journal of Computer Applications 64.3 (2013).

IJCA'13 Vaidya, Tushar S., et al. "Performance augmentation of video retrieval using even-odd videos with multilevel block truncation coding." International Journal of Computer Applications 64.9 (2013).

Interviews and Talks

ICRA'16 **Humanitarian Robotics and Automation Technology Challenge(HRATC)** Interview Covered By Radio Sweden

University Of Pune **Talk on Content Based Video Retrieval**

Notable Projects

AICoE **Machine Vision Pipeline as microservice**, Mentored and coordinated the team which developed GPU supported version of **bundle adjustment and sensory calibration** pipeline for applications related to 3D mapping and SLAM, Total 1800 stores across nation are making API calls to our service to convert their physical retail stores into 3D assets.

Standard Cognition **Deeply trained point based features**, Training variants of SIFT and SURF features using self supervision methods. Competing to simillar methods in the market like **Superpoint, Superglue**. These features emerge with ability to perform matching on image pairs with diverse viewpoints.

Standard Cognition **Sparse Bundle Adjustment on Multi Camera Systems**, Bundle Adjustment system to Multi Camera Extrinsic Calibration. Methods used to solve association problem over here is deeply trained point features and matching technique.

ExplorerAI **City-Scale Mapping**, Project dealt with the construction of MAPs from multimodal data viz LIDAR, Camera, GPS, IMU using multimodal sensor fusion techniques and using semantic cues of road junctions recognition, loop closure detection and using graph optimization techniques to refine MAPS. Target maps are intended for the task of Self Driving for cars

- *Project Deployment Link*

CAIR and RRC collaboration **Multi-Robot Visual SLAM for Autonomous Ground Vehicles:**, Project Deals with Localization and mapping of multiple Autonomous Vehicles in the environment by exchanging their observations and giving better predictions of the environment to participate in navigational planning

- *Project Demo Link*

RRC-IIITH **Topological Mapping for warehouse robotics:**, Project aims to merge mapping localization and navigation pipeline to inherit human-like navigation for robots, Application to warehouse environment for autonomous navigation

- work showcased at 36^{th} International Conference on Robotics and Automation(ICRA)

ICRA'16 **Autonomous navigation in stochastic Environment and an application to demining**, Project proposes online optimal sensor-based coverage approach to scan for potential landmines using landmine sensor mounted on autonomous robot, Project is sponsored by Nvidia and IEEE under Humanitarian Robotics and Automation Society

- *Link-1*
- *Link-2*

Undergrad 2013 **Content Based Video Retrieval**, Project proposes achieving the faster, efficient and human intuitive video retrieval from a large video database. The content of videos are represented by scale and gradient invariant unique signature using Thepade's sorted Ternary Block Truncation Coding

- Advisor: Dr. Sudeep Thepade

Leadership Activities and Curricular impacts

T-Hub, Hyderabad — Mentor for AI and Robotics Startups at **T-Hub** — Asia's largest incubation centre, Hyderabad, Telangana, India

IROS'25,'26 — Recognised as **Best Reviewer** at the IEEE/RSJ International Conference on Intelligent Robots and Systems (**IROS**) for the years 2025 and 2026

NeurIPS'20 — Organizing Volunteer at 34^{th} Conference on Neural Information Processing Systems (NeurIPS'20)

ICRA'20 — Organizing Volunteer at International Conference on Robotics & Automation 2020 (**ICRA'20**)

2016 — Honorary Mention as 'Runner Up' for Submission of work at International Conference on Robotics and Automation 2016 (**ICRA'16**) for work on Autonomous Navigation and Landmine Detection robot.

2016 — Received Grants from Center for AI and Robotics (**CAIR , Ministry Of Defence Govt. of India**) and also from Nvidia for a research project in the domain of Robotics Vision and Navigational Planning for Autonomous Ground Vehicle

Undergrad, 2013 — Represented at ACM-Inter Collegiate Programming Contest (**ICPC'13**) (First representative from University at ICPC)

Undergrad 2011 — Founder of Team for University and representing the team for **ROBOCON'11** achieving a title of best-performing debutant team and also giving road-map for future students.

Languages

English — Fluent

Marathi — Fluent

Hindi — Fluent

Sanskrit — Intermediate