Siddharth Singh

singh.sid930@gmail.com | /Singh-sid930 | singh-sid930.github.io/ | /in/siddharth-singh-764b49bb/

SUMMARY

Applied Scientist and published researcher with over 5 years of experience developing end-to-end scientific solutions for embodied AI systems and autonomous robots. Expertise in development of multi-modal foundation models, synthetic data generation pipelines and robotic simulation systems from the ground up.

EDUCATION

University of Pennsylvania August 2018 - May 2020
M.S. Electrical Engineering

JSS Academy of Technical Education August 2013 - May 2017
B.S. Mechanical and Automotive Engineering

KEY SKILLS

Programming and Scripting Languages: C/C++, Python, Java, Kotlin, Bash
Libraries: Pytorch, Numpy, Scipy, Scikit-learn, OpenCV, TensorRT, HuggingFace, Transformers, LangChain
Platforms: Linux, ROS, Unity, IsaacSim, IsaacLab, Omniverse, ComfyUI, Gazebo, Docker, AWS, Weights & Biases

INDUSTRIAL EXPERIENCE

Applied Scientist-II, Product : Alexa Devices October 2024 - Present
Lab126, Amazon Sunnyvale, CA

- Led a team of scientists and engineers to develop end-to-end synthetic data generation, validation and fine-tuning pipelines for Audio-visual foundation models for targeted scene understanding use cases.
- Developed pipelines for both real-to-augmented and simulated-to-augmented videos utilizing Wan2.2, Cosmos and multi-ControlNet family models along with sensor noise and environment noise injection.
- Implemented a audio data labeling and gating system using AFCLAP, DASS and SSLAM to improve AQA generation and data imbalance in training of Audio Foundation Models.
- Developed a synthetic data validation pipeline to analyze sim-to-real gap by performing activation and embedding map analysis and zero shot inference.
- Fine tuned foundation models for scene understading use-cases using using synthetic data capturing time-of-day variations ethnicity variations, object replacement, etc.
- Guided and contributed to the development of a 3d Gaussian Splatting based digital twin creation framework from scratch starting from data collection critereas to deployment over Isaac-Sim NuREC for Novel-View-Synthesis.
- Integrated environment generation with human activity generation to create semantically rich simulation environments for embodied AI and Human-Robot-Interactions(Published)

Applied Scientist, Product : Astro and Alexa Devices November 2022 - September 2024
Lab126, Amazon Sunnyvale, CA

- Developed a novel simulation environments and digital cousins generation pipeline for end-to-end robot stack development and testing.
- Integrated LLMs over Bedrock, CLIP, SBERT and optimization methods to create a pipeline to generate environments from natural language prompts.
- Designed a probabilistic semantic environment mapping system to create digital cousins by passing RGBD data of Astro through several computational geometry and deep learning layers.
- Built scalable human motion synthesis pipeline for generating diverse activity datasets in simulation, supporting human-robot interaction research and activity recognition model training.

- Improved the digital twin generation framework to achieve >90 percent occupancy accuracy and accelerated resolution of algorithmic bugs in real world from >2 months to <3 weeks.
- Delivered ~$>300K cost savings through environment generation, digital twins and human motion synthesis pipelines, while reducing operational overhead and accelerating development velocity.

Software Development Engineer, Product : Astro August 2020 - October 2022
Lab126, Amazon Seattle, WA

- Engineered physics-based simulation scenarios with user-configurable parameters for robotic algorithm validation, including slip dynamics, wheel obstruction, and terrain interaction models.
- Refactored and improved end-to-end simulated visual perception pipeline to achieve high-fidelity sim-to-real transfer for detection and classification, ensuring consistency between simulated and on-device perception systems.
- Debugged and optimized simulated sensor suite (LiDAR, IMU, cameras) to ensure reliable SLAM execution and accurate localization in complex indoor environments.

ACADEMIC RESEARCH EXPERIENCE

Research Assistant April 2019 - May 2020
GRASP Laboratory, University of Pennsylvania Philadelphia, PA

- **Project: RoboNet: Visual foresight based learning for action prediction and planning (published)**
- Engineered data collection framework for diverse robotic platforms and environments to train video prediction models
- Implemented state space augmentations to improve few-shot learning across different robot morphologies to enable few shot learning and generalization capabilities of the model.
- **Project: Embodied AI for active semantic goal navigation for indoor environments (published)**
- Implemented semantic mapping to generate ground truth labels for ensemble of semantic prediction models and investigated different approaches for predicting spatial distribution of object semantics and scene understanding.
- Achieved state-of-the-art performance in object goal navigation on Habitat-AI benchmarks (MP3D & Gibson datasets), surpassing previous baselines.
- **Project: Curiosity Increase equality in Competitive Resource Allocation (published)**
- Designed curiosity-driven algorithms for multi-agent exploration and foraging tasks in competitive resource allocation
- Established theoretical and empirical evidence that curiosity-driven agents achieve more equitable resource distribution in competitive multi-agent scenarios.

Research Assistant November 2018 - May 2020
PRECISE Center, University of Pennsylvania Philadelphia, PA

- **Project: Development of 1/10th scale autonomous vehicles using Jetson TX2 and multimodal sensor suite (NeuRIPS Best honorable mention demo)**
- Built real-time perception pipeline integrating RGBD cameras for semantic depth map generation and video prediction, enabling proactive obstacle detection.
- Developed VAE-based models for time-series prediction with novel loss functions incorporating temporal and gradient consistency, improving prediction accuracy.
- Implemented Model Predictive Control (MPC) framework for trajectory generation and tracking using IPOPT optimization, achieving robust obstacle avoidance in dynamic scenarios.

PUBLICATIONS

- S. Dasari, F. Ebert, S. Tian, S. Nair, A. Xie, B. Bucher, S. Singh, K. Schmeckpeper, S. Levine, and C. Finn. Robonet: Learning and generalizing across robots through large-scale visual prediction and planning. In *CoRL*, 2019
- I. Idrees, S. Singh, et al. A framework for realistic simulation of daily human activity. In *ROMAN*, 2023
- I. Idrees, S. Singh, et al. A framework for realistic simulation of daily human activity. In *ROMAN*, 2023
- G. Georgakis, B. Bucher, K. Schmeckpeper, S. Singh, et al. Learning to map for active semantic goal navigation. In *ICLR*, 2022

- S. Singh, B. Bucher, et al. Curiosity increases equality in competitive resource allocation. In *ICLR*, 2020
- H. Zheng J. Auckley S. Singh M. O'Kelly, D. Karthik et al. F1/10: An open-source 1/10th scale platform for autonomous racing and reinforcement learning. In *NeuRIPS*, 2019
- S. Singh, A. Modh, et al. Gradient aware-shrinking domain based control design for reactive planning frameworks used in autonomous vehicles. In *Advances in Robotics*, 2019