Siddharth Singh

siddharth93305@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

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
- **S. Singh**, I. Idrees, et al. Realistic synthetic household data generation at scale. In *AAAI*, 2026
- 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
- M. O'Kelly, D. Karthik, H. Zheng, J. Auckley, **S. Singh**, 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

INDUSTRIAL EXPERIENCE

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

- Architected end-to-end synthetic data pipeline for edge-deployed audio-visual foundation models, targeting scene understanding failure modes; now leading a cross-functional team (2 engineers, 1 scientist) through productionization and model fine-tuning.
- Developed real-to-augmented and sim-to-augmented video pipelines using Wan2.1, ControlNet, SAM2, and DepthAnything—enabling targeted augmentations (time-of-day, object replacement, human activity injection) to address model failure modes and expand long-tail data coverage.
- Integrated realistic sensor noise modeling (shot noise, read noise, motion blur, defocus, vignetting) into synthetic data pipelines to reduce sim-to-real domain gap.
- Built audio data labeling and gating framework using CLAP-based embeddings and sound event detection models (SSLAM, DASS) to generate strong labels for AQA supervision and enable balanced training distributions across sound event categories.
- Developed synthetic data validation framework using DINOv2 and SigLIP embeddings to quantify sim-to-real gap—comparing patch-level activation distributions across augmentation types, validating behavioral consistency between synthetic and real data through zero-shot inference on Qwen-8B/32B, and benchmarking against in-development edge models.
- Fine-tuned edge-deployed VLMs using LoRA, experimenting with synthetic/real data mixture ratios to optimize scene understanding performance across targeted failure modes.

- Guided development of 3D Gaussian Splatting-based digital twin framework for remote RGBD data collection—reducing need for repeated physical site visits.
- Implemented depth regularization in gsplat and introduced preprocessing pipeline (super-resolution, motion deblurring) to improve reconstruction quality from noisy captures.
- Deployed 3DGS digital twins to Isaac Sim via NuREC, enabling integration of simulated sensors and robot models within photorealistic reconstructed environments.
- Designed iterative co-generation algorithm where environment and human activity generation inform each other, improving semantic grounding; conducted statistical analysis validating alignment between generated outputs and input personas. (Published at ROMAN 2023)

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

- Developed probabilistic digital cousin pipeline reconstructing real environments from Astro RGBD data, achieving >90% occupancy accuracy against ground truth layouts.
- Built natural language environment generation pipeline: persona and household descriptions → LLM-generated floor plan schematics → CLIP/SBERT asset retrieval → optimization-based spatial placement. Reduced environment creation from 3 months (manual) to 1 week, generating 50+ simulation environments.
- Built geometric reconstruction pipeline combining line regularization for wall detection, depth-projected 2D segmentation for asset extraction, and semantic matching with scale/rotation alignment.
- Enabled rapid algorithm iteration by providing digital cousin environments for testing—reducing bug resolution cycles from >2 months (physical site revisits) to <3 weeks.
- Delivered >$300K cost savings through automated environment generation, reducing reliance on manual tech artist work.

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

- Developed physics-based simulation scenarios for Astro algorithm validation—modeling slip dynamics, wheel obstruction, and terrain interactions to expose edge cases difficult to reproduce physically.
- Extended simulated visual perception pipeline with bounding box generation, pet detection, and instance segmentation to ensure parity between simulation and on-device perception systems.
- Fixed IMU sensor simulation by correcting faulty trapezoidal integration of acceleration and velocity, eliminating erroneous spikes that disrupted SLAM execution.

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 across diverse robotic platforms for video prediction model training.
- Implemented state space augmentations to improve few-shot learning and generalization across robot morphologies.

- **Project: Embodied AI for active semantic goal navigation for indoor environments (published)**
- Implemented semantic mapping for ground truth label generation; investigated approaches for predicting spatial distribution of object semantics.
- Achieved state-of-the-art object goal navigation on Habitat-AI benchmarks (MP3D & Gibson).
- **Project: Curiosity Increases Equality in Competitive Resource Allocation (published)**
- Designed curiosity-driven algorithms for multi-agent foraging in competitive resource allocation.
- Established theoretical and empirical evidence that curiosity improves equitable resource distribution.

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.

KEY SKILLS

Programming and Scripting Languages: C/C++, Python, Java, Kotlin, Bash, CUDA

Libraries: Pytorch, Numpy, Scipy, Scikit-learn, OpenCV, TensorRT, HuggingFace, Transformers, LangChain, gsplat, Diffusers, PEFT/LoRA

Platforms: Linux, ROS, Unity, IsaacSim, IsaacLab, Omniverse, ComfyUI, Gazebo, Docker, AWS (Bedrock, SageMaker), Weights & Biases