Sai Navaneet

sainavaneet76@gmail.com +82 10-5831-6619 Daegu, South Korea | Work Authorization: Requires visa sponsorship LinkedIn

https://sainavaneet.github.io/portfolio/

About Me

Vision-Language-Action (VLA) Model Engineer with extensive experience designing and deploying scalable state space architectures for autonomous robotic systems. Proven success developing advanced VLA models like MambaVLA, achieving faster inference latency and superior task performance on LIBERO and RoboCasa benchmarks. Skilled in AI frameworks (PyTorch, TensorFlow, JAX), transformers, diffusion models, and integrating vision-language policies with robotic control pipelines. Proficient in Python, C, C++, and ROS/ROS2 for developing and deploying robotics software. Adept at end-to-end system integration from dataset curation to real-world deployment on platforms including Franka and ViperX.

Experience

Researcher – Physical Intelligence Lab — Kyungpook National University *Feb 2024 – Dec 2025 | Daegu, South Korea*

- Designed and deployed end-to-end Robotics control pipelines on real hardware using ROS and ROS2, integrating perception, policy execution, and evaluation.
- Built state handling, long-horizon task logic, and control loops for robotic manipulation systems.
- Developed learning-based manipulation policies achieving 94% success on LIBERO and 88% on RoboCasa benchmarks using PyTorch, TensorFlow, and JAX frameworks for transformers and state space models like Mamba.
- Implemented navigation-aware task execution by combining learned policies with classical control strategies.
- Led system-level integration from dataset preparation and training to real-robot deployment and validation.

Research & Robotics Engineer – Airobotics *Apr 2025 – Present | Daegu, South Korea*

- Worked on industrial autonomous systems using Yaskawa robots for automotive manufacturing.
- Integrated vision-based perception with robotic control for weld bead detection and inspection.
- Contributed to robot motion planning and system integration in production environments.

Robotics Engineer – Dexweaver *Jul 2024 – Dec 2024 | Daegu, South Korea*

- Developed a vision-guided robotic manipulation system using ViperX arms for tissue processing.
- Built teleoperation, data collection, and autonomous execution pipelines using ROS.
- Integrated Sensor Fusion of LiDAR and point-cloud data for environment perception and workspace understanding.
- Achieved 85.7% autonomous success rate, comparable to human teleoperation (92.4%).
- Engineered multi-modal datasets (RGB, point clouds, joint states) and deployed trained policies on real robotic hardware.

Research Intern – Physical Intelligence Lab *Sep 2022 – Feb 2024 | Daegu, South Korea*

- Implemented navigation, obstacle avoidance, and state estimation for differential-drive mobile robots.
- Worked with LiDAR-based SLAM, localization, and mapping pipelines in ROS.
- Applied reinforcement learning and classical control for autonomous mobile robot navigation.
- Integrated perception, planning, and control into closed-loop autonomous systems.

Education

Kyungpook National University – Daegu, South Korea *Expected Dec 2025 | GPA: 4.07/4.3*

M.S. in Electronics and Electrical Engineering

Kyungpook National University – Daegu, South Korea *Feb 2024 | GPA: 3.8/4.3*

B.S. in Electronics Engineering (Double Degree)

Christ University – Bangalore, India *Dec 2021 | GPA: 3.7/4*

B.Tech. in Electronics and Communication Engineering

Publications

MambaVLA: A Scalable and Efficient Vision-Language-Action Model with State Space Architecture – CCNC 2026

Jan 2026

LegMamba: A Scalable and Efficient State Space model for Quadrupedal Locomotion – POSTER IROS 2025
Nov 2025

QROOT: An Integrated Diffusion Transformer and Reinforcement Learning Approach for Quadrupedal Locomotion – NeurIPS 2025 (under review)
Dec 2025

Vision-Guided Predictive Action Imitation Learning with Discrete Latent Encoding for Multitasking Robots – Submitted to Neural Networks
Jun 2025

Discrete Latent Diffusion Motion Planning – The International Conference on Nonlinear Dynamics (NODYCON 2025)
Jun 2025

Vision-Guided Imitation Learning Using Action Chunk Transformers – IEMEK Symposium on Embedded Technology 2024 (IEMEK 2024)
Oct 2024

Hybrid Model Predictive and Iterative Learning Control for Enhanced Leader-Follower Robotic Tracking – KNU-EERC 2024
May 2025

Robotics Software Skills

ROS/ROS2, LiDAR-based SLAM, Navigation, Robot control, Sensor fusion, Gazebo, RViz, MuJoCo, Isaac Sim, C++, Python, C, MATLAB, OpenCV, PCL, Docker, Git, Imitation Learning, Reinforcement Learning, Diffusion Models, Transformers, State-Space Models (Mamba), Vision-Language-Action (VLA), Embedded Linux

Programming Languages: C++, Python, C, MATLAB

Robotics & Systems: ROS/ROS2, Gazebo, RViz, MuJoCo, Isaac Sim

Tools & Libraries: OpenCV, PCL, Docker, Git

Core Expertise: Imitation Learning, Reinforcement Learning, Diffusion Models, Transformers, State-Space Models (Mamba)

Applied Areas: Vision-Language-Action (VLA), Multimodal perception, Sensor fusion for robotics, Learning-based control

Specialization: Robot learning, Policy optimization, Navigation-aware manipulation, Path planning, Embedded Linux

Robots Worked With: Franka, ViperX, Go2, AGV, TurtleBots, Drones

Projects

MambaVLA: Vision–Language–Action Model with Mamba State Space Architecture

- Built an efficient VLA pipeline combining Mamba SSM, Eagle2 visual backbone, Qwen-7B encoder, and diffusion-based action generation. Achieved 5-10ms faster inference latency compared to baseline transformer models, enabling faster and more efficient robot control across multiple robotics benchmarks. Project website: https://sainavaneet.github.io/MambaVLA.gihub.io/

Transformer Based Vision Guided Tissue Processing

- Developed an automation of tissue packing using ViperX robotic arms.

Action Chunk Transformer on Franka Robot

- Implemented Act on Franka robot to do vision guided imitation learning on pick and place tasks.

QROOT: An Integrated Diffusion Transformer and Reinforcement Learning Approach for Quadrupedal Locomotion

- Introduced a control stack that combines diffusion transformer with a reinforcement learning-based stabilizer (PPO), enabling smooth and robust execution on real-world hardware.

Vision-Guided Predictive Action Imitation Learning with Discrete Latent Encoding for Multitasking Robots

- Introduced a control stack that combines diffusion transformer with a reinforcement learning-based stabilizer (PPO), enabling smooth and robust execution on real-world hardware.