

Arjun Srinivas

+65 86780668 | arjun022@e.ntu.edu.sg | arjunsrinivas.in | [github](https://github.com) | [linkedin](https://www.linkedin.com/in/arjunsrinivas)
Robotics & Machine Learning | Availability: Part-time until May 2027 | Full-time from July 2027

EDUCATION

Nanyang Technological University, Singapore

Aug 2023 – Jul 2027

Bachelor of Engineering in Electrical and Electronic Engineering

Relevant Modules: Robotics and Automation, Image Processing and Computer Vision, Artificial Intelligence and Data Mining, Digital Signal Processing, Microprocessors, Modelling and Control Theory

WORK EXPERIENCE

Institute for Infocomm Research (I²R), A*STAR

May 2025 – Dec 2025

Audio / Multimodal LLM Research Intern

- Developed and tuned multilingual speaker-diarisation systems on NVIDIA NeMo (custom configs, Optuna hyperparameter search, WhisperX/OpenSMILE feature extraction) across 10+ experiments on NSCC and A*STAR GPU clusters.
- Built core components of a multimodal speech-emotion pipeline for an ICASSP paper submission; benchmarked audio-LLMs (Qwen3-Omni, Audio Flamingo 3) on MELD and MSP-Podcast, with Qwen3-Omni reaching ~0.57 UAR / ~0.54 macro-F1 on MELD and ~0.35 macro-F1 on MSP-Podcast, and identified prosodic features (pitch, loudness) as most discriminative.
- Diagnosed non-deterministic Audio Flamingo 3 outputs across 5+ datasets and stabilised them with majority-vote, mean-softmax and logit-sum fusion, evaluated via Cohen's kappa, Accuracy, F1 and UAR.
- Rebuilt the audio-LLM inference backend for explicit audio conditioning: upgraded the AF3 base from LLaMA-2B to LLaMA-7B, added a Whisper-3 Large "sound tower" front-end, and extended experiments to Qwen3-Omni-30B.

PROJECTS

Final Year Project, Learning-Based 3D Autonomous Exploration

Aug 2026 – Present

- Benchmarked classical LiDAR exploration planners (EPIC, FUEL) on custom point-cloud maps with coverage and containment tooling, then built an attention-based deep-RL planner (extending ARIADNE) that selects destination viewpoints rather than motion primitives.
- Ported the policy from 2D to 3D multi-storey environments, warm-starting the 3D policy from the 2D one; consolidating results towards a paper submission targeting IROS 2027.

Robotics and Computer Vision Engineer, Medication Assistance Automation System

Jan 2026 – Apr 2026

- Built an end-to-end medication-assistance prototype spanning prescription scheduling, vision-based basket identification, 3D pose estimation and robot-arm simulation using OpenCV, EasyOCR, YOLOv8, ROS 2, Gazebo and ZED stereo vision.
- Developed the CV/OCR pipeline for real-time basket detection and label verification (dataset creation, augmentation, ROI preprocessing, multi-frame confirmation), reaching 95% detection accuracy.
- Engineered ZED-based 3D pose estimation (5x5 depth sampling, median filtering, temporal smoothing, grab-point calibration) and integrated ROS 2 / Gazebo digital-twin simulation for pick-and-place execution.

Full Stack Developer, RAG Streamlit App

May 2025 – Jun 2025

- Built a Streamlit RAG app over a local Ollama endpoint: chunked and vectorised documents with Chroma (LRU-cached embeddings, 256 entries) and served top-5 cosine-similarity passages with source attribution.

CO-CURRICULAR ACTIVITIES

NTU-EEE Machine Learning & Data Analytics Lab (MLDA)

Aug 2024 – May 2026

Vice President

- Promoted from member to Director then Vice President within a year; represented the lab in sponsorship and strategy discussions across 3 flagship events.

Deep Learning Week 2025

Oct 2024 – Feb 2025

Vice President, Communications & Publicity

- Led publicity for NTU's largest hackathon (750+ participants) and secured sponsorships from industry partners.

NTU Dragonboat Team, Varsity

Aug 2023 – Feb 2025

- Won Gold at the Singapore National Men's Open and Tertiary races (2024); represented NTU across 5+ events nationwide.

SKILLS

Languages: Python, C, C++, Java, MATLAB, SQL

ML & Audio: PyTorch, TensorFlow, Scikit-Learn, Optuna, NVIDIA NeMo, WhisperX, OpenSMILE

Vision & Robotics: OpenCV, YOLOv8, EasyOCR, ROS 2, MoveIt 2, Gazebo, ZED SDK

Embedded & Tools: STM32, ESP32, Arduino, Raspberry Pi, KiCad, LTSpice, Vivado, Git, Linux

Spoken: English, Hindi, Tamil, Bengali, Mandarin, German (A2)