Ankit Kumar Shaw

M.Sc. Graduate | Robotics & AI Enthusiast

Kolkata, India +917904534830 shawak10@mails.tsinghua.edu.cn https://ankit-zefan.github.io

Summary

Graduate student researching Embodied AI with a focus on Autonomous Driving, Haptic shared Robot Learning, and Multimodal LLMs. Passionate about building context-aware robotic agents that collaborate with humans through language and touch to solve complex manipulation tasks.

Education

School of Vehicle and Mobility, Tsinghua University — **Sep 2020 - Dec 2024**
Mechanical Engineering — Master of Science
3.87/4.0, Rank: 5th

Thesis: **"Multimodal LLM based Data Cleansing Model for Confidence Score driven Crowdsourced Data Fusion"** under Dr. Yang Diange in Autonomous Driving Lab.
[NB: Almost 2 years of study gap due Covid-19]

Vellore Institute of Technology — **Jul 2016 - Jun 2020**
Mechanical Engineering — Bachelor of Technology
8.78/10.0

Thesis: "**Investigation into the effects of Waste Plastic Oil in Biogas fuelled Dual-Fuel Engine**" under Dr. M. Feroskhan.

Experience

National University of Singapore — **Mar 2025 - Present**
Research Assistant

Visuo-Tactile based Robot Learning and Manipulation under Prof. Mike Shou at Show Lab in the School of Computing.

King Abdullah University of Science and Technology — **Jun 2022 - Dec 2022**
Remote Intern as well as Visiting Student — Saudi Arabia

Learned to implement simple Offline Reinforcement Learning models in different Mujuco Environment under Prof. Mohamed Elhoseiny at Vision-CAIR group in Visual Computing Center.

Tsinghua AI for Student Club (TAIS) — **2023 - 2024**
President — Tsinghua University

Actively participate in AI hackathons, debates, and global collaborations with university and industry teams in China, fostering a vibrant AI community through talks and panel discussions.

Future Robotics Club — **2023 - 2024**
Team Member — Tsinghua University

Design and develop the "Tinker" home-based indoor robot for RoboCup @home.

Conference Reviewer — **IROS 2025, Hangzhou, China.**

Research Interests

Embodied AI, Visuo-Tactile based Robot Learning and Dexterous Manipulation, Human-Robot Interaction (HRI), Multimodal Sensing and Robot Perception, Multimodal Large Language Models, Vision Language Action Models, Scene Understanding in Autonomous Driving

Research Projects and Publications

CleanMAP: Distilling Multimodal LLMs for Confidence-Driven Crowdsourced HD Map Updates **Feb 2024 - Dec 2024**

A. K. Shaw, K. Jiang, T. Wen, C. K. Sah, Y. Shi, M. Yang, D. Yang, and X. Lian, "CleanMAP: Distilling Multimodal LLMs for Confidence-Driven Crowdsourced HD Map Updates," *arXiv preprint arXiv:2504.10738*, 2025. [Online]. Available: https://arxiv.org/abs/2504.10738.

Accepted at the **CVPR 2025** Workshop on Distillation of Foundation Models for Autonomous Driving (**WDFM-AD**) and will be published as **CVPR 2025 Workshop Proceedings (IEEE).**

Advancing Autonomous Vehicle Intelligence: Deep Learning and MLLM for Traffic Sign Recognition and Robust Lane Detection **Feb 2024 - Jan 2025**

C. K. Sah, **A. K. Shaw**, X. Lian, A. S. Baig, T. Wen, K. Jiang, M. Yang, and D. Yang, "Advancing Autonomous Vehicle Intelligence: Deep Learning and Multimodal LLM for Traffic Sign Recognition and Robust Lane Detection," *arXiv preprint arXiv:2503.06313*, 2025. [Online]. Available: https://arxiv.org/abs/2503.06313

Joint first author – Contributed to the development of a multimodal large language model (MLLM) for robust understanding of road elements, including lane detection under adverse conditions.

Under review at **ICCV 2025**.

ViTaMIn: Learning Contact-Rich Tasks Through Robot Free Visuotactile Manipulation Interface **Aug 2024 - Jan 2025**

F. Liu, C. Li, Y. Qin, **A. Shaw**, J. Xu, P. Abbeel, and R. Chen, "ViTaMIn: Learning Contact-Rich Tasks Through Robot-Free Visuo-Tactile Manipulation Interface," *arXiv preprint arXiv:2504.06156*, 2025. [Online]. Available: https://arxiv.org/abs/2504.06156.

Joint second author – Contributed equally to the design and implementation of the multimodal fusion-driven diffusion model.

Accepted at **ICRA 2025 CRM Workshop**, Submitted to **CoRL 2025**.

Done in collaboration with UC Berkeley.

Stereo Camera based Object detection at different depths and performing Pick and Place Task at specific Bin using the 7 DoF Robotic Arm **Sep 2023 - Jan 2024**

Machine Vision course final project

A Privacy-Preserving Data Storage and Service Framework Based on Deep Learning and Blockchain for Construction Workers' Wearable IoT Sensors. **2022**

X. Zhou, **A. K. Shaw**, and P.-C. Liao, "A privacy-preserving data storage and service framework based on deep learning and blockchain for construction workers' wearable IoT sensors," *arXiv preprint arXiv:2211.10713*, 2022. [Online]. Available: https://arxiv.org/abs/2211.10713.

Published in IEEE Access.

Skills

Programming & Frameworks: Python, C++, PyTorch, CUDA, TensorFlow, OpenCV

Robotics Tools: ROS, Gazebo, MoveIt, RViz

Collaboration & Development: Git, Docker, Linux, Jupyter, VS Code, Github

AI & ML Tools: Deep Learning, Imitation Learning, LLM, VLM, VLA

Additional Tools: Open3D, Scikit-learn, Matplotlib, NumPy, Pandas

Achievements

Best Solution Award — **2022**
AWS Disaster Response Hackathon

Second Class Meritorious Scholarship (英才二等奖学 金) — **Oct 2021**
Tsinghua University

Awarded for having excellent overall academic records

Chinese Government Scholarship — **Sep 2020 - Jun 2024**
Tsinghua University

Fully Funded Master's Study Scholarship

Languages

English

●●●●●

Hindi

●●●●○

Chinese

●●○○○