Daichi Yashima

CONTACT INFORMATION

ydaichi1207@keio.jp
github.com/daichi1207
daichi1207.github.io

RESEARCH INTERESTS

Foundation Models: Multimodal Large Language Models, Vision-Language-Action Models
Vision & Language: Video Understanding, Self-Supervised Learning
Embodied AI: Mobile Manipulation, Service Robots

EDUCATION

MS in Computer Science, **Keio University**, Kanagawa, Japan — Apr. 2025 – Current
Advisor: Prof. Komei Sugiura

BS in Computer Science, **Keio University**, Kanagawa, Japan — Apr. 2021 – Mar. 2025

RESEARCH EXPERIENCE

SMILab, Keio University, Kanagawa, Japan — Jan. 2025 – Current
Research Assistant

HSRTX, Tokyo University, Tokyo, Japan (remote) — Jan. 2025 – Current
Research Assistant

LLMjp, NII, Tokyo, Japan (remote) — Sep. 2024 – Current
Research Assistant

WORK EXPERIENCE

ZenTech, Tokyo, Japan — Aug. 2024 – Current
Lead AI Engineer

Wanderlust, Tokyo, Japan — Dec. 2022 – Current
Co-Founder, Software Engineer

Cryptoeconomics Lab, Tokyo, Japan (remote) — Apr. 2022 – Oct. 2023
Smart Contract Engineer

SHINTA VR, Jakarta, Indonesia (remote) — Mar. 2022 – Jun. 2022
Smart Contract Engineer

PUBLICATIONS

D. Yashima, R. Korekata, and K. Sugiura, "Open-Vocabulary Mobile Manipulation Based on Double Relaxed Contrastive Learning With Dense Labeling", **IEEE RA-L**, vol. 10, no. 2, pp. 1728-1735, 2025. **(IF: 5.2)**

K. Katsumata, M. Kambara, **D. Yashima**, R. Korekata, and K. Sugiura, "Mobile Manipulation Instruction Generation from Multiple Images with Automatic Metric Enhancement", **IEEE RA-L**, vol. 10, no. 3, pp. 3022-3029, 2025. **(IF: 5.2)**

R. Takanami, P. Khrapchenkov, S. Morikuni, J. Arima, Y. Takaba, S. Maeda, T. Okubo, G. Sano, S. Sekioka, A. Kadoya, M. Kambara, N. Nishiura, H. Suzuki, T. Yoshimoto, K. Sakamoto, S. Ono, H. Yang, **D. Yashima**, A. Horo, T. Motoda, K. Chiyoma, H. Ito, K. Fukuda, A. Goto, K. Morinaga, Y. Ikeda, R. Kawada, M. Yoshikawa, N. Kosuge, Y. Noguchi, K. Ota, T. Matsushima, Y. Iwasawa, Y. Matsuo, and T. Ogata (2025). "AIRoA MoMa Dataset: A Large-Scale Hierarchical Dataset for Mobile Manipulation", arXiv preprint arXiv:2509.25032.

Skills	**Programming Languages**: Python, C/C++, Triton, Typescript, Rust, Bash, Lua, Vimscript **Tools**: Deep Learning Frameworks (Pytorch, TensorFlow, JAX), Python libraries (NumPy, OpenCV, SciPy, matplotlib, etc), ROS, Gazebo
Presentations	Invited Talk: "Multimodal Large Language Models based on Deep State Space Models for Video Understanding," CV Study Group #135, AIST (Tsukuba, Japan), August 22, 2025