

SHENGXIANG SUN

☎ +1 4169026176 ◇ ✉ owen.sun@mail.utoronto.ca ◇ 🌐 Personal Website ◇ 🎓 Google Scholar ◇ 🐙 GitHub

EDUCATION

University of Toronto

Sep 2022 – (expected) Apr 2026

Honours Bachelors of Science in Computer Science

- GPA: 3.82/4.00

RESEARCH EXPERIENCE

- Visiting Research Assistant, University of Utah
Topics: Multi-skill In-context Imitation Learning from Human Videos
May 2025 – Present
Advisor: Prof. Weiyu Liu
- Visiting Research Assistant, National University of Singapore
Topic: Learning Robotic Assembly from Abstract Manuals
Oct 2024 – Sept 2025
Advisor: Prof. Lin Shao
- Undergraduate Research Assistant, University of Toronto
Topic: VLA Failure Detection, Point-Cloud Forecasting
May 2024 – Aug 2024
Advisor: Prof. Florian Shkurti

PUBLICATIONS

* indicates equal contribution

- 1 [In Submission] C. Tie*, S. Sun*, Y. Lin, Y. Wang, Z. Li, Z. Zhong, J. Zhu, Y. Pang, H. Chen, J. Chen, R. Wu, L. Shao, “Manual2Skill++: Connector-Aware General Robotic Assembly from Instruction Manuals via Vision–Language Models”
- 2 [NeurIPS 2025] Q. Gu, Y. Ju, S. Sun, I. Gilitschenski, H. Nishimura, M. Itkina, F. Shkurti, “SAFE: Multitask Failure Estimation for Vision-Language-Action Models” [Paper] [Website]
- 3 [RSS 2025] C. Tie*, S. Sun*, J. Zhu, Y. Liu, J. Guo, Y. Hu, H. Chen, J. Chen, R. Wu, L. Shao, “Manual2Skill: Learning to Read Manuals and Acquire Robotic Skills for Furniture Assembly Using Vision-Language Models” [Paper] [Website]

RESEARCH PROJECTS

Multi-skill In-context Imitation Learning (ICIL) from Human Videos (Ongoing)

Advisor: Prof. Weiyu Liu, Assistant Professor, Utah, CS

Aug 2025 – Present

- Training and evaluating a video-conditioned Diffusion Policy for assembly tasks within IsaacGym
- Created a dataset of long-horizon furniture assembly using FurnitureBench, IsaacGym, and LeRobot

Manual2Skill++: Connector-Aware General Robotic Assembly from Instruction Manuals via Vision–Language Models

Advisor: Prof. Lin Shao, Assistant Professor, NUS, CS

Apr 2025 – Sept 2025

- Developed a novel dataset representing connector placements (e.g., screws, nails) for assembly objects, converting abstract manual illustrations into a unified hierarchical graph representation for connector-aware assembly
- Proposed a benchmark for automatic extraction of connector placements and graph generation from manuals

Improving Point-Cloud Forecasting Accuracy, CS Project Course

Advisor: Prof. Florian Shkurti, Assistant Professor, UofT, CS

May 2025 – Aug 2025

- Developed an end-to-end model that integrates BEV features from multi-view RGB images with temporal point cloud data to improve future forecasting on the NuScenes dataset

SAFE: Multitask Failure Estimation for Vision-Language-Action Models

Advisor: Prof. Florian Shkurti, Assistant Professor, UofT, CS

May 2024 – May 2025

- Performed ablation studies on input representations (raw images vs. VLA’s final layer embeddings) for training and evaluating the failure estimation module
- Developed a PyTorch & SimplenEnv pipeline to fine-tune VLAs on mixed datasets from OXE with SLURM

Manual2Skill: Learning to Read Manuals and Acquire Robotic Skills for Furniture Assembly Using Vision-Language Model

Advisor: Prof. Lin Shao, Assistant Professor, NUS, CS

Nov 2024 – Feb 2025

- Employed VLMs to generate high-level furniture assembly plans from IKEA manuals, achieving generalization across diverse furniture types and exceeding previous baselines by over 300%.
- Generated 10,000+ furniture parts via a novel, automated Blender pipeline to simulate realistic assembly scenes

RESEARCH INTERESTS

My research spans **Robotics** and **3D Vision** to enable **generalizable robot manipulation**. I aim to develop systems that execute complex, long-horizon tasks from simple instructions (e.g., “prepare a dish from this cookbook”) by learning from existing human knowledge with foundation models, rather than collecting data from scratch.

SCHOLARSHIPS & AWARDS

- **2022-2025 General In-Course Scholarship** (For maintaining a GPA of at least 3.7/4.0) (CAD 12,000)
- **2023-2025 Dean List Scholar**
- **2024 Summer NSERC Math & Computer Science Research Award** (CAD 8,000)
- **2024 First Place, GenAI Genesis (Canada’s Largest AI Hackathon) – Best Safety AI**

WORK EXPERIENCE

Loblaw Digital

Toronto (CA)

Machine Learning Engineer Co-op - Generative AI Team

Jan 2024 – Apr 2024

- Enhanced an automated email reply system using Google’s Gemini Pro, Python, Docker, CI/CD, Few-Shot and Chain-of-Thought prompt engineering, which resulted in over 3400 correctly automated email replies per week.
- Developed an end-to-end machine learning pipeline for enhanced shopping experience, with OpenAI’s GPT-4 Vision, Python, Pandas, SQL, Apache Airflow DAGs, and Google Cloud Platform, which automatically generated product descriptions for 154,286 products sold at Loblaws, Shoppers Drug Mart, and Joe Fresh

New H3C Technologies

Beijing (CN)

Machine Learning Research Intern

Jul 2023 – Aug 2023

- Designed training & testing pipelines of Llama2, Dreambooth, InstructPix2Pix on MobaXterm and WebUI, by using PyTorch and HuggingFace, which doubled the team’s testing data outputs
- Enabled automated downloads of Python dependencies with bash scripting, reducing installation steps by 40%

PROGRAMMING SKILLS & LANGUAGE SKILLS

Proficient Python, LaTeX, HTML

Familiar PyTorch, Linux, C, Java, Git

Chinese (Native), English (Fluent), French (Intermediate)