

SHENGXIANG SUN

⌂ <https://owensun2004.github.io/>

📞 +1 4169026176

✉️ owen.sun@mail.utoronto.ca

🎓 Google Scholar

🐙 GitHub

EDUCATION

University of Toronto

Sep 2022 – (expected) Apr 2026

Honours Bachelor of Science in Computer Science

- GPA: 3.82/4.00

RESEARCH EXPERIENCE

- Visiting Research Assistant, Stanford University

Topics: Multi-skill In-Context Imitation Learning from Human Videos

May 2025 – Present

Advisor: Dr. 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 2025

Advisor: Prof. Florian Shkurti

PUBLICATIONS

* indicates equal contribution

1 [ICRA 2026] Chenrui Tie*, Shengxiang Sun*, Yudi Lin, Yanbo Wang, Zhongrui Li, Zhouhan Zhong, Jinxuan Zhu, Yiman Pang, Haonan Chen, Junting Chen, Ruihai Wu, Lin Shao, “Manual2Skill++: Connector-Aware General Robotic Assembly from Instruction Manuals via Vision–Language Models” [\[Paper\]](#)

2 [NeurIPS 2025] Qiao Gu, Yuanliang Ju, Shengxiang Sun, Igor Gilitschenski, Haruki Nishimura, Masha Itkina, Florian Shkurti, “SAFE: Multitask Failure Estimation for Vision-Language-Action Models” [\[Paper\]](#) [\[Website\]](#)

3 [RSS 2025] Chenrui Tie*, Shengxiang Sun*, Jinxuan Zhu, Yiwei Liu, Jingxiang Guo, Yue Hu, Haonan Chen, Junting Chen, Ruihai Wu, Lin Shao, “Manual2Skill: Learning to Read Manuals and Acquire Robotic Skills for Furniture Assembly Using Vision-Language Models” [\[Paper\]](#) [\[Website\]](#)

ACADEMIC SERVICE

Reviewer for ICRA 2026, CVPR 2025 Robo-3DVLM

RESEARCH PROJECTS

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

Advisor: Dr. Weiyu Liu, Postdoc, Stanford, CS; incoming Prof., University of Utah

May 2025 – Present

- Training and evaluating a video-conditioned Diffusion Policy for assembly tasks within Isaac Gym
- Created a dataset of long-horizon furniture assembly using FurnitureBench, Isaac Gym, 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 21 assembly objects using Blender, converting manual illustrations into an Assembly 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 multi-view RGB images with temporal LIDAR scans to improve future point cloud 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 (DINO & CLIP features vs. VLA's final layer embeddings) for training and evaluating the failure estimation module
- Developed a PyTorch & SimplerEnv 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

Oct 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 9,000)
- **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

Machine Learning Engineer Co-op - Generative AI Team

Toronto (CA)

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

Machine Learning Research Co-op

Beijing (CN)

Jul 2023 – Aug 2023

- Designed testing pipelines of Llama2, Dreambooth with PyTorch, 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

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