Yuchen You

Unit 5030, 1929 Plymouth Rd., Ann Arbor, MI 48105

Email : yuchenxr@umich.edu
Alt : wesley_you@sjtu.edu.cn
Mobile : (+1) 734 510 0456

Education

- **University of Michigan** 2024 – Present
  B.S.E in Computer Science - GPA 3.96/4.00 *Ann Arbor, MI, US*
  - **Honors**: Cheng-Family Scholarship (Top 2%)
- **Shanghai Jiao Tong University** 2022 – 2024
  B.S.E in Mechanical Engineering - GPA 3.81/4.00 *Shanghai, China*
  - **Honors**: John Wu & Jane Sun Sunshine Scholarship (Top 6%), SJTU Excellence Scholarship (Top 3%)

Research Experience

- **Agentic Distributed System Ops** May 2024 – Present
  Order Lab, University of Michigan; Advisor: Prof. Ryan (Peng) Huang *Ann Arbor, MI, US*
  - **Closed-loop control plane**: Engineered a **closed-loop control plane** for **ZooKeeper** clusters to mitigate overload and network fluctuation, integrating **Prometheus** telemetry, **HAProxy** for traffic shaping, and **Resilience4j** for circuit breaking.
  - **Mitigation action library**: Implemented a library of **atomic mitigation actions**—including network throttling (**tcconfig**), dynamic load balancing (**HAProxy Runtime API**), **JVM** heap tuning, and disk I/O control (**fsync**)—packaged as auditable scripts that any controller can invoke safely.
  - **Load injection & evaluator**: Built a custom load generator (zkbench) that supports **multi-stage traffic** (warmup → spike → cooldown) and **skewed distributions** to stress-test resilience, and developed a closed-loop evaluator that correlates injection rates, SLO violations, and throughput for different control strategies.
  - **Agentic controller variant**: Prototyped an **LLM-based controller** (GPT-4o with **Model Context Protocol** tools) that consumes Prometheus snapshots and log summaries to select mitigation actions, enabling comparison between static policies and agentic reasoning on complex failure patterns.
- **SoftRobot Electronic Control** Sept. 2024 – Present
  HDR Lab, University of Michigan; Advisor: Prof. Xiaonan (Sean) Huang *Ann Arbor, MI, US*
  - **Real-time embedded control stack**: Architected a **real-time control stack** across **STM32** (low-level actuation) and **Orange Pi** (planning); designed custom PCBs and implemented **multi-threaded task isolation** to decouple high-frequency **CAN**/$\mathbf{I^2C}$ telemetry from planning and logging.
  - **Control algorithms and kinematics**: Implemented feedback control and a **Piecewise Constant Curvature** kinematic solver with Jacobian-based IK to handle high-DoF redundancy in real time, supporting smooth and repeatable soft-arm motion.
  - **Validation and Recognition**: Validated robustness through collaboration with **General Motors**; work recognized with an **ICRA 2025 Best Poster Award** and spotlight talks at **RoboSoft 2025** and **ICON 2025**.
- **Control Developer** Feb. 2024 – Sept. 2024
  Sirius Lab, Shanghai Jiao Tong University; Advisor: Prof. Yutong Ban *Shanghai, China*
  - **Embodied manipulation pipeline**: Built an embodied manipulation pipeline that integrates an **LLM-based task planner**, **ZED depth camera** perception, and a **Flexiv 7-DoF** arm to execute language-specified tabletop tasks such as jigsaw assembly.
  - **Motion control and sim-to-real**: Led motion control on top of **Flexiv-RDK**, implementing inverse kinematics and collision-aware trajectory generation, and validated trajectories in simulation before deploying on the physical robot to improve robustness.

Publications

- **Origami Inspired Soft Robotic Arm: A Modular Platform for Manipulation** May 2025
  IEEE ICRA Workshop, Atlanta, USA *Best Poster Award*
  - **Authors**: Jiyang Wang, **Yuchen You**, Xinqi Zhang, Haobo Fang, Jiaqi Wang, Xiaonan Huang
  - **Award**: Gave a **spotlight talk**; poster received the **Best Poster Award** at the IEEE ICRA 2025 Workshop.

Patent

- **U.S. Patent, 2025**: J. Wang, Y. You, X. Xhang, H. Fang, J. Wang, and X. Huang, "Lightweight, Proprioceptive, Origami-Inspired Soft Robotic Arm for High Payload, Low-Cost Reconfigurable Manipulation," U.S Patent, (Pending), 2025.

Selected Projects

- **CUDA Proxy Player** Aug. 2025 – Dec. 2025
  CSE 582 (Advanced Operating Systems) – CUDA runtime optimization for LLM inference *University of Michigan*
  - **Multi-path CUDA runtime**: Designed a multi-path CUDA execution proxy (baseline / graph / persistent) that routes requests by batch size and micro-op shape to reduce CPU launch overhead on launch-bounded LLM inference workloads.
  - **Graph & persistent execution paths**: Implemented CUDA Graph capture and caching plus persistent worker kernels with device queues and shape bucketing, achieving up to **1.5–3**× lower end-to-end latency on repetitive, launch-bounded workloads.
  - **Microbenchmarks and instrumentation**: Designed microbenchmarks and runtime instrumentation (path statistics, latency, graph hit rate) to evaluate performance across MoE-like forward passes and routing strategies.
- **Edge–Cloud Collaborative VLM System for Autonomous Driving** Aug. 2025 – Dec. 2025
  CSE 589 (Advanced Networks) *University of Michigan*
  - **Edge–cloud perception pipeline**: Designed an edge–cloud pipeline that splits perception between an onboard encoder and a cloud VLM, enabling trajectory planning under bandwidth and latency constraints.
  - **Communication stack & cloud server**: Implemented a low-latency TCP-based communication stack and cloud inference server with request batching and caching to support multi-model offloading.
  - **Trajectory evaluation**: Designed evaluation procedures for trajectory quality (e.g., collision risk, smoothness, goal reaching) and compared split points / compression levels to quantify accuracy–latency tradeoffs.
- **Simulated Distributed System** Aug. 2025 – Dec. 2025
  EECS 491 (Introduction to Distributed System)
  - **Primary–backup KV store**: Implemented a primary–backup key–value store in Go with lexical confinement and request concurrency, designing the protocol to provide linearizable GET/PUT/APPEND operations.
  - **Paxos-based replicated state machine**: Designed a three-layer Paxos replicated state machine and safely optimized the protocol with accept-phase skipping to reduce normal-case latency.
  - **Sharded KV & reconfiguration**: Implemented sharded key placement and dynamic reconfiguration using Paxos-based configuration management, debugging consistency and recovery under concurrent shard moves.
- **Simulated Basic Operating System** Jan. 2025 – Apr. 2025
  EECS 482 (Introduction to Operating System) Lecture Project
  - **User-level threading library**: Built a lightweight user-level multicore threading library (context switching, Mesa monitors, synchronization primitives) and designed scheduling to handle concurrent workloads without deadlock.
  - **Pager & virtual memory**: Implemented a minimal pager and MMU layer managing page tables and dirty/reference bits, designing a clock-based eviction policy and copy-on-write support for fork/mmap.
  - **Network file system & kernel tracing**: Implemented a Unix-style network file system with synchronized inode operations and extended Linux `ptrace` with selective memory snapshot/restore to support efficient debugging.
- **Network Simulation** Jan. 2025 – Apr. 2025
  EECS 489: Introduction to Computer Networks
  - **Mininet-based experiments**: Designed Mininet topologies and C++ socket experiments to measure RTT/throughput and reproduce bufferbloat, observing how queue sizes and AQM impact latency.
  - **Video proxy & transport**: Implemented a video proxy with basic load balancing and adaptive streaming, and built a TCP-like reliable transport over UDP plus an L3 router with ARP/ICMP handling.
  - **SDN controller**: Implemented a POX-based SDN controller that assigns flows to QoS queues, designing rules to mitigate bufferbloat for latency-sensitive traffic.
- **Digital Forensics** Jan. 2025 – Apr. 2025
  EECS 388 (Introduction to Computer Security) Lecture Project
  - **Cryptanalysis & Cracking**: length-extension, padding-oracle; John the Ripper (PDF/ODT), Hydra (SSH).
  - **Web Exploitation**: auth bypass via XSS/SQLi/CSRF.

- **Binary Exploitation**: ROP/NOP-sled against DEP/ASLR.
- **Reverse Engineering**: Ghidra decompilation and PWNing.
- **Protocols**: TLS 1.3 handshake; Google-style TOTP.

- **Auto Sentry Robot Control** 2023 – 2024
  Chinese Univ. National Robot Competition – Robomaster Championship
  - **Autonomous control**: Autonomous decision making and engagement with dual gimbals and 4-wheel chassis on STM32-F407.
  - **Circuit & control**: Lead circuit design; dual-gimbal control stabilization; high-speed 4-wheel chassis response.
  - **Perception & pose**: Developed CAN/UART pipelines for CV and LiDAR data; implemented IMU-based absolute-pose control.

Skills

- **Programming**: C/C++, Java, Rust, Golang, Python, Bash; Git; CMake, Makefile, Maven, uv, cargo
- **Systems**: Arch/Ubuntu Linux; concurrency (boost locks); MMU/paging; POSIX sockets (select/poll)
- **Networking**: tc(config); HAProxy; Mininet, POX; TCP (GBN/SR), L3 routing
- **Distributed**: Docker (Compose), Kubernetes; ZooKeeper, HDFS; Prometheus+Grafana (JMX Exporter)
- **Security**: Wireshark, Ghidra, John the Ripper, Hydra, sqlmap, Autopsy, Stegseek; ROP chains
- **ML**: PyTorch, CoT, MCP, Qwen-VLM Tuning
- **Databases**: SQLite, Oracle(SQL*Plus), MongoDB
- **Robotics**: STM32, FreeRTOS; CAN/I^2C; Flexiv RDK; PID/dynamics; C++/Rust firmware, MATLAB
- **Other**: JavaScript, HTML, Markdown, LaTeX; Neovim (LSP via Mason), SSH, tmux, GDB/LLDB

Honors & Awards

- **ICRA Best Poster Award (May 2025)**: Presented by IEEE Robotics and Automation Society.
- **Cheng–Family Scholarship (Jun. 2024)**: (Top 2%).
- **RoboMaster University Championship (May 2024)**: (Eastern Region Champion).
- **RoboMaster University League (Apr. 2024)**: (Shanghai Division Champion).
- **University Physics Competition (Nov. 2023)**: (Silver Prize).
- **SJTU Excellence Scholarship, Level B (Dec. 2023)**: (Top 3%).
- **John Wu & Jane Sun Sunshine Scholarship (Nov. 2023)**: (Top 6%).
- **RoboMaster University Championship (Aug. 2023)**: (National Champion).
- **SJTU Social Practice, Third Prize (Aug. 2023)**:

Teaching Experience

- **Teaching Assistant (2024)**: **Teaching Assistant** at Shanghai Jiao Tong University, ENGR 1000J (Introduction to Software Engineering).

Extra Curriculars

- **Undergraduate Research Assistant (May 2025 – Ongoing)**: Undergraduate Research Assistant at University of Michigan College of Engineering, MI, USA.
- **ICRA Volunteer (2025)**: **Volunteer** at IEEE International Conference on Robotics and Automation (ICRA), Atlanta, GA, May 2025.
- **UM-SJTU Joint Institute Youth Volunteer Team (2023)**: UM-SJTU Joint Institute Youth Volunteer Team member (Shanghai, China).
- **Old Friends Youth Team (2023)**: Old Friends Youth Team, Shanghai, Facilitated intergenerational communication activities.

Personal Details

- **Language**: English (TOEFL 107/120), Chinese (Native).
- **Hobbies**: Badminton, Rubik's Cube, Linux Customization (Arch Linux + Hyprland + NeoVim + Fcitx5).