

# Yifeng Cao

☎ +1 6467502060 | ✉ yc4317@columbia.edu | 🔗 LinkedIn | 🐙 GitHub | 📁 Portfolio | 📍 New York

## EDUCATION

---

### Columbia University

*Master of Electrical and Electronics; GPA: 3.82/4.00*

New York, USA

Sep 2023 – Dec 2024

### North China University of Technology (NCUT)

*Bachelor of Engineering in Automation; GPA: 3.88/4.00*

Beijing, CHN

Sep 2018 – Jun 2022

## ACADEMIC EXPERIENCE

---

### Columbia University, ROAM LAB

New York, USA

*Supervisor: Ph.D. Zhanpeng He and Prof. Matei Ciocarlie*

Sep 2023 – Present

- Pioneered a novel human-in-the-loop framework leveraging diffusion models to identify high-entropy states in robotic manipulation tasks, enabling autonomous determination of optimal intervention points for human expert guidance.
- Demonstrated deep technical expertise by independently implementing and training diffusion models from scratch, achieving robust policy learning across diverse manipulation scenarios in RLbench and Robomimic environments.
- Engineered comprehensive modifications to RLbench and Robomimic simulators at the source code level, designing manipulation tasks specifically crafted to exhibit multi-modal solution spaces and capture high-variance states.
- Architected an interactive human-robot interface utilizing Gaussian Mixture Models (GMM) to decompose multi-modal action distributions, allowing human experts to efficiently guide policy selection at critical decision points.
- Achieved end-to-end system integration, from environment modification and data collection to model training and human-robot interface development, with minimal supervision while receiving only high-level strategic guidance from a human expert.
- Conducted extensive experimentation with state-of-the-art offline reinforcement learning algorithms, implementing and evaluating SAC, BCQ, and CQL for human-in-the-loop robotics applications in D4RL benchmark environments.
- Investigated the fundamental challenges of Q-value overestimation in offline RL, particularly in the context of uncertainty estimation using auxiliary Q-networks, revealing critical instabilities in variance measurements for human intervention timing.

### New York University, CILVR LAB

New York, USA

*Supervisor: Ph.D. Raunaq Bhirangi and Prof. Lerrel Pinto*

Jun 2024 – Present

- Optimized sensor data integration methods, including FiLM layers, frequency adjustments, and data history utilization. Through this process, identified and implemented the most effective model structure, improving policy performance by 15.2% across various tasks.
- Designed contact-rich manipulation tasks (e.g., USB insertion, card swiping), collected extensive data by VR, trained policies, and evaluated them on real-world robotic systems, demonstrating the essential role of tactile sensors in successful task execution.
- Conducted comprehensive ablation studies comparing the performance of AnySkin, DIGIT, and ReSkin sensors, as well as vision-only models, revealing a 43% performance improvement using AnySkin over DIGIT and a 27.5% improvement over vision-only models.
- Developed and executed experiments evaluating AnySkin's signal consistency, demonstrating enhanced reliability in long-term usage across varied environments through signal strength, alignment, and sensor offset tests.

### Columbia University, Creative Machines Lab

New York, USA

*Supervisor: Ph.D. Yuhang Hu and Prof. Hod Lipson*

May 2024 – Present

- Architected an effective human-robot interaction pipeline, integrating Large Language Models (LLMs), facial expression recognition, speech processing, and robotic control to enable naturalistic social interactions.
- Implemented a real-time multi-modal perception system combining Automatic Speech Recognition (ASR) and facial expression recognition algorithms to capture and analyze human emotional and verbal inputs.
- Engineered a novel video synthesis pipeline leveraging text-to-speech technology and advanced animation techniques to generate photo-realistic robot facial animations with synchronized lip movements and natural expressions.

- Designed an innovative control system utilizing Variational Autoencoder (VAE) framework to transform synthesized facial animations into precise motor control signals, enabling accurate physical replication of generated expressions.
- Pioneered a novel application of Diffusion Transformer (DiT) to learn aesthetic object arrangement principles through an end-to-end model, enabling an automatic transformation of cluttered scenes into organized "knolling" patterns with objects aligned in parallel or at right angles.
- Leveraged the model's adaptive layer normalization (adaLN) architecture to effectively process and incorporate messy scene images as conditioning inputs, establishing robust mappings between cluttered arrangements and their corresponding tidy configurations.

## University of Chinese Academy of Sciences (UCAS)

Beijing, CHN

Supervisor: Assoc. Prof. Haoran Li and Prof. Dongbin Zhao

Sep 2019 – Jul 2021

- Developed a high-fidelity Unity-based simulation environment for the Mecanum-wheel chassis and gimbal through Real2Sim dynamic modeling, facilitating efficient and cost-effective testing and training of control algorithms.
- Implemented Sim2Real transfer by refining the simulation model using real-world vehicle data and deploying Python-based reinforcement learning algorithms on the physical robot, resulting in a 20% improvement in control accuracy.
- Created ROS1-based interaction between low-level control system and high-level identification.
- Applied Reinforcement Learning to achieve the autonomous decision-making of the robot, such as path planning, striking and tactics selection.
- Employed YOLOv5 model to train the labeled data for visual recognition development in Python.
- Added A-star algorithm to complete the global planning and navigation, combined visual odometry and wheel odometry to obtain the accurate coordinate of the robot, and applied Bezier Smooth algorithm to assure more stable robot motion. This integrated approach decreased the navigation error from 15cm to 5cm.

## North China University of Technology (NCUT)

Beijing, CHN

Supervisor: Assoc. Prof. Bin Song

Sep 2019 – Jul 2021

- Assisted in a Two-Wheeled Self-balancing Car mechanical design and test by modeling in SolidWorks.
- Programmed with C in Keil v5 to realize the control of the robot, achieved motor control of omnidirectional motion of Mecanum wheel based on CAN communication and PID controller.
- Partook in the PCB design and packaging of the core chip of the robot based on Altium designer.
- Applied Kalman filter and median filter to obtain robust gyroscope data to better ensure the balance of the car.

## ACADEMIC EXPERIENCE

### Team leader, CS Retake Mode with Behavioral Cloning

Sep 2023 – Dec 2023

- Assigned tasks based on each person's expertise and developed corresponding timetables
- Created a Behavior Cloning agent for Counter-Strike: Global Offensive (CSGO) retake mode using Video Vision Transformers (ViViT)

### Large Language Model (LLM) Game NPC

May 2023 – Sep 2023

- Developed a game testing environment with C# and Unity, integrating Python scripts to interact with the LLM component, enabling the creation of intelligent, context-aware NPCs.
- Implemented Attention Mechanism to enhance NPC's context-awareness, and fine-tuned prompts to endow NPCs with game-specific self-perception and adaptive backstories, ensuring coherent and immersive player-NPC interactions.

## PUBLICATION

Pattabiraman, V., **Cao Yifeng**., Haldar, S., Pinto, L. and Bhirangi, R., 2024. Learning Precise, Contact-Rich Manipulation through Uncalibrated Tactile Skins. arXiv preprint arXiv:2410.17246.

Bhirangi, R., Pattabiraman, V., Erciyes, E., **Cao Yifeng**., Hellebrekers, T. and Pinto, L., 2024. AnySkin: Plug-and-play Skin Sensing for Robotic Touch. arXiv preprint arXiv:2409.08276.

**Cao Yifeng**., Wu, Y., Tian, Z. and Yu, X., 2021, September. An auxiliary tool for preliminary tests of skin cancer: A self-modifying meta-learning method for clean and noisy data. In 2021 2nd International Conference on Big Data & Artificial Intelligence & Software Engineering (ICBASE) (pp. 172-176). IEEE.

## SKILLS

**Languages and IT Skills:** C/C++, C#, Python, Matlab/Simulink, ROS1, ROS2, Altium Designer, Unity, SolidWorks, LabVIEW, G-code, Adobe Premiere, Linux, LaTeX