

Jacob Gong

Address: NYC ✉ yuwendgong2c@gmail.com ☎ 646-512-1792

🌐 Personal Page

🐙 GitHub

🌐 LinkedIn

EDUCATION

Columbia University

Aug 2025 - Present

BA in Computer Science

- GPA: 4.26/4.33
- Coursework: Advanced Programming, Project in MECH ENG

City University of Hong Kong

Aug 2023 - June 2025

BS in Computer Science

- iGPA: 3.90/4.0 (top 7%)
- Coursework: Computational Probability Modeling, Advanced Algorithms, Software Design, Machine Learning
- **Dean's List**: All Semesters

RESEARCH & JOB EXPERIENCE

Research Assistant

Creative Machine Lab @Columbia University

Aug 2025 - present

- Track: Robotics, Computer Vision, GPU Simulation
- Skills: **Model Implementation & Optimization**, Supervised Fine Tuning, Reinforcement Learning, Data Synthesis, Mass-Spring Physics, NEAT/CPPN, Taichi Lang, NVIDIA Warp, Reverse-Mode AD
- Results: Add vision modality into a robot self model pipeline;
 - Exploiting fine-grained segmentation in preliminary vision stage;
 - Built two GPU-accelerated frameworks for joint morphology-control co-optimization of voxel-based soft robots (see Projects)

Distributed and Mobile Sensing Computing Group @City University of HK

Jan 2025 - Aug 2025

- Track: Edge AI
- Skills: Robust Server Implementation, Policy Design
- Results: Design & Develop mixed inference middleware for edge inference (lowered 98.82% memory usage & 7.32% latency)

PROJECTS

Voxel-Based Soft Robot Co-Design

see personal page

- Track: Robotics, GPU Simulation, Evolutionary Optimization
- Skills: Mass-Spring Physics, NEAT, CPPN, Taichi Lang, NVIDIA Warp, Reverse-Mode AD (wp.Tape)
- Results & Findings: Built two GPU-accelerated frameworks for **joint morphology-control co-optimization** of voxel-based soft robots; **Diff-CoDesign** (Taichi): NEAT-evolved CPPNs decode 10^3 voxel morphologies with 5 material types; batch-evaluated 200 robots/generation across 35×1000 -generation runs (7M total evaluations, mean fitness 1.33); **VoxelCoDesign** (Warp): re-engineered physics with a **zero-atomic-operation, two-pass gather architecture**, enabling deterministic gradients via wp.Tape for hybrid evolutionary + gradient-based co-design; Achieved $10.8 \times$ topology build speedup and validated gradient backprop through full physics pipeline; Compared against **voxJax** (JAX, implicit atomics, $30 \times$ softer stiffness) and **PhysT-win** (Warp, explicit atomics, single-robot system ID): our atomic-free design trades compilation simplicity for **determinism and portability**

DriftWorld

World Model Framework for Robot Manipulation

- Track: Generative Models, Robotics, World Models
- Skills: Diffusion Models, Flow Matching, VAE, **Drifting Models (novel architecture)**, DiT, SD-VAE, DINOv2
- Results & Findings: Built a modular world model framework predicting **next visual state from (image, action)** for robot manipulation; Implemented and benchmarked **4 generative architectures** (VAE, DDPM, Flow Matching, Drifting Models) on CALVIN with a shared DiT-S/2 backbone operating in SD-VAE latent space; Trained RL policy in imagination using the world model, demonstrating a **complete world-model-to-policy pipeline** with zero-shot sim transfer; Drifting Models achieve **1-step inference** ($50 \times$ faster rollouts than diffusion), enabling real-time planning

Task-agnostic Robot Modeling

Task-agnostic Robot Modeling experiment record(2025 Fall)

- Track: Computer Vision, Robotics, 3D Reconstruction
- Skills: Vision Foundation Models (DINO v2), 3D Geometry (VGGT), URDF Modeling, Video Segmentation
- Results & Findings: Developed a pipeline to **reconstruct robot URDFs from 2D video**, replacing the need for expensive 3D sensors; Proposed an **”early-stage mask fusion”** strategy, feeding DEVA-masked images into VGGT to resolve 2D artifacts; Fine-tuned **DINO v2** on PartNet-Mobility for semantic link-level segmentation and kinematic signature extraction; Conducted **failure analysis on 4D trajectory fields**, identifying depth jitter as a bottleneck for motion clustering

SKILLS

- **Languages:** Python, C++, PyTorch, Taichi Lang, Linux (Ubuntu)
- **Comp. Vision:** 3D/4D Reconstruction (VGGT, DUST3R), Video Object Segmentation (DEVA, SAM), 4D Trajectory Fields
- **Deep Learning:** Vision Foundation Models (DINO v2), Diffusion/Flow Matching, World Models, SFT, Transformer (DiT)
- **Robotics:** Kinematic Modeling (URDF, SDF, MJCF), Physics Engines (Taichi, NVIDIA Warp, MuJoCo), Model-Based RL
- **Robustness:** Failure Case Diagnosis (Depth Jitter), Noise/Artifact Mitigation (Early-stage Mask Fusion)