# Chuyao Fu

12310306@mail.sustech.edu.cn | github.com/ChuyaoZellFu | chuyaozellfu.github.io | +86 153 4733 9211
Shenzhen, China / Undergraduate researcher in Vision-Language-Action Models, World Models, and Autonomous Driving.

## Education

**Southern University of Science and Technology** Sep. 2023 – Jun. 2027 (expected)
*B.Eng. in Electronic Information Engineering* *GPA: 3.90/4.00 Rank: 3/46*

- **Selected Coursework:** Data Structures and Algorithm Analysis, Probability and Statistics, Machine Learning, Robotic Motion and Control, Digital Signal Processing, Signals and Systems, Analog Circuits, Digital Circuits

## Publications

**ProDrive:** Proactive Planning for Autonomous Driving via Ego-Environment Co-Evolution
**Chuyao Fu**, S. Gan, Z. Ouyang, Y. Rui, Z. Qian, X. Chi, S. Han, J. Wang, Z. Hong
*CVPR 2026 GigaBrain Challenge Workshop (Poster)*

**Token-World:** Native Latent Dynamic World Model for Vision-Language-Action Models
**Chuyao Fu**, X. Chi, H. Li, Y. Rui, Z. Qian, Y.-K. Wang, X. Zhang, Y. Lou, H. Cheng, Y. Guo, S. Han, S. Zhang.
*ECCV 2026 (under review)*

**EchoArena:** Mutual Benchmarking of World Models and VLA Policies
Y.-K. Wang, K. Zhang, X. Chi, T. Chen, S. Huang, **Chuyao Fu**, T. Guo, P. Jia, Y. Qin, K. Ge, S. Qian, W. Mi, Z. Qian, J. Li, Q. Wuwu, S. Zhang
*ECCV 2026 (under review)*

**FORCE:** Efficient VLA Reinforcement Fine-Tuning via Value-Calibrated Warm-up and Self-Distillation
S. Zhang, Y. Lou, H. Cheng, Y. Guo, **Chuyao Fu**, Y. Lyu, X. Zhang, H. Li, P. Wang, Z. Wang, S. Zhang
*ICML 2026 (under review)*

**Predicting What Matters:** Robust Generalist Robot Policy Learning via Future Semantic Mask
Y. Lou, X. Chi, X. Zhang, Z. Qian, C. Li, R. Zhang, Y. Lyu, G. Song, **Chuyao Fu**, H. Xu, P. Wang, S. Zhang
*ICML 2026 (under review)*

## Research Experience

**Robotics and Computer Vision Lab, Southern University of Science and Technology** Sep. 2024 – Present
*Undergraduate Researcher* *Supervisor: Prof. Hong Zhang*

- **Topics:** Social Navigation and Autonomous Driving.
  - Contributed to **TPT-Bench** data collection for social navigation and gained hands-on experience with ROS bag processing, multisensor synchronization, and data preprocessing (*paper under review at IJRR*).
  - Independently proposed **ProDrive**, a proactive autonomous driving framework that jointly trains a driving planner and a BEV world model via *ego-environment co-evolution*.
  - Demonstrated stronger performance than several prior world-model-based driving frameworks on NAVSIM v1, improving both navigation efficiency and safety.
  - **First author**; led method design, experiments, and manuscript preparation. Accepted by the *CVPR 2026 GigaBrain Challenge Workshop*.

**HMI Lab, School of Computer Science, Peking University** Jun. 2025 – Present
*Undergraduate Researcher* *Supervisor: Prof. Shanghang Zhang*

- **Topic:** Vision-Language-Action Models, World Models, and Latent-Space Simulation.
  - Proposed **Token-World**, a latent world model for VLA systems that predicts future visual tokens directly in token space, eliminating intermediate RGB reconstruction.
  - Designed a spatio-temporal multimodal Transformer with GRU-enhanced temporal modeling for action-conditioned long-horizon prediction.
  - Combined flow matching with diffusion forcing to improve rollout stability and reduce autoregressive error accumulation.
  - Developed a latent reward model and imagined rollout evaluation pipeline for policy assessment in latent space.
  - Outperformed strong RGB-based world-model baselines on RoboTwin in rollout error, semantic similarity, and inference speed.
  - **First author**; led method design, experiments, and manuscript preparation. Submitted to *ECCV 2026* (under review).

**Beijing Academy of Artificial Intelligence** Jun. 2025 – Nov. 2025
*Research Intern, Embodied Multimodal Large Models Research Center* *Mentor: Xiaojie Zhang*

- **Topic:** Reinforcement Learning for Vision-Language-Action Models and Flow-Based Robot Policies.
  - Reproduced representative VLA-RL and robot policy post-training frameworks, including *RLinf*, *SimpleVLA-RL*, *RL4VLA*, and *DSRL*.
  - Investigated reinforcement learning objectives for flow-based robot policies, focusing on training stability, inference efficiency, and compatibility with modern VLA backbones.
  - Contributed to the early development of the **FORCE** project through idea exploration, baseline execution, and result organization.
  - Supported manuscript preparation, including paper polishing and figure refinement; co-authored an *ICML 2026* submission (under review).

## Work Experience

**XtalPi** Jan. 2026 – Present
*Research Intern, Future Chemistry Department* *Shenzhen*

- **Topic:** Humanoid Robots for Laboratory Automation; Vision-Language-Action Policy Learning and Deployment.
  - Trained VLA policies for humanoid robot manipulation in laboratory environments, focusing on tool use and structured object interaction.
  - Deployed policies on real humanoid robot platforms and integrated perception, policy inference, and control for real-world execution.
  - Developed manipulation capabilities for representative laboratory instruments and containers in automated bio-lab scenarios.
  - Conducted iterative real-robot evaluation and failure analysis to improve deployment robustness.

**Beijing Academy of Artificial Intelligence** Jun. 2025 – Nov. 2025
*Research Intern (Data Pipeline), Embodied Multimodal Large Models Research Center* *Beijing*

- **Topic:** Pretraining Data Pipelines for Vision-Language-Action and Vision-Language Models.
  - Processed large-scale embodied datasets including *AgiBot-World* and *DROID* for VLA pretraining.
  - Performed action normalization, action tokenization, and dataset reorganization into structured `JSON` training format.
  - Standardized multimodal trajectory data for compatibility with large-scale VLA pretraining pipelines.
  - Contributed to VLM data preprocessing by adding grounding boxes as object-aware visual-language supervision.

## Selected Course Projects

**Raspberry Pi-based Robotic Arm Control** Nov. 2025 – Jan. 2026
*Embedded Systems / Robotics* *Project Lead*

- Developed a Raspberry Pi-based control system for a 3D-printed robotic arm, using GPIO PWM signals to actuate an SG52 servo for gripper control and four-phase eight-step drive signals for 4-phase 5-wire stepper motors at the arm joints.
- Implemented two interaction modes: infrared remote control via `lirc` and camera-based gesture control using MediaPipe for real-time arm manipulation.

**Desktop Manipulation via 3D Diffusion Policy** Mar. 2025 – Jun. 2025
*Robot Learning / 3D Vision* *Project Lead*

- Built and deployed a real-world 3D Diffusion Policy (DP3) pipeline for tabletop manipulation, including multimodal data collection, camera calibration, point-cloud preprocessing, policy training, and online inference.
- Validated the system on real tasks including lemon-in-bowl placement, bowl lifting, and drawer pushing using point-cloud observations from a RealSense D435 and an ARX X5 robotic arm.

## Leadership and Service

**Class Leader, Class 2303** 2024 – 2025
*Class Committee*

- Led class operations and activities, including event planning, logistics coordination, routine administration, and award application materials; the class received honors including *Excellent Undergraduate Academic Style*, *Best Sports Team*, and *Best Popularity Class*.

**House Division Lead, Street Dance Club** 2024 – 2025
*Student Organization*

- Led the House division by coordinating weekly training with instructors and organizing two campus street dance performances, including rehearsal planning and member coordination.