

Jianman Lin

linjianmancjx@gmail.com | [Homepage](#) | [Google Scholar](#)

RESEARCH PROFILE

Pursuing an M.Eng. in Electronic Information at **South China University of Technology (SCUT)** (Sep. 2024 – Present), advised by [Prof. Tianshui Chen](#) and [Prof. Chunmei Qing](#). Previously earned a B.Eng. in Industrial Engineering at **Guangdong University of Technology (GDUT)** (Sep. 2020 – June 2024; GPA: 87.2/100).

My research spans **embodied intelligence, multimodal learning, and generative modeling**, with the goal of building agents that generalize beyond their training environments and translate perception and language into reliable physical interaction.

SELECTED PUBLICATIONS

- [1] [Breaking the Vision–Action Shortcut: Latent Interface Training for Generalizable Robot Foundation Models](#)
Co-first author; **ICRA** (under review).
- [2] [Learning Spatial-Temporal Coherent Correlations for Speech-Preserving Facial Expression Manipulation](#)
Co-first author; **IEEE TPAMI**, 2026.
- [3] [Neural Scene Designer: Self-Styled Semantic Image Manipulation](#)
First author; **IEEE TIP**, 2025.
- [4] [Contrastive Decoupled Representation Learning and Regularization for Speech-Preserving Facial Expression Manipulation](#)
Co-first author; **IJCV**, 2025.
- [5] [Learning Adaptive Spatial Coherent Correlations for Speech-Preserving Facial Expression Manipulation](#)
Co-author; **CVPR**, 2024. **Highlight (top 2.8%)**.

RESEARCH EXPERIENCE

LIT: Generalizable Robot Foundation Models [1] 2026 – Present

- Developed two-stage training to **improve robot-policy generalization**: image-free, spatial-goal-conditioned action pretraining followed by a pose-supervised interface as the action expert’s only visual conditioning pathway.
- Improved LIBERO-Plus success by **3.87–10.70 percentage points** across four VLA/WAM architectures and real-world success by **13.3–16.7 points** across three tasks under unseen cameras, lighting, and distractors.

Multimodal Representation Learning & Image Editing [2–5] 2023 – 2025

- Developed speech-preserving expression editing and controllable diffusion-based image manipulation through representation disentanglement, spatial-temporal correlation learning, and style-preserving generation.

WORK EXPERIENCE

Research Intern: AI & Robotics

[X-era.ai](#) & [ExpanderaAI Lab](#)

Dec. 2023 – Present

Guangzhou, China

- Trained and deployed VLA policies for autonomous ~5-minute manipulation; built pipelines spanning hardware setup, teleoperation, data collection, training, and deployment on AgileX Piper, UR5, ARX, and Zerith arms.
- Earlier projects: VLM/LLM-based interior design and StyleGAN-based digital human systems.

ACADEMIC SERVICE & SKILLS

Service & IP: Reviewer for ICRA, CVPR, NeurIPS, ECCV; **three granted Chinese patents**.

Technical: PyTorch, Transformers/Diffusers, Python, C/C++, CUDA, Linux; teleoperation and real-robot deployment.

Language: English – **IELTS Overall: 7.0**.