

Bo (Albert) Gao

✉ bogao@andrew.cmu.edu | ✉ golbert2002@163.com | 🌐 Homepage | 🐙 GitHub | 📄 Google Scholar

Education

Carnegie Mellon University

MSc in Artificial Intelligence Engineering – Electrical and Computer Engineering

- GPA: 4.00/4.00. Advisor: Prof. Andrea Zanette and Prof. Changliu Liu.

Pittsburgh, USA

Aug. 2025 – Dec. 2026 (expected)

Sun Yat-sen University

B.Eng. in Intelligence Engineering

- GPA: 3.9/4.0; Rank: 3/110. Advisor: Prof. Xiaodan Liang.

Shenzhen, China

Sep. 2021 – Jun. 2025

Publications

* Equal contribution.

What Makes Two Robot Failures the Same? Learning Recoverability Functions for Zero-Shot Robot Recovery

Siyuan Ma, Juntao Cheng, Liwei Hou, **Albert Gao**, Xin Yan, Qianshan Wei, Yanbo Zhao, Weida Zhao, Zhisheng Chen, Qixin Zhang, Zheng Zhu
arXiv preprint, 2026 Under Review

Learning to Outgrow a Theory: Experimental Discovery Beyond the Initial Hypothesis Space

Siyuan Ma*, **Albert Gao***, Chunzheng Zhu, Xin Yan, Wenlong Zhang, Wenxin Zhang, Luqi Gong, Tianlin Li, Qixin Zhang
arXiv preprint, 2026 Under Review

SLVR: Structured Latent Visual Reasoning via Human-like Reasoning Flows

Albert Gao*, Bing Xue, Andrea Zanette
The Fortieth Annual Conference on Neural Information Processing Systems (NeurIPS), 2026 — Poster Accepted

Linear Fitness Subspace in Protein Language Models Enables Sample-Efficient Directed Evolution

Siyuan Ma*, Canran Xiao, Zikai Xiao, **Albert Gao**, Liang He, Xuan-Yu Wang, Shuying Cao, Xiaojun Jia
Empirical Methods in Natural Language Processing (EMNLP), 2026 — Findings / Poster Accepted

BOOKAGENT: Orchestrating Safety-Aware Visual Narratives via Multi-Agent Cognitive Calibration

Bo Gao*, Chang Liu, Yuyang Miao, Siyuan Ma, Ser-Nam Lim
Annual Meeting of the Association for Computational Linguistics (ACL), 2026 — Findings/ Poster Accepted

Free-Mask: A Novel Paradigm of Integration Between the Segmentation Diffusion Model and Image Editing to Improve Segmentation Ability

Bo Gao*, Jianhui Wang, Xinyuan Song, Yangfan He, Fangxu Xing, Tianyu Shi
ACM International Conference on Multimedia (ACM MM), 2025 — Oral Accepted

Exploring Warping-Guided Features via Adaptive Latent Diffusion Model for Virtual Try-on

Bo Gao*, Junchi Ren, Fei Shen, Mengwan Wei, Zijun Huang
IEEE International Conference on Multimedia and Expo (ICME), 2024 — Oral Accepted

Research Experience

Carnegie Mellon University

Researcher Intern — Advisor: Prof. Andrea Zanette

- Developed structured latent visual reasoning methods that replace verbose chain-of-thought with compact, typed intermediate representations for efficient and interpretable multimodal reasoning.
- Investigating reflection as an explicit action in VLM agents, learning when to gather visual evidence, use tools, reflect on failures, or stop through counterfactual and verifiable feedback.

Pittsburgh, PA

2025 – Present

Carnegie Mellon University

Research Intern — Advisor: Prof. Changliu Liu

- Developing controllable 3D generation for LEGO assembly, using LLMs to translate language instructions into structured spatial representations such as coordinates and assembly plans under geometric constraints.

Pittsburgh, PA

2026 – Present

Everlyn

Research Intern — Advisor: Dr. Ser-Nam Lim

- Developed multi-agent frameworks for long-form visual storytelling, coordinating narrative planning and generation to improve controllability, coherence, and consistency across scenes.

Hong Kong, China

Mar. 2025 – May. 2026

Tsinghua University

Research Intern — Advisor: Prof. Jun Zhu

- Developed disentangled diffusion-based image editing with controllable representations for color, shape, and structure, using Q-Former and attention-based feature integration.

Beijing, China

Jul. 2024 – Oct. 2024

Honors & Awards

- National Scholarship; First Prize Scholarship; Academic Innovation Scholarship; Special Research Scholarship, Sun Yat-sen University.
- Top 3% (24/929), Kaggle Image Matching Challenge 2024 – Hexathlon (CVPR 2024 Workshop).
- First Prize, National College Students Mathematical Competition, 2022.
- Meritorious Winner, Mathematical Contest in Modeling (MCM), 2023.

Academic Service

Reviewer: npj Digital Medicine (Nature Portfolio), AAAI, NeurIPS, ICLR, and CVPR.

Skills

Programming: Python, C/C++, MATLAB, L^AT_EX, Linux.

ML: PyTorch; multimodal LLM/VLM training; diffusion models; SFT/RL post-training.