

Yiming Qin

Beijing, China | ymk4474@gmail.com | wakals.github.io

Bio

I am currently a senior undergraduate student at **Zhi Class of Peking University (PKU)**. I am also a visiting student at UC Berkeley and doing research with my mentor **Prof. XuDong Wang** and advisor **Trevor Darrell** at **Berkeley Artificial Intelligence Research Lab (BAIR)**. Before that, I had a wonderful time at **Wangxuan Institute of Computer Technology (WICT)**, advised by **Prof. Yang Liu**. I have also spent a good time with Dr. Jialun Liu at TeleAI.

My research interest is Multimodal Learning, and especially focuses on **Vision-Language Models and Generative Models**. My research vision is to develop AI system with precise reasoning and generation capability.

Education

Zhi Class, Peking University , BS in Yuanpei College	2022.09 – 2026.06 (expc.)
• GPA: 3.71/4.00 rank: 9/69	

University of California, Berkeley Extension , FES in EECS	2025.01 - 2025.06
---	-------------------

Awards

- THE THIRD SCHOLARSHIP AT PEKING UNIVERSITY
- PEKING UNIVERSITY FRESHMAN SCHOLARSHIP
- YUANPEI RESEARCH EXCELLENCE AWARD

Research Experience

Trevor Darrell @ BAIR , Vision Group Intern - Berkeley, CA, USA	2025.01 – now
• Mentor: Prof. XuDong Wang and Prof. Trevor Darrell	
• Topic: Enhance the visual reasoning ability of Large Vision-Language Model.	

Jialun Liu @ TeleAI , AIGC Intern – Beijing, CN	2025.08 – 2025.10
• Topic: Video Editing.	

Yang Liu @ Wangxuan Institute , Computer Vision Intern – Beijing, CN	2023.06 – 2025.02
• Topic: Multimodal Generative model.	

Publications

Chain-of-Visual-Thought: Teaching VLMs to See and Think Better with Continuous Visual Tokens , Preprint	2025.11
<i>Yiming Qin</i> , Bomin Wei, Jiaxin Ge, . . . , Trevor Darrell, XuDong Wang	
Link: Arxiv, Project Page, Github	
Apply Hierarchical-Chain-of-Generation to Complex Attributes Text-to-3D Generation , Computer Vision and Pattern Recognition (CVPR) 2025	2025.02
<i>Yiming Qin</i> , Zhu Xu, Yang Liu	
Link: Arxiv, Project Page, Github	
Diff-BGM: A Diffusion Model for Video Background Music Generation , Computer Vision and Pattern Recognition (CVPR) 2024	2024.02
<i>Sizhe Li, Yiming Qin, Minghang Zheng, Xin Jin, Yang Liu</i>	
Link: Arxiv, Project Page, Github	