CHUN-HSIAO (DANIEL) YEH

daniel_yeh@berkeley.edu, danielchyeh.github.io, (+1) 510-910-0254

RESEARCH INTEREST

My research bridges multimodal understanding and controllable generation. I am interested in enhancing the fundamental perception of Multimodal Large-Language Models (MLLMs), particularly focusing on spatial reasoning and geometric awareness to ground models in the physical world. Recently, I have been leveraging these capabilities to advance on controllable generative pipelines that can decompose and execute complex, instruction-based image editing and synthesis with high controllability.

EDUCATION

University of California, Berkeley — Aug. 2021 - Present
Ph.D. in Vision Science (Computational Vision Track) — *Berkeley, CA*
Advisors: Yi Ma (EECS), Meng C. Lin (VS)

National Taiwan University — Sept. 2015 - Mar. 2019
M.S. in Engineering Science, Division of Computer Science — *Taipei, Taiwan*

RESEARCH EXPERIENCE

FAIR, Meta — May 2025 - Nov. 2025
Research Scientist Intern — *Menlo Park, CA*
Mentors: Fanyi Xiao, Shengyi Qian, Manchen Wang, Joseph Tighe
Project: VLM with Geometric Priors for Spatial Reasoning and 3D Understanding

Adobe Research — May 2024 - Mar. 2025
Research Scientist Intern — *San Jose, CA*
Mentors: Krishna Singh, Yilin Wang, Nanxuan Zhao, Yuheng Li, Richard Zhang
Project: X-Planner: Complex Instruction-Based Image Editing (AAAI 2026) [C6]

Adobe Research — May 2022 - Aug. 2022
Research Scientist Intern — *San Francisco, CA*
Mentors: Simon Jenni, Fabian Caba, Bryan Russell, Josef Sivic
Project: Personalized Vision-Language Models for Video Retrieval (CVPR 2023) [C3]

FAIR, Meta — NYU — Sep. 2020 - Aug. 2021
Research Collaboration — *Remote*
Collaborators: Yubei Chen, Yann LeCun
Project: Decoupled Contrastive Representation Learning (ECCV 2022) [C2]

SELECTED PUBLICATION

(C: Conference, *: Equal Contribution)

[C6] Beyond Simple Edits: X-Planner for Complex Instruction-Based Image Editing
Chun-Hsiao Yeh, Yilin Wang, Nanxuan Zhao, Richard Zhang, Yuheng Li, Yi Ma, Krishna Kumar Singh
Association for the Advancement of Artificial Intelligence (**AAAI**), 2026

[C5] Seeing from Another Perspective: Evaluating Multi-View Understanding in MLLMs
Chun-Hsiao Yeh*, Chenyu Wang*, Shengbang Tong, Ta-Ying Cheng, Ruoyu Wang, Tianzhe Chu, Yuexiang Zhai, Yubei Chen, Shenghua Gao, Yi Ma
Association for the Advancement of Artificial Intelligence (**AAAI**), 2026

[C4] Gen4Gen: Generative Data Pipeline for Generative Multi-Concept Composition
Chun-Hsiao Yeh*, Ta-Ying Cheng, He-Yen Hsieh, David Chuan-En Lin, Yi Ma, Andrew Markham, Niki Trigoni, H.T. Kung, Yubei Chen
British Machine Vision Conference (**BMVC**), 2025

[C3] Insight: A Multi-Modal Diagnostic Pipeline using LLMs for Ocular Surface Disease Diagnosis
Chun-Hsiao Yeh, Jiayun Wang, Andrew D. Graham, Andrea J. Liu, Bo Tan, Yubei Chen, Yi Ma, Meng C. Lin
Medical Image Computing and Computer Assisted Intervention (**MICCAI**), 2024

[C2] Meta-Personalizing Vision-Language Models to Find Named Instances in Video
Chun-Hsiao Yeh, Byran Russell, Josef Sivic, Fabian David Caba Heilbron, Simon Jenni
Conference on Computer Vision and Pattern Recognition (**CVPR**), 2023

[C1] Decoupled Contrastive Learning
Chun-Hsiao Yeh, Cheng-Yao Hong, Yen-Chi Hsu, Tyng-Luh Liu, Yubei Chen, Yann LeCun
European Conference on Computer Vision (**ECCV**), 2022

PROFESSIONAL ACTIVITIES

Conference Reviewer
ICLR '25–'26, NeurIPS '25, CVPR '23–'26, ECCV '22 & '24, ICCV '23 & '25, AAAI '26, SIGGRAPH '25, MICCAI '24–'25, CPAL '24–'25