

Mingzhe Zheng

2-year PhD Student @ HKUST | 5 years of AI | Pretraining/Post-training/Agent | Philosopher

 <https://mingzhezheng.homes/>  mzhengar@connect.ust.hk  [LinkedIn](#)  [GitHub](#)  [Google Scholar](#)

EDUCATION

Hong Kong University of Science and Technology, HongKong SAR

Feb.2025 - Present

Ph.D. in Computer Vision and Machine Learning

Supervisor: Prof. Harry Yang, Prof. Qifeng Chen

Northwestern Polytechnical University, Xi'an, China (1985)

Sept.2020 - July.2024

B.Eng. in Artificial Intelligence

Outstanding Graduate Award

RESEARCH INTERESTS

Field: Foundation Models, Agentic AI, Multimodal AI, Movie Creation

Orientation: My long-term research goal is to create a visual-centric agentic system with a self-evolving "brain" that can comprehend the environment and generate reactions to real-life interactions, then produce cinematic-level visual content similar to how human directors make films, titled "AI Movie Industry."

SELECTED PAPERS

Agentic System, Agentic RL, RL, LLM Bench

Mingzhe Zheng, et al *EmoReward: Self-Evolving Reward Calibration for Emotion-Aware Video Generation RLHF ongoing project (For HunyuanVideo2.0)*

Mingzhe Zheng, et al *Agentic Rubrics synthesis and judging for Image-Editing Reward Modeling ongoing project (For HunyuanImage3.5)*

Mingzhe Zheng*, Weijie Kong*, Yue Wu, Dengyang Jiang, Yue Ma, Xuanhua He, Bin Lin, Kaixiong Gong, Zhao Zhong, Liefeng Bo, Qifeng Chen, Harry Yang. *Manifold-Aware Exploration for Reinforcement Learning in Video Generation arXiv preprint arXiv:2603.21872*

Mingzhe Zheng, Dingjie Song, Guanyu Zhou, Jun You, Jiahao Zhan, Xuran Ma, Xinyuan Song, Ser-Nam Lim, Qifeng Chen, Harry Yang. *CML-Bench: A Framework for Evaluating and Enhancing LLM-Powered Movie Scripts Generation arXiv preprint arXiv:2510.06231*

Dengyang Jiang, Xin Jin, Dongyang Liu, Zanyi Wang, **Mingzhe Zheng**, Ruoyi Du, Xiangpeng Yang, Qilong Wu, Zhen Li, Peng Gao, Harry Yang, Steven Hoi. *D-OPSD: On-Policy Self-Distillation for Continuously Tuning Step-Distilled Diffusion Models arXiv preprint arXiv:2605.05204*

Foundation Models, Video Generation/Editing/Understanding, Image Generation/Editing
Hunyuanvideo 1.5 technical report arXiv preprint arXiv:2511.18870

Mingzhe Zheng, Yongqi Xu, Haojian Huang, Xuran Ma, Yexin Liu, Wenjie Shu, Yatian Pang, Feilong Tang, Qifeng Chen, Harry Yang, Ser-Nam Lim. *VideoGen-of-Thought: Step-by-step generating multi-shot video with minimal manual intervention NeurIPS 2025 NextVid Workshop (Oral)*

Zeqian Long*, **Mingzhe Zheng*(co-first)**, Kunyu Feng*, Xinhua Zhang, Hongyu Liu, Harry Yang, Linfeng Zhang, Qifeng Chen, Yue Ma. *Follow-Your-Shape: Shape-Aware Image Editing via Trajectory-Guided Region Control ICLR 2026*

Kaixiong Gong, Xin Cai, Bin Lin, Hao Wang, Yunlong Lin, **Mingzhe Zheng**, Bohao Li, Jian-Wei Zhang, Miles Yang, Zhao Zhong, Liefeng Bo, Xiangyu Yue *Twins: Learn to Predict Unified Representations with Focal Loss ICML 2026*

Yue Ma, Xinyu Wang, Qianli Ma, Qinghe Wang, **Mingzhe Zheng**, Xiangpeng Yang, Hao Li, Chongbo Zhao, Jixuan Ying, Harry Yang, Hongyu Liu, Qifeng Chen *Group Editing: Edit Multiple Images in One Go CVPR 2026*

Yue Ma, Zhikai Wang, Tianhao Ren, **Mingzhe Zheng**, Hongyu Liu, Jiayi Guo, Mark Fong, Yuxuan Xue, Zixiang Zhao, Konrad Schindler, Qifeng Chen, Linfeng Zhang *FastVMT: Eliminating Redundancy in Video Motion Transfer ICLR 2026*

Xuran Ma, Yexin Liu, Yaofu Liu, Xianfeng Wu, **Mingzhe Zheng**, Zihao Wang, Ser-Nam Lim, Harry Yang. *Model Reveals What to Cache: Profiling-Based Feature Reuse for Video Diffusion Models ICCV 2025*

Yatian Pang, Bin Zhu, Bin Lin, **Mingzhe Zheng**, Francis EH Tay, Ser-Nam Lim, Harry Yang, Li Yuan. *DreamDance: Animating Human Images by Enriching 3D Geometry Cues from 2D Poses ICCV 2025*

Haodong Chen, Haojian Huang, Junhao Dong, **Mingzhe Zheng**, Dian Shao. *Finecliper: Multi-modal fine-grained clip for dynamic facial expression recognition with adapters ACM MM 2024*

INDUSTRY EXPERIENCES

Tencent, Shenzhen, China	July. 2025 – Present
Department: Tencent Hunyuan Foundation Model Team	Mentor: Dr. Yue Wu
Position: Qingyun Research Intern (Top Talent Internship)	Internship/Full-time

- Contributed to the large-scale Pre-training process of Hunyuan Video 1.5.
- Developed the cornerstone of the GRPO Post-training paradigm for Hunyuan Video (SAGE-GRPO).
- Participating in the agentic data recipe for Hunyuan Video foundation models, responsible for the emotion data construction.
- Building cutting-edge Agentic RL paradigms for video generation, contributing to Hunyuan Video foundation models.
- Building frontier Agentic Rubrics Reward system, contributing to Hunyuan Image foundation models.

Everlyn, Shenzhen, China	July. 2024 – July. 2025
Department: EVERLYN LABS	Mentor: Prof. Sernam Lim
Position: Researcher and Algorithm Engineer	Internship/Full-time

- Developed AutoEncoder for Video Generation, built pipelines, and conducted experiments on the image level, prepared for the video level.
- Researching cutting-edge video models and video generation paradigms.

HeyGen, Shenzhen, China	July.2022 - Sept.2022
Department: Surreal Engine Algorithm group (SEA)	Mentor: Dr. Zhibin (Charly) Hong
Position: Computer Vision Algorithm Intern	Internship/Full-time

- Developed the Multilingual Force Alignment Algorithm to support the time matching between input text in 26 languages and corresponding text-to-speech (TTS) audio, which contributed to the first generation of the virtual avatar-driven module in [HeyGen](#).
- Researched and implemented Audio-driven and Video-driven Talking Head Generation technologies.

Shenzhen Qianhai Paisu Technology Co., Ltd. , Shenzhen, China

June.2022 - July.2022

Department: Core Algorithm Department

Position: Machine Learning Algorithm Intern

Internship/Full-time

- Developed a Sleeping Status Prediction System based on heart rate and body movement data acquired by millimeter wave radar, with LightGBM, which is used in [SleepKnow](#).
- Optimized the accuracy of the original algorithm from 72% to 86%, and recorded my model training tricks in the technical document to guide future usage and arrangement.

ACADEMIA EXPERIENCE

Hong Kong University of Science and Technology, Researcher, PhD student

Advisor: [Prof. Harry Yang](#), [Prof. Oifeng Chen](#)

Feb. 2025 - Present

Topic: Generative Models & Multimodality & Movie Generation

- PhD research focused on advancing agentic and multimodal visual generation and movie creation.

StatsLE Lab, University of Toronto, Research Assistant

Advisor: [Prof. Qiang Sun](#) (Department of Statistical Sciences)

Sept.2023 - Mar.2024

Topic: Generative Models & Text-To-3D

- Studying existing text-to-3D architectures and advanced acceleration strategies, writing research surveys and proposals.
- Implementing efficient text-to-3D methodologies, aiming to optimize the 3D generation process with statistical machine learning knowledge.

DianLab, Northwestern Polytechnical University, Research Assistant

Advisor: [Prof. Dian Shao](#) (Unmanned System Research Institute)

May.2023 - June.2024

Topic: Generative Models & Fine-grained Human Face Generation & Multimodality

- Authored an article about Visual Generation with the help of LLMs.
- Authored “FineCLIPER: Multi-modal Fine-grained CLIP for Dynamic Facial Expression Recognition with AdaptERs” accepted by ACM MM 24
- Completed my graduate design titled “High-Fidelity Text-Guided 3D Face Generation based on Diffusion Model in Neural Radiance Field”

National Engineering Lab for Integrated Aerospace-Ground-Ocean Big Data Application

Technology, Northwestern Polytechnical University, Research Assistant

Advisor: [Prof. Wei Wei](#) (Department of Computer Science)

Nov.2022 - May.2023

Topic: Generative Models & Super Resolution

- Developed a project that combines Latent Diffusion Model and VQGAN for Super Resolution.
- Authored an article about Modified Latent Diffusion Models for Visual Classification.

Advisor: [Prof. XinBo Zhao](#) (Department of Computer Science)

Sept.2021 - May.2023

Topic: Human-Computer Interaction & Computer Vision (Reconigntion Direction) & EEG

- Explored Raspberry Pi car controlled by EEG and Fighter Game played with Eyes and Awareness.
- Developed an Eye-animation system that can paint with eye movement.
- Completed an Undergraduate Science and Technology Innovation Project *Cognitive Perception Binary Fatigue Driving Detection based on Deep Learning* as Project Leader.

APPENDIX

Invited Talk: 2026 China Generative AI Conference – Video Generation Venue

Awards: Meshy Finalist Fellowship 2026

Review Services: ECCV, ICCV, CVPR, NeurIPS, ICLR, ICML