WENTAO HU

Phone: (+86) 18980399950 ◇ Email: wentao002@e.ntu.edu.sg
Homepage: wentaohu1208.github.io
Github ◇ LinkedIn

EDUCATION

Nanyang Technological University (QS15) *Nov. 2025 (expected)*
M.Eng in Computer Science and Engineering (research-based program)
GPA: 3.83/5.0
Supervisor: Prof.Hanwang Zhang

Hunan University (Project 985) *Jun. 2023*
B.Sc in Statistics
GPA: 3.55/4.0

RESEARCH INTERESTS

Multimodal Large Language Model, Video Generation

PUBLICATIONS

[1] Anonymous authors (Co-first authors). "Reasoning Physical Video Generation with Diffusion Timestep Tokens via Reinforcement Learning." In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2025. (under review)

[2] Anonymous authors. "On Path to Multimodal Generalist: Levels and Benchmarks." In Proceedings of the International Conference on Machine Learning (ICML), 2025. (under review)

RESEARCH EXPERIENCE

Huawei Central Media Technology Institute, 2012 Laboratory Aug. 2024 - Present
Research Intern Singapore

Autoregressive Video Generation with Reinforcement Learning Aug. 2024 - Mar. 2025
Supervisors: Dr. Wang Lin, Dr. Zhongqi Yue

- Explored autoregressive models for video generation using reinforcement learning, with a focus on capturing real-world physical dynamics.
- Conducted an extensive literature review on autoregressive video generation. Built a training framework on both GPU and NPU platforms using Megatron-LM, ModelLink, and MindSpeed. Based on phyworld dataset, curated and constructed a custom dataset designed to simulate real-world physical dynamics.
- The project paper has been submitted to ICCV 2025.

A Diffusion Timestep Tokenizer for Visual Tasks Aug. 2024 - Present
Supervisors: Dr. Fengda Zhang, Dr. Zhongqi Yue and Prof. Hanwang Zhang

- Explored a timestep-based visual tokenizer that abandons spatial properties and performs excellently on tasks such as image editing and video generation.
- Conducted the development of an inference and evaluation pipeline for tokenizer, participated in the design and implementation of part of ablation experiments, and curated and constructed a large-scale high-resolution dataset for training.

SKILLS/HOBBIES

Language	Mandarin(native), English(IELTS: 6.5, GRE: 330)
Programming Languages	Python, MATLAB, R
Hobbies	table tennis, xiangqi