Guian **Fang**

VIDEO GENERATION · WORLD MODEL

(+65) 84072627 | guianfang@u.nus.edu | enderfga.cn | enderfga | enderfga | Guian Fang

"Think Twice, Code Once."

Education

NUS(National University of Singapore) — *Singapore*

PhD in Electrical and Computer Engineering — *Aug. 2024 - Current*

- Supervisor: **Mike Zheng Shou** (Show Lab)

SYSU(Sun Yat-sen University) — *Shenzhen, China*

B.E. in Intelligent Science and Technology — *Sept. 2020 - Jul. 2024*

- Supervisor: **Xiaodan Liang** (HCP-I2 Lab SYSU)

Skills

Frameworks	PyTorch, TensorFlow, Deepspeed, Accelerate
Programming Languages	Python, C/C++, Matlab, LaTeX, Web (HTML, CSS, JavaScript)
	Chinese (Mandarin, mother tongue), English (IELTS 6.5)
DevTool	Git (version control system), Linux (operating system), Docker (containerization tool), Spack (package management tool), Slurm (cluster management and job scheduling system)

Experience

Character animation — *Mikomiko Pte. Ltd.*

Research Scientist — *Jul. 2024 - Sept. 2025*

Focus on controlable image generation and ID-consistency video animation.

Text to 3D World & Autonomous Agents and Multiagent Systems — *Utopai Studios | Cybever*

Research Scientist — *Sept. 2023 - Jul. 2024*

Remote internship, mostly works on 3D computer vision and Multimodal Large language model.

Evaluation of text-to-video generation. — *NUS | SHOW LAB*

Primary author — *Jun. 2023 - Nov. 2023*

- Towards A Better Metric for Text-to-Video Generation

Moltimodel LLM & Prompt engineering. — *Alibaba | DAMO Academy*

Main participant — *Aug. 2023 - Nov. 2023*

- ChartThinker: A Contextual Chain-of-Thought Approach to Optimized Chart Summarization(**COLING accepted**)

Fine-tuning large language models & Controllable image generation. — *Shanghai AI Laboratory*

Main participant — *Jul. 2023 - Nov. 2023*

- **LLaMA-Adapter(5.9k stars)**, **LLaMA2-Accessory(2.8k stars)**
- Improving Compositional Text-to-image Generation with Large Vision-Language Models
- FocusDiff: Foreground Self-Decoupling Customization for Text-to-Image Generation

Enhancing Generative AI Models for 2D or 3D Using RLHF. — *HUAWEI | NOAH'S ARK LAB*

Primary author — *Mar. 2023 - Nov. 2023*

- RealignDiff: Boosting Text-to-Image Diffusion Model with Coarse-to-fine Semantic Re-alignment(**TNNLS accepted**)
- HumanRefiner: Benchmarking Abnormal Human Generation and Refining with Coarse-to-fine Pose-Reversible Guidance(**ECCV accepted**)

National Undergraduate Innovation and Entrepreneurship Training Program. — *Sun Yat-sen University*

Principal Investigator — *Jun. 2021 - Mar. 2023*

- Online Evidence Enhanced and Fine-Grained Fake News Detection in Long Text During the COVID-19 Pandemic.(**NCAA accepted**)
- Alzheimer's disease diagnosis based on multimodal data including images, voiceprints, poses and biomarkers.(**under Lancet submission**)
- IEEE UV 2022 Paper Publication: Wise in Vaccine Allocation

Honors & Awards

Competition

2021	**1st Place**, "Smart Campus" Hackathon competition	*Guangzhou, China*
2021	**3rd Place**, National College Students Mathematical Modeling Competition (Huashu Cup)	*Tianjin, China*
2022	**Silver Prize**, Intelligent Algorithm Competition of the Game Cup	*Shenzhen, China*
2022	**3rd Place**, 3rd National College Computer Ability Challenge	*Shenzhen, China*
2022	**2nd Place**, Asia and Pacific Mathematical Contest in Modeling	*Beijing, China*
2022	**Silver Medal**, China Collegiate Algorithm Design & Programming Challenge Contest	*Shanghai, China*
2022	**2nd Place**, The 2022 Social Computing Innovation Competition	*Hangzhou, China*
2023	**1st Place**, 4th National College Computer Ability Challenge	*Shenzhen, China*

Scholarship

2021	**Recipient**, Ethic Award scholarship	*Shenzhen, China*
2022	**Recipient**, Huawei Intelligent Foundation Scholarship	*Shenzhen, China*
2022	**Recipient**, National Encouragement scholarship	*Shenzhen, China*
2023	**1st Place**, SYSU Outstanding Student Scholarship (Top **5%**)	*Shenzhen, China*
2023	**1st Place**, National Scholarship (Rank **1st**)	*Shenzhen, China*
2023	**1st Place**, Li Xuerou Foundation Scholarship (Among the top **10** students in the college)	*Shenzhen, China*

Publication

- **Guian Fang**, Wenbiao Yan, Yuanfan Guo, Jianhua Han, Zutao Jiang, Hang Xu, Shengcai Liao, & Xiaodan Liang. (2024). HumanRefiner: Benchmarking Abnormal Human Generation and Refining with Coarse-to-fine Pose-Reversible Guidance. In European conference on computer vision. Cham: Springer Nature Switzerland, 2024
- Mengsha Liu, Daoyuan Chen, Yaliang Li, **Guian Fang**, & Ying Shen. (2024). ChartThinker: A Contextual Chain-of-Thought Approach to Optimized Chart Summarization. In Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024) (pp. 3057-3074).
- **Guian Fang**, Zutao Jiang, Jianhua Han, Guansong Lu, Hang Xu, Shengcai Liao, & Xiaodan Liang. (2023). RealignDiff: Boosting Text-to-Image Diffusion Model with Coarse-to-fine Semantic Re-alignment.
- Ziyang Ma, Mengsha Liu, **Guian Fang**, & Ying Shen. (2023). LTCR: Long-Text Chinese Rumor Detection Dataset.
- Baiqiao Yin, Jiaqing Yuan, Weichen Lv, Jiehui Huang, & **Guian Fang**. (2023). Wise in Vaccine Allocation. In 2022 6th International Conference on Universal Village (UV) (pp. 1-6).

Preprints

- Jay Zhangjie Wu, **Guian Fang**, Haoning Wu, Xintao Wang, Yixiao Ge, Xiaodong Cun, David Junhao Zhang, Jia-Wei Liu, Yuchao Gu, Rui Zhao, Weisi Lin, Wynne Hsu, Ying Shan, & Mike Zheng Shou. (2024). Towards A Better Metric for Text-to-Video Generation.
- Song Wen, **Guian Fang**, Renrui Zhang, Peng Gao, Hao Dong, & Dimitris Metaxas. (2023). Improving Compositional Text-to-image Generation with Large Vision-Language Models.
- **Guian Fang**, Mengsha Liu, Yi Zhong, Zhuolin Zhang, Jiehui Huang, Zhenchao Tang, & Calvin Yu-Chian Chen. (2023). Multimodal Identification of Alzheimer's Disease: A Review.
- Xudong Lu, Renrui Zhang, **Guian Fang**, Fu-Yun Wang, Zhaoyang Huang, Aojun Zhou, Peng Gao, & Hongsheng Li. (2024). FocusDiff: Foreground Self-Decoupling Customization for Text-to-Image Generation.
- **Guian Fang**, Yuchao Gu, & Mike Zheng Shou. (2025). FramePrompt: In-context Controllable Animation with Zero Structural Changes.
- Jay Zhangjie Wu, **Guian Fang**, Dongrong Joe Fu, Vijay Anand Raghava Kanakagiri, Forrest Iandola, Kurt Keutzer, Wynne Hsu, Zhen Dong, & Mike Zheng Shou. (2025). VEditBench: Holistic Benchmark for Text-Guided Video Editing.