Yatian Pang

yatian_pang@u.nus.edu | +65 8941 0019 | Homepage | LinkedIn | 1.2k+ citations

Education

Ph.D. **National University of Singapore** — Aug 2021 - Now
- Advisor: Prof. Tay Eng Hock, Francis
- Working closely with Prof. Li Yuan from Peking University.
- Research interest: Multi-modal understanding and generation.

M.Sc. **National University of Singapore**, College of Design and Engineering — Aug 2020 - June 2021
- GPA: 4.50/5.00

B.Eng. **Soochow University**, College of Mechanical and Electrical Engineering — Sep 2016 - June 2020
- GPA: 3.80/4.00

Experience

Alibaba Qwen team, Research Intern — Hangzhou, China, April 2025 - Now
- Mentor: Dr. Shuai Bai and Dr. Hang Zhang
- Working on Qwen3-VL with focus on video understanding, including long video and streaming video understanding.

Everlyn AI, Researcher — Shenzhen, China, Oct 2024 - March 2025
- Mentor: Prof. Harry Yang and Prof. Sernam Lim
- Developing advanced algorithms for multi-modal generation, including image-to-video generation and Auto-regressive image generation.

A*STAR, Research Engineer & Research Intern — Singapore, Oct 2021 - July 2023
- Developing an abnormal detection method for industrial systems.
- Conducting industrial data cleaning with machine learning algorithms.

Projects

UniWorld, one of the latest unified models following GPT-4o — 2025
- The project aims to design a unified framework for image understanding, generation, and especially editing. We solve a key challenge when connecting frozen VLMs with Diffusion generators that an extra image semantic encoder is necessary for image editing quality.

UniWorld
Technical Report

Open-Sora-Plan, one of the earliest projects trying to reproduce Sora, 12k+ stars — 2024
- The project aims to contribute a large-scale generation model for generating high-resolution videos with long durations based on various user inputs.

Open-Sora-Plan
Technical Report

Publications & Preprints

Next Patch Prediction for Autoregressive Visual Generation. — 2024
Yatian Pang, Peng Jin, Shuo Yang, Bin Lin, Bin Zhu, Zhenyu Tang, Liuhan Chen, Francis EH Tay, Ser-Nam Lim, Harry Yang, Li Yuan
UNDER REVIEW. [Arxiv 2024]

DreamDance: Animating Human Images by Enriching 3D Geometry Cues from 2D Poses. — 2024

Yatian Pang, Bin Zhu, Bin Lin, Mingzhe Zheng, Francis EH Tay, Ser-Nam Lim, Harry Yang, Li Yuan
UNDER REVIEW. [Arxiv 2024]

Envision3D: One Image to 3D with Anchor Views Interpolation. 2024
Yatian Pang, Tanghui Jia, Yujun Shi, Zhenyu Tang, Junwu Zhang, Xinhua Cheng, Xing Zhou, Francis EH Tay, Li Yuan
[Arxiv 2024]

Masked Autoencoders for 3D Point Cloud Self-Supervised Learning. 2023
Yatian Pang, Zhenghua Chen, Li Yuan
Book Chapter in Deep Learning For 3D Vision: Algorithms And Applications. [link]

Masked autoencoders for point cloud self-supervised learning. 2022
Yatian Pang, Wenxiao Wang, Francis EH Tay, Wei Liu, Yonghong Tian, Li Yuan
European Conference on Computer Vision (ECCV) 2022. [link]

VideoGen-of-Thought: A Collaborative Framework for Multi-Shot Video Generation. 2024
Mingzhe Zheng, Yongqi Xu, Haojian Huang, Xuran Ma, Yexin Liu, Wenjie Shu, ***Yatian Pang***, Feilong Tang, Qifeng Chen, Harry Yang, Ser-Nam Lim
UNDER REVIEW. [Arxiv 2024]

Cycle3D: High-quality and Consistent Image-to-3D Generation via Generation-Reconstruction Cycle. 2024
Zhenyu Tang, Junwu Zhang, Xinhua Cheng, Wangbo Yu, Chaoran Feng, ***Yatian Pang***, Bin Lin, Li Yuan
AAAI 2025. [link]

Repaint123: Fast and high-quality one image to 3D generation with progressive controllable 2D repainting. 2024
Junwu Zhang, Zhenyu Tang, ***Yatian Pang***, Xinhua Cheng, Peng Jin, Yida Wei, Wangbo Yu, Munan Ning, Li Yuan
European Conference on Computer Vision (ECCV) 2024. [link]

Moe-llava: Mixture of experts for large vision-language models. 2023
Bin Lin, Zhenyu Tang, Yang Ye, Jiaxi Cui, Bin Zhu, Peng Jin, Jinfa Huang, Junwu Zhang, ***Yatian Pang***, Munan Ning, Li Yuan
UNDER REVIEW. [Arxiv 2024]

Languagebind: Extending video-language pretraining to n-modality by language-based semantic alignment. 2023
Bin Zhu, Bin Lin, Munan Ning, Yang Yan, Jiaxi Cui, HongFa Wang, ***Yatian Pang***, Wenhao Jiang, Junwu Zhang, Zongwei Li, Wancai Zhang, Zhifeng Li, Wei Liu, Li Yuan
International Conference on Learning Representations (ICLR) 2024. [link]

Abnormal Wedge Bond Detection Using Convolutional Autoencoders in Industrial Vision Systems. 2022
Ji-Yan Wu, ***Yatian Pang***, Xiang Li, Wen Feng Lu
International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICCME) 2022. [link]

Skills

Languages: Chinese, English, Python

Deep Learning Technologies: Pytorch, Deepspeed, Megatron