

PENGFEI ZHOU

Phone: +8616601535077 • E-mail: zpf4wp@outlook.com • Homepage: lancezpf.github.io

EDUCATION

Doctor of Philosophy School of Computing National Univeristy of Singapore , Singapore	Advisor: Prof. Yang You Graduating: Dec. 2028 Funded by NUS Research Scholarship
Master of Engineering Institute of Computing Technology Chinese Academy of Sciences , Beijing, China	Advisor: Prof. Shuqiang Jiang and Prof. Weiqing Min Graduated: Jul. 2024 GPA: 3.62/4
Bachelor of Engineering College of Computer Science Zhejiang University of Technology , Hangzhou, China	Advisor: Prof. Cong Bai Graduated: Jun. 2021 Grade: 86.0/100 (<i>Ranking: 2/46</i>)

MAJOR AWARDS & HONORS

Top-tier

- Outstanding Graduate of Beijing (*First Class Honors, 10K/200K+ in Beijing*) May 2024
- Outstanding Graduate of Zhejiang Province (*First Class Honors, 10K/300K+ in Zhejiang*) May 2021
- **National Scholarship** (*3/300+ in VIPL Group and 25/1K+ in ICT, CAS*) Dec. 2023
- **National Scholarship** (*5/1K+ in College of Computer Science and 50/4K+ in ZJUT*) Dec. 2020

International

- Silver Award in Zhejiang International "Internet+" Undergraduate Entrepreneurship Competition Sep. 2021
- Second Place in International Information Retrieval Competition of LMRT in ImageCLEF Lifelog 2019 Sep. 2019

National

- E-Funds Fintech Fellowship by ICT, Chinese Academy of Sciences (*Top Honors, 10/300+ in ICT*) Jan. 2023
- "Jianxing" Honors Student by Zhejiang University of Technology (*Top Honors, 86/4K+ in ZJUT*) Jun. 2021
- First Prize in "HUAWEI Cup" National Undergraduate IoT Design Contest Aug. 2020
- Second Prize in National Service Outsourcing Competition Aug. 2020

PUBLICATIONS

Representative

- **Pengfei Zhou**, Xiaopeng Peng, Jiajun Song, Chuanhao Li, Zhaopan Xu, Yue Yang, Ziyao Guo, Xiaojun Chang, Yu Qiao, Wenqi Shao and Kaipeng Zhang. 2025. *OpenING: A Comprehensive Benchmark for Judging Open-ended Interleaved Image-Text Generation*. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (**Oral**, 0.7% in 13K Submissions)
- **Pengfei Zhou**, Weiqing Min, Chaoran Fu, Ying Jin, Mingyu Huang, Xiangyang Li, Shuhuan Mei and Shuqiang Jiang. 2025. *FoodSky: A Food-oriented Large Language Model that Can Pass the Chef and Dietetic Examination*. Patterns (Cell Press)
- **Pengfei Zhou**, Weiqing Min, Jiajun Song, Yang Zhang, and Shuqiang Jiang. 2023. *Knowledge-Enhanced Features Synthesizer for Real-World Zero-Shot Food Detection*. IEEE Transactions on Image Processing (JCR-Q1 Top-tier Journal)
- **Pengfei Zhou**, Weiqing Min, Yang Zhang, Jiajun Song, Ying Jin, and Shuqiang Jiang. 2023. *SeeDS: Semantic Separable Diffusion Synthesizer for Zero-shot Food Detection*. Proceedings of the ACM International Conference on Multimedia (Top-tier Conference in Multimedia, 29.3% Acceptance Rate)
- **Pengfei Zhou**, Kaining Ying, Zhenhua Wang, Dongyan Guo, and Cong Bai. 2022. *Self-Supervised Enhancement for Named Entity Disambiguation via Multimodal Graph Convolution*. IEEE Transactions on Neural Networks and Learning Systems (JCR-Q1 Top Journal)
- **Pengfei Zhou**, Cong Bai, Jie Xia, and Shengyong Chen. 2022. *CMRDF: A Real-Time Food Alerting System Based on Multimodal Data*. IEEE Internet of Things Journal (JCR-Q1 Top Journal)

Under Review

- **Pengfei Zhou**, Jie Xia, Xiaopeng Peng, Wangbo Zhao, Zilong Ye, Zekai Li, Suorong Yang, Jiadong Pan, Yuanxiang Chen, Ziqiao Wang, Kai Wang, Qian Zheng, Xiaojun Chang, Gang Pan, Shurong Dong, Kaipeng Zhang, Yang You. *Neural-Driven Image Editing*.
- **Pengfei Zhou**, Fanrui Zhang, Xiaopeng Peng, Zhaopan Xu, ..., Wangbo Zhao, Kai Wang, Xiaojun Chang, Wenqi Shao, Yang You, Kaipeng Zhang. *MDK12-Bench: A Multi-Discipline Benchmark for Evaluating Reasoning in Multimodal Large Language Models*.

RESEARCH EXPERIENCE

Research Engineer

Shanghai Artificial Intelligence Laboratory, Shanghai, China

Jul. 2024 – Jan. 2025

- Conducted research on MLLMs for interleaved image-text generation
- Built a benchmark for interleaved image-text generation and published final outcomes in CVPR 2025
- Developed large-scale multimodal training datasets and integrated multimodal chain-of-thought techniques to train a novel interleaved image-text generation MLLM

Technical Director

Institute of Intelligent Computing Technology, CAS, Suzhou, China

Apr. 2022 – Oct. 2023

- Led a team of 8 employees to develop a visual system for dietary management and nutritional estimation
- Proposed food detection algorithm via zero-shot learning techniques with 99%+ accuracy in real-world scenarios
- Deployed dietary systems in Changzhou First People's Hospital, Puxiang Co. LTD and Xujiang Experimental Middle School with 100M+ usage, leading to a 3× checkout rate and 80% manpower requirement reduction
- Authorized a Chinese patent, published related papers in ACM MM 2023 and IEEE TIP

Co-founder, CTO

Hangzhou Xiao Bin Technology Co. LTD, Hangzhou, China

Jan. 2020 – May 2021

- Director of Smart Home System Research and Development Laboratory with 10+ employees
- Developed a smart bin recognition system (*99%+ classification accuracy, 30%+ higher than competitors*)
- Sold products in major cities in China including Shanghai and Hangzhou, with an annual turnover of CNY 100K+
- Won 10+ prizes in competitions including Silver Award in Zhejiang International "Internet+" Undergraduate Entrepreneurship Competition (*2/120*) and First Prize in National Undergraduate IoT Design Contest (*5/100*)

Research Assistant

Institute of Computer Vision, Zhejiang University of Technology, Hangzhou, China

Apr. 2019 – Jun. 2021

- Worked with Prof. Cong Bai and Prof. Shengyong Chen to research multimodal retrieval technique supported by National Natural Science Fund of China Project (*No. 61976192*)
- Led a team of 5 researchers in the Undergraduate Innovation Entrepreneurship Training Program-funded project
- Won the second place in LMRT of ImageCLEFlog2019 Competition (*as the team leader*)
- Authored 3 published international papers, 3 Chinese patents (*granted*) and 2 software copyrights (*granted*)

ACADEMIC ACTIVITIES

Reviewer: IEEE TNNLS/TMM/TCYB, AAAI 2026, NeurIPS 2025/24, ICML 2025, ICLR 2024, ACM MM 2025/24/23

Co-organizer: China Multimedia 2023 Conference (*1K+ attendees*): Website designer and system manager

SKILLS

Language: English (**7.0 IELTS**), Chinese (*Native*)

Self-evaluation: Good at leading team work, active creative thinking, love and skilled in all kinds of writing

RESEARCH INTERESTS

Current: Working on Computer Vision (*MLLMs*) and Multimodal Content Analysis (*Food Computing*)

Interested in: Generative Models (*LLMs, Diffusion*), AI-enabled Healthcare, and Open-world Topics (*World Models*)