

# Yang Zhongyu

+86 13437555149  Email  Homepage  Google Scholar  GitHub

## RESEARCH INTERESTS

**Fields:** Multimodal Learning, MultiAgent, Vision-Language Models, Generative AI

**Topics:** MLLM; MultiAgent; Diffusion-based Generation; Retrieval-Augmented Reasoning

**Objective:** My long-term goal is to build general-purpose multimodal systems that can perceive, reason, and communicate effectively across visual, textual, and behavioral modalities in dynamic, real-world environments.

## EDUCATION

- **Lanzhou University (Project 985)** Sept. 2021 - June.2025  
B.S. in Mathematics(the Basic Theory Class)(Main major) and Management (Minor) Lanzhou, China
  - Relevant courses: Mathematical Analysis, Advanced Algebra, Python, C++, Probability Theory, Ordinary Differential Equations, Numerical Analysis, Microeconomics, Differential Geometry, Functional Analysis, etc.
  - **Thesis:** Adaptive Multi-task Medical Image Restoration, Supervised by [Prof.Yu-Mei Huang](#) and rated as A<sup>+</sup>.

## EXPERIENCE

- **SenseTime Research** Shenzhen, China  
Research Intern, General Perceptual Computing Group Feb. 2025 – Present
  - Research on streaming MLLMs, focusing on hallucination mitigation and consistent vision-text alignment.
- **Heriot-Watt University** Edinburgh, UK  
Remote Research Intern, Advisor: [Prof.Wei Pang](#) and [Dr.Yingfang Yuan](#) March. 2024 – Present
  - Research on diffusion-based image generation and multi-agent systems for task-specialized visual synthesis.
- **KAUST** Saudi Arabia  
Remote Research Intern, Advisor: [Dr.Jun Chen](#) and [Prof.Mohamed Elhoseiny](#) Dec. 2024 – Jan. 2026
  - Exploring multi-agent coordination for web-scale retrieval to augment the reasoning capability of MLLMs.
- **The Chinese University of Hong Kong, Shenzhen** Shenzhen, China  
Research Assistant, Advisor: [Prof.Ruimao Zhang](#) Apr. 2024 – Nov. 2024
  - Developed MLLMs' alignment between visual and textual cues to enable human-centric multimodal generation.
- **iFLYTEK Co., Ltd.** June 2023 - Aug. 2023  
Data Analysis Assistant Intern Lanzhou, China
  - Developed multimodal recommendation models from multi-source behavior-aware cues.

## PAPERS AND PUBLICATIONS(CORE A\* × 6 & CORE A × 1)

C = Conference, J=Journal, T = Tech Report, R = Under Review

\* means Corresponding Author, † means equal contribution

### Multimodal Understanding, Agents & Efficient MLLMs

- [C.1] [Zhongyu Yang](#), Junhao Song, Siyang Song, Wei Pang, Yingfang Yuan. **MERMAID: Multi-perspective Self-reflective Agents with Generative Augmentation for Emotion Recognition**. EMNLP Main 2025 (**Core A**).
- [C.2] [Zhongyu Yang](#)<sup>†</sup>, Yingfang Yuan<sup>†</sup>, Xuanming Jiang, Baoyi An, Wei Pang. **InEx: Hallucination Mitigation via Introspection and Cross-Modal Multi-Agent Collaboration**. AAAI 2026, (Acceptance Rate = 17.6%, **Core A\***).
- [C.3] [Zhongyu Yang](#), Wei Pang, Yingfang Yuan. **X<sup>R</sup>: Cross-Modal Agents for Composed Image Retrieval**. WWW 2026 (Acceptance Rate = 20.1%, **Core A\***).
- [C.4] [Zhongyu Yang](#), Zuhao Yang, Wei Pang, Yingfang Yuan. **SVAgent: Storyline-guided Long Video Understanding via Cross-modal Multi-agent Collaboration**. CVPR 2026 (**Core A\***).
- [J.1] [Zhongyu Yang](#)<sup>†</sup>, Dannong Xu<sup>†</sup>, Wei Pang, Yingfang Yuan. **Script: Graph-Structured and Query-Conditioned Semantic Token Pruning for Multimodal Large Language Models**. Transactions on Machine Learning Research.
- [R.1] [Zhongyu Yang](#), Dannong Xu, Yonghan Zhang, Kefan Chen, Xinyi Wang, Yang Xu, Wei Pang, Yingfang Yuan. **Do Vision and Text Cues Exhibit Evidential Coupling? UFO: A Benchmark for Compositional Multimodal Reasoning in Unified Models**. Under review.

### Multimodal Generation & Creative Applications

- [C.1] [Zhongyu Yang](#)<sup>†</sup>, Jun Chen<sup>†</sup>, Dannong Xu, Junjie Fei, Xiaoqian Shen, Liangbing Zhao, Chun-Mei Feng, Mohamed Elhoseiny. **WikiAutoGen: Towards Multi-Modal Wikipedia-Style Article Generation**. ICCV 2025 (**Core A\***).
- [C.2] [Zhongyu Yang](#), Junhao Song, Zhang Luo, Zuhao Yang, Yang Xu, Jingfen Lan, Yonghan Zhang, Wei Pang, Siyang Song, Yingfang Yuan. **ReChar: Revitalizing Characters with Structure-Preserved and User-Specified Aesthetic Enhancements**. Siggraph Asia 2025 Oral (**Core A\***).

- [C.3] Xuanming Jiang, Baoyi An, Dingyu Nie, Haoyu Ren, Zhengwei Zou, Yizhe Yang, Jialie Shen, Zhiwen Jin, Xueming Qian, **Zhongyu Yang\***, Guoshuai Zhao. **MeteoNet: Next-Generation Efficient Reliable Reconstruction Solution for Full-Scale Non-Uniform Meteorological Fields**. *CVPR 2026 (Core A\*)*.

### Economics, Math & Others

- [J.1] **Zhongyu Yang**, Ziyue Xue. **Analysis and Forecast of GDP of Gansu Province based on ARIMA Model**. *Chinese Market (IF=0.6)*
- [J.2] Mengying Su, **Zhongyu Yang\***, Shujaat Abbas, Yuriy Bilan, Agnieszka Majewska. **Toward Enhancing Environment Quality in OECD Countries: Role of Municipal Waste, Renewable Energy, Environment Innovation and Environmental Policy**. *Renewable energy (SCI Q1 Top, IF=9.0)*.
- [J.3] Zhichao Yu, Wenlan Xie, Junjie Guo, **Zhongyu Yang\***. **Green Effect of Energy Transition Policy: A quasi-natural Experiment Based on New Energy Demonstration Cities**. *Finance Research Letters (SSCI Q1 Top, IF=10.4)*.
- [T.1] **Zhongyu Yang**, Hao Wu. **Tropical Linear Representation of Involute Chinese Monoids**. Technology Report.

### PATENTS

P=Patent, S=Software

- [P.1] **Zhongyu Yang**. A mathematics teaching system based on virtual reality. (CN116312091A)
- [S.1] **Zhongyu Yang**. Green and Low-carbon Integrated Monitoring Software.(2023SR1355487)
- [S.2] **Zhongyu Yang**. Fully automatic spatial sound field environment perception system. (2024SR0538446)

### HONORS AND AWARDS

- **Conference Volunteer**, EMNLP 2025 Top 1%
- **Broadening Participation Award**, ICCV 2025 Top 1%
- **Excellence Graduation thesis**, Lanzhou University (2025) Top 5%
- **Best Wiki Winner**, International Directed Evolution Competition (IDEC) (2024) Top 5%
- **Silver Medal**, International Genetically Engineered Machine Competition (iGEM) (2024) Top 15%
- **Meritorious Winner**, ICM Mathematical Modeling Competition (2023) Top 6%
- **Honorable Mention**, MCM (2023), focus on multimodal data modeling Top 25%
- **Best Hardware & Target Design**, IDEC (2023), integrating ML with real-world application Top 1%
- **1st Prize**, National Data Analysis Competition (2022), applied to visual data regression Top 3%
- **1st Prize**, China Big Data Challenge (2022) Top 8%
- **Outstanding Student Scholarship**, Lanzhou University (2022, 2024) Top 15%

### SKILLS AND SERVICES

- **Programming Languages:** Python, R, Stata, Latex
- **Languages:** Mandarin (Native), Cantonese (Native), English (Fluent)
- **Operation System:** Windows (advanced), Linux (advanced)
- **Reviewer Services:** CVPR, ICCV, ECCV, ICML, ICLR, Siggraph Asia, ARR(ACL, EMNLP, NAACL, TACL), AAAI, ICME; EMFT (Q1), ESPR (Q1), IJER (Q2)