Tianshuo Peng

Tel: +86 152 251 67693   **Email:** pengtianshuo@link.cuhk.edu.hk   **Homepage:** https://pengts.github.io
Address: The Chinese University of Hong Kong, Shatin, NT, Hong Kong SAR, China

EDUCATION

Chinese University of Hong Kong, PhD   Hong Kong, China
Multimedia Lab   *2025-present*
- Advised by Xiangyu Yue

Wuhan University, B.S.   Wuhan, China
Artificial Intelligence, School of Computer Science   *2021-2025*
- **GPA: 3.95**/4.00   **Rank: 1**/51
- Advised by Zuchao Li and Lefei Zhang

RESEARCH INTEREST

- **Agent System & Agentic Model**: Multi Agent Collaboration, Planning and Execution in Complex Scenarios
- **Multi-Modal Understanding and Generating:** Multi-Modal Representation Learning, Visual-Language Generation, Unified Understanding-Generation Model
- **Large Multi-modal Models (LMMs):** Visual-Language Reasoning Model

SCHOLARSHIPS AND AWARDS

- **Outstanding Graduate of Wuhan University** (Top 10% schoolwide)   *May. 2025*
- **Hong Kong PhD Fellowship Scheme**   *Apr. 2025*
- **Lei Jun Breakthrough Scholarship** (5 students schoolwide)   *Apr. 2025*
- **Academic Pioneer of the School of Computer Science** (2 students schoolwide)   *Mar. 2025*
- **China National Scholarship** (Top 2% nationwide)   *Oct. 2024*
- **China National Scholarship** (Top 2% nationwide)   *Oct. 2023*
- **CCF (China Computer Federation) Elite Collegiate Award** (101 students nationwide)   *Aug. 2024*

SELECTED PUBLICATION

Chimera: Improving Generalist Model with Domain-Specific Experts   [ICCV 2025]
- *Tianshuo Peng, Mingsheng Li, Jiakang Yuan, Hongbin Zhou, Renqiu Xia, Renrui Zhang, Lei Bai, Song Mao, Bin Wang, Aojun Zhou, Botian Shi, Tao Chen, Bo Zhang, Xiangyu Yue*

Multi-modal Auto-regressive Modeling via Visual Tokens   [ACM MM2024]
- *Tianshuo Peng, Zuchao Li, Lefei Zhang, Hai Zhao, Ping Wang, Bo Du*

A Novel Energy Based Model Mechanism for Multi-Modal Aspect-Based Sentiment Analysis   [AAAI2024]
- *Tianshuo Peng, Zuchao Li, Ping Wang, Lefei Zhang, and Hai Zhao*

FSUIE: A Novel Fuzzy Span Mechanism for Universal Information Extraction   [ACL2023]
- *Tianshuo Peng, Zuchao Li, Lefei Zhang, Bo Du, Hai Zhao*

MME-Reasoning: A Comprehensive Benchmark for Logical Reasoning in MLLMs   [Preprint]
- *Jiakang Yuan, Tianshuo Peng, Yilei Jiang, Yiting Lu, Renrui Zhang, Kaituo Feng,*
- *Chaoyou Fu, Tao Chen, Lei Bai, Bo Zhang, Xiangyu Yue*

ACADEMIC SERVICE

- **Invited Reviewer:** ICLR, NeurIPS, ACL

RESEARCH EXPERIENCE

InternAgent: Closed-loop Multi-agent Framework for Autonomous Research   Feb. 2025 - Jul. 2025
Shanghai Artificial Intelligence Laboratory   *Advisor: Xiangyu Yue, Bo Zhang*

- We introduce InternAgent, a unified closed-loop multi-agent framework to conduct Autonomous Scientific Research (ASR) across various scientific research fields, enabling researchers to tackle complicated problems in these fields with unprecedented speed and precision.
- InternAgent has demonstrated its versatility across 12 scientific research tasks, capable of generating innovative ideas to enhance the performance of baseline code.

Comprehensive Logical Reasoning Evaluation for LMMs   Feb. 2025 - May. 2025
Shanghai Artificial Intelligence Laboratory   *Advisor: Xiangyu Yue, Bo Zhang*

- According to Charles Sanders Peirce's classification of reasoning, we developed multi-modal reasoning benchmark that includes all three basic types of reasoning: Deductive, Inductive, and Abductive
- Our evaluation results reveal an imbalance in the reasoning capabilities of existing thinking&chat LMMs in different reasoning scenarios. A systematic analysis was conducted on the performance improvement brought by thinking mode and the stimulation of model reasoning ability by existing RL methods.

Improving Generalist Model with Domain-Specific Experts   Aug. 2024 - Nov. 2024
Shanghai Artificial Intelligence Laboratory   *Advisor: Xiangyu Yue, Bo Zhang*

- We introduce a scalable and low-cost multi-modal pipeline, cooperated with a novel Generalist-Specialist Collaboration Masking (GSCM) mechanism, designed to boost the ability of existing LMMs with domain-specific experts. This results in a versatile model that excels across the chart, table, math, and document domains
- Our proposed method achieve SOTA performance on multi-modal reasoning and visual content extraction tasks
- ICCV 2025 poster presentation (**first author**)

Unified Multi-modal Auto-regressive Modeling Framework   Sept. 2023 - Apr. 2024
Wuhan University, Sigma Lab   *Advisor: Zuchao Li*

- We perform multi-modal auto-regressive modeling with a unified objective over multi-modal large language models for the first time.
- Our proposed model achieves superior visual-language understanding performance compared to the 13B model on five VQA datasets and four Benchmark Toolkits, using only 7B parameters. It even demonstrates competitive performance against larger-scale models like 33B.
- MM 2024 poster presentation (**first author**)

Energy Based Model for Multi-Modal Aspect-Based Sentiment Analysis   Mar. 2023 - Aug. 2023
Wuhan University, Sigma Lab   *Advisor: Zuchao Li*

- We proposed a Dual-Query Mechanism using the prompt as both visual query and language query to extract prompt-aware visual information. Additionally, we introduce an Energy-based Pairwise Expert that predicts aspect or sentiment span based on pairwise stability.
- Experiments on three widely used benchmarks demonstrate that our method outperforms previous approaches and achieves a new state-of-the-art performance.
- AAAI 2024 poster presentation (**first author**)

Fuzzy Span Mechanism for Universal Information Extraction   Spet. 2022 - Feb. 2023
Wuhan University, Sigma Lab   *Advisor: Zuchao Li*

- We proposed the Fuzzy Span Loss to alleviate the span-based model's excessive reliance on precise annotations. Additionally, we introduced Fuzzy Span Attention to adjust the model's focus on semantic information within limited spans in information extraction tasks.
- Our proposed method gets SOTA or competitive performance on a series of main IE tasks.
- ACL 2023 poster presentation (**first author**)

OTHERS

- **Languages:** Mandarin(Native), English(IELTS: 7.0)
- **Skills:** Python, Pytorch, Deep Learning, Pattern Recognition, Computer Vision and Natural Language Processing