

Tianshuo Peng

Tel: +86 152 251 67693

Email: pengts@whu.edu.cn

Homepage: <https://pengts.github.io>

Address: No. 299 Bayi Road, Wuchang District, Wuhan City, Hubei Province, China

EDUCATION

Wuhan University B.S.

Major: Artificial Intelligence, School of Computer Science

Wuhan, China

September 2021- present

- **GPA: 3.95/4.00** **Rank: 1/51**
- **Core courses:** Advanced Mathematics, Linear Algebra, Optimality Method, Probability and Mathematical Statistics, Stochastic Processes, Machine Learning, Computer Vision, Intelligent Robots, Intelligent System Design and Implementation

SCHOLARSHIPS AND AWARDS

- **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
- **Lei Jun Computer Research Fund** (Top 0.4% schoolwide) Jun. 2024
- **First-class Scholarship** (Top 5% schoolwide) Oct. 2024
- **First-class Scholarship** (Top 5% schoolwide) Oct. 2023
- **Second-class Scholarship** (Top 10% schoolwide) Oct. 2022
- **Merit Student** (Top 5% schoolwide) Oct. 2024
- **Merit Student** (Top 5% schoolwide) Oct. 2023
- **Merit Student** (Top 5% schoolwide) Oct. 2022
- **Winning Award, “Sinan Cup” Quantum Computing Programming Challenge** Aug. 2023

SELECTED PUBLICATION (*CORRESPONDING AUTHORS, †EQUAL CONTRIBUTION)

- Chimera: Improving Generalist Model with Domain-Specific Experts** [Under Review]
- Tianshuo Peng[†], Mingsheng Li[†], Hongbin Zhou, Renqiu Xia, Renrui Zhang, Lei Bai, Song Mao, Bin Wang, Conghui He, 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

RESEARCH INTEREST

- **Multi-Modal Understanding and Generating:** Multi-Modal Representation Learning, Visual-Language Pretraining, Image Generation, Video Generation
- **Multi-Modal Large Language Models (MLLMs):** Visual Question Answering, Visual-Language Chatbot, Multi-Modal Instruction Following Tuning
- **Natural Language Processing:** Large Language Models, Information extraction, Sentiment Analysis

RESEARCH EXPERIENCE

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
- CVPR 2025 (**under review**)

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**)

PROJECT EXPERIENCE

PDF-Extract-Kit (Open-source Project)

Jul. 2024 - Nov. 2024

- Assist in building powerful document parsing toolkit: Incorporates state-of-the-art models for layout detection, formula detection, formula recognition, OCR, and other core document parsing tasks. (**6k stars**)
- Assist in fine-tuning models with diverse document annotation data to deliver high-quality results across various complex document types. (**51k downloads**)

Computer Vision & Deep Learning Integrated Project (Course Project)

Jun. 2023

- Perform image classification on the CIFAR-10 dataset using the classic LeNet model and further improve the model structure to enhance performance.
- Conduct medical image segmentation on the ISBI-2012 dataset using the Unet model and further enhance the model to improve performance.
- Fine-tune the pre-trained DeeplabV3 model on the PASCAL VOC 2012 dataset to achieve semantic segmentation.

Computer Vision Based Autonomous Driving Project (Course Project)

Jul. 2023

- Learn to use digital image processing techniques to identify and label lane lines in video frames.
- Implement vehicle object detection using MobileNets for low computational cost.

OTHERS

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