Wenyuan Wang

github.com/codemaker-123 Google Scholar
ww462@scarletmail.rutgers.edu

RESEARCH INTERESTS

Multimodal Learning, Computer Vision, Reinforcement learning, Trustworthy AI

EDUCATION

Wuhan University
B.S. in Electronic Information Engineering *Sep 2020 - June 2024*

– GPA: 3.60(87.2)/4.0 (40%)
– Third Class Scholarship of School of Electronic Information 2021

PUBLICATIONS

- **Generalizable Geometric Image Caption Synthesis.**
  Wenyuan Wang*, Yue Xin*, Rui Pan*, BingXu Meng*, Renjie Pi, Tong Zhang
  Submitted to NeurIPS Datasets and Benchmarks Track(5443)
- **Probabilistic Residual User Clustering.**
  Wenyuan Wang, Yusong Zhao, Zihao Xu, Hengyi Wang, Shreya Venugopal, Desmond Lobo, Chengzhi Mao, Qi Xu, Zhigang Hua, Yan Xie, Bo Long, Shuang Yang, Hao Wang.
  IJCAI2025 workshop Causal Learning for Recommendation Systems / Submitted to TMLR
- **Multimodal Needle in a Haystack: Benchmarking Long-Context Capability of Multimodal Large Language Models.**
  Hengyi Wang, Haizhou Shi, Shiwei Tan, Weiyi Qin, **Wenyuan Wang**, Tunyu Zhang, Akshay Nambi, Tanuja Ganu, Hao Wang.
  NAACL 2025 Main
- **Continual Learning of Large Language Models: A Comprehensive Survey.**
  Haizhou Shi, Zihao Xu, Hengyi Wang, Weiyi Qin, **Wenyuan Wang**, Yibin Wang, Zifeng Wang, Sayna Ebrahimi, Hao Wang.
  ACM Computing Surveys
- **Multi-tailed vision transformer for efficient inference.**
  Yunke Wang, Bo Du, **Wenyuan Wang**, Chang Xu.
  Neural Networks, 2024, 174: 106235.

RESEARCH EXPERIENCE

Geo-Image-Textualization *Mar 2025 - May 2025*
Research Assistant, Department of Computer Science, UIUC Advisor: Prof. Tong Zhang

- Proposed Geo-Image-Textualization, a reinforcement learning-based framework for generating semantically aligned geometry image-caption pairs.
- Constructed GeoReasoning-10K, the first dataset with full modality equivalence for geometric reasoning, enhancing MLLMs' cross-modal alignment.
- Demonstrated significant improvements in Qwen-2.5-vl performance across geometry, arithmetic, algebraic, and numeric domains.

Interpretability of MLLM *May 2024 - Sep 2024*
Research Assistant, Department of Computer Science, Rutgers Advisor: Prof. Hao Wang

- Identified strong correlations between visual inputs and token-level outputs in the LLaVA.
- Enhanced the interpretability of LlaVA by uncovering the underlying attention mechanisms.

- Proposed an adaptive pruning technique to selectively prune hierarchical attention layers, preserving high-attention heads to improve both accuracy and efficiency.

Probabilistic Residual User Clustering. *Jul 2024 - Jan 2025*
Research Assistant, Department of Computer Science, Rutgers Advisor: Prof. Hao Wang

- Proposed PRUC, a causal Bayesian framework that clusters users and models residuals between predicted and true ratings to enhance recommendation accuracy.
- Introduced a plug-and-play architecture compatible with diverse deep learning recommenders, improving performance in cold-start and domain-shift settings.
- Demonstrated significant improvement across benchmark datasets while uncovering meaningful user clusters via latent variable inference.

Multimodal Needle Benchmark & The survey for Continual learning of LLM *Mar 2024 - May 2024*
Research Assistant, Department of Computer Science, Rutgers Advisor: Prof. Hao Wang

- Evaluated performance of InstructBLIP vicuna/t5 on custom-developed benchmark.
- Researched advancements in continual learning for multimodal large language models.
- Contributed to comprehensive survey on multimodal LLM continual learning

Dynamic scene graph generation *Jun 2023- Sep 2023*
Research Assistant, Shenzhen University Advisor: Prof. Runhao Zeng, Chuang Gan

- Based on the insights that class, box and relationship triples can help to forecast each other mutually, deployed the Multimodal-LLM following the MiniGPT4, and fine-tuned it on AG datasets to generate scene graph.
- Fine-tuned class, box and add these models in MiniGPT4's base model by transfer learning to enhance relationship predictions.

Multi-Tailed vision Transformers *Feb 2023- Apr 2023*
Research Assistant, Department of Computer Science, University of Sydney Advisor: Prof. Chang Xu

- Designed multiple tails to generate visual sequences of different lengths for the following Transformer encoder
- Employed a tail predictor to determine which tail produces the most accurate prediction for the image.
- Achieved a significant reduction in FLOPs with no accuracy degradation and outperformed other compared methods in both accuracy and FLOPs. Confirmed that our backbone is a general backbone that can be used in downstream tasks, such as object detection

Dynamic Confidence Calibration Based on Adaptive Temperature Adjustment *Dec 2022- Jan 2023*
Research Assistant, Department of Computer Science, University of Sydney Advisor: Prof. Chang Xu

- Proposed a dynamic neural network allocate computing resources of the neural network according to the difficulty of the sample.
- Used adaptive temperature adjustment for network confidence calibration to more accurately verify the difficulty of samples in the dynamic neural network.

WORK EXPERIENCE

Visual alignment in industrial Settings *Aug 2023 - Jan 2024*
Intern, Siemens, Shanghai, China

- Simulated the AGV to detect the cargo in Gazebo based on ROS2.
- Designed a visual algorithm that calculated the pose and position of cargo by determining its frontal area.

SKILLS

- **Programming:** Proficient in Python, C++ and C
- **Development Tools & Platforms:** Experienced with ROS, Linux, and embedded systems