Haochuan Li

haochuan [at] u.nus.edu haochuan-li.github.io

Education

National University of Singapore, Singapore *Aug 2023 – Jan 2025*
M.S. in Computer Science

- GPA: 4.42/5.00
- **Thesis:** Multimodal Generation and Retrieval

Soochow University, Suzhou, Jiangsu, China *Sept 2019 – June 2023*
B.S. in Telecommunication Engineering

- GPA: 3.82/4.0 (Rank: 4/107)

Research Interest

Generative Recommendation, Large Multimodal Model(LMM)

Publications

1. Leigang Qu, **Haochuan Li**, Tan Wang, Wenjie Wang, Yongqi Li, Liqiang Nie, Tat-Seng Chua, **TIGeR**: Unifying **T**ext-to-**I**mage **Ge**neration and **R**etrieval with Large Multimodal Models. *The Thirteenth International Conference on Learning Representations* (**ICLR'25**).
   - Driven by the complementary roles of **text-to-image generation and retrieval** in visual information access, we propose **unifying both tasks** to meet complex human information needs.
   - Inspected the **intrinsic cross-modal discriminative abilities of LMMs** and proposed TIGeR-ONE, **a model-agnostic framework** for the TIGeR task. TIGeR-ONE performs text-to-image generation and retrieval in **a training-free autoregressive manner**, selecting the best-matched result autonomously and efficiently.
   - Constructed a **comprehensive image acquisition benchmark**, TIGeR-Bench, to evaluate the performance of TIGeR on LMMs in **creative and knowledge-intensive domains**. Extensive experiments on TIGeR-Bench and two T2I-G benchmarks including Flickr30K and MS-COCO verify the effectiveness of TIGeR-ONE.
2. Leigang Qu, **Haochuan Li**, Wenjie Wang, Xiang Liu, Juncheng Li, Liqiang Nie, Tat-Seng Chua, Self-Improving Large Multimodal Models for Compositional Text-to-Image Generation. *The IEEE/CVF Conference on Computer Vision and Pattern Recognition* (**CVPR'25**).
   - First to focus on the task of **LMMs' self-improvement for T2I**. Proposed a **model-agnostic self-improvement framework** to enable LMMs to achieve high-quality self-feedback and learning.
   - Introduced a **dropout-based strategy** for continuous LMMs to diversify image representations, along with a **continuous DPO approach**, KC-DPO, to optimize LMMs with preference representation pairs
   - Conducted extensive experiments on three compositional T2I benchmarks, validating the superiority of SILMM, **30% improvements on T2I-CompBench++**.

Experience

Algorithm Engineer *Hangzhou, China*
TAOBAO & TMALL GROUP, Alibaba *Mar 2025 – Current*

- Working on large-scale personalized recommendation for Taobao.

Multimodal Large Language Model Intern *Shanghai, China*
Bytedance *Mar 2024 – Aug 2024*

- Contributed to the **pretraining of TikTok e-commerce Large Multimodal Model(LMM)**, conducting experimental explorations on Large Language Model(LLM) backbones (LLaMA/Mistral), vision models (CLIP/SigLIP), and image resolution strategies (high-resolution ViT, tiled image inputs). Pretrained the model in three stages using both open-source and internal e-commerce multimodal data, **achieving an**

average +3.88pp improvement on indoor e-commerce task benchmarks compared to the previous version.

- Led the **Supervised Fine-Tuning(SFT) and deployment of TikTok e-commerce multimodal product matching model**. Optimized data quality (including noise reduction and constructing Chain-of-Thoughts(CoT) multi-task data) and injected knowledge (rule-based and explicit feature injection). The final model **outperformed the online XGB model in nationwide precision-recall metrics by an average of +5pp, achieving parity with human annotations** and providing high-quality labeled data for multilingual and new market scenarios.
- Developed a **cold-start product sales prediction model** for TikTok e-commerce by leveraging in-platform and external product matching information, pricing power, and other features. Modeled the likelihood of successful sales post cold-start traffic incubation, **replacing rule-based selection with a model-based scoring system**. **Successfully launched in multiple regions**, including the US, UK, and Southeast Asia.

Applied Scientist Intern *Suzhou, China*
Microsoft *May 2022 – Dec 2022*

- Worked on **Fine-grained Ranking models for Windows and Edge homepage news recommendation feed** which provide content services to over 1 billion Windows users.
- Contributed to the development of feature crossing, sequence modeling and multi-task modules in ranking model, which resulted in a **+0.35% gain in online DAU(Daily Active User), +0.83% gain in online CI/UU(Content Interaction Per Unique User), +1.493% Click/UU**.
- Developed **model-level and instance-level explanation summaries and visualization tools** using SHAP, LIME, and counterfactual methods for better understanding of the ranking model's predictions.

Honor and Awards

- **Outstanding Academic Performance Scholarship** *2020-2022*
- **Globalink Research Internship Scholarship, Globalink Mitacs Alumni** *2022*
- **Hui-Chun Chin and Tsung-Dao Lee Chinese Undergraduate Research Endowment(CURE)** *2022*
- **Outstanding Graduate of Soochow University, China** *2023*