Hao Li

⚲ Personal Homepage   ✉ 18th.leolee@gmail.com   ⌂ Github

RESEARCH INTEREST

- My major interest is the multimodal learning, especially for **Vision and Language**, which could be regarded as two main inputs of an agent (*e.g.*, humans or robots). I believe that research on vision and language perception is essential for achieving strong artificial intelligence.
- Current multimodal works generally focus on the performance, and resultant little work explore multimodal theory. I pursue the construction of **multimodal theory**, and now my research mainly focuses on **Diversity**, **Uncertainty**, **Robust Learning**, and **High-dimension Theory**.

EDUCATION

University of Electronic Science and Technology of China
Master, computer science   *Sep 2021 – Present*
Co-advised by Prof. Jingkuan Song, Prof. Lianli Gao, and Prof. Heng Tao Shen

Northeast Forest University
B.S., computer science (GPA: 3.82/4.00)   *Sep 2017 – Jul 2021*

AWARDS AND COMPETITIONS

- RoboMaster University Championship 2018   (**Regional Champion**)
- RoboMaster University Technical Challenge 2018   (**Global Third Place**)
- RoboMaster University Championship 2019   (**Global Third Prize**)
- RoboMaster University Technical Challenge 2019   (**Global Second Prize**)
- China Undergraduate Mathematical Contest in Modeling 2020   (**National Second Prize**)
- National Artificial Intelligence Innovation and Application Competition 2022 (**National First Prize**)

PUBLICATIONS

- A Differentiable Semantic Metric Approximation in Probabilistic Embedding for Cross-Modal Retrieval. *NeurIPS2022*. **Hao Li**, Jingkuan Song, Lianli Gao, Pengpeng Zeng, Haonan Zhang, Gongfu Li. pdf

RESEARCH EXPERIENCE

Center for future media (CFM)   *Sep 2021 – Present*
- The Research on Diversity of Cross-modal Retrieval
- Traditional multimodal benchmarks are hard to support training and evaluating the semantic diversity of multimodal data, so I propose a potential correlation excavating method. Then a reasonable metric and a metric optimization approach are designed to assess and boost the diversity of multimodal model, respectively. The paper has been published in NeurIPS2022.

- The Research on Uncertainty of Cross-modal Retrieval
- The predictions of cross-modal models are often unreliable due to the aleatoric uncertainty and the epistemic uncertainty. The former is caused by the inherent noise and ambiguity in data, while the latter is inherent to the model. I give a clear definition of the inherent aleatoric uncertainty in multi-modal data. Then, a prototype-based uncertainty quantification approach is proposed to provide trustworthy results. The paper is planned to be submitted to NeurIPS2023.