

侣庆一 (QINGYI SI)

(+86) 131-6168-5288 · siqingyi@iie.ac.cn · Page phoebussi.github.io · GitHub @PhoebusSi

教育背景

中国科学院大学-信息工程研究所, 网络空间安全, 博士研究生 2019.9 - 2024.6
博士国家奖学金 (前 2%), 朱李月华优秀博士 (前 1%), 所长奖 (前 2%), 优秀干部, 三好学生, 师从于林政研究员和王伟平研究员

北京语言大学, 计算机科学与技术, 工学学士 2015.9 - 2019.6
排名 1/50, 国家奖学金 (前 2%), 北京市普通高等学校优秀毕业生, 北京市优秀毕业设计 (论文)(前 0.7%)

研究兴趣

多模态 (language & vision); 大语言模型 (LLMs); 视觉问答 (VQA); 分布外鲁棒性 (OOD robustness)

学术论文 (一作论文 11 篇, 9 篇已录用, * 表示共同一作)

- [1] **Combo of Thinking and Observing for Outside-Knowledge VQA** *ACL'23*
• 【大模型 & 多模态】我们将跨模态编码器拓展进 LLM, 使之兼顾视觉特征和 LLM 的隐含知识, 大幅度提高开放域视觉问答能力。与此同时, 我们用 wikipedia 进行检索增强以提高模型的知识输入。
- [2] **Compressing And Debiasing Vision-Language Pre-Trained Models for VQA** *EMNLP'23*
• 【大模型 & 多模态】我们探究了大型多模态预训练模型的压缩和分布外鲁棒性的联系和双赢策略。
- [3] **An Empirical Study of Instruction-tuning Large Language Models in Chinese** *EMNLP'23*
• 【大模型】我们分别探讨了一系列的 LLM 底座、参数效率方法和不同的指令微调数据在中文指令微调中的影响。此外, 我们还实验探究了思维链数据, prompt 语种, 中文词表及人类价值取向等重要元素的影响。
- [4] **Language Prior Is Not the Only Shortcut: A Benchmark for Shortcut Learning in VQA** *EMNLP'22*
• 【多模态】我们对 VQA 鲁棒性评测基准拨乱反正, 受到了 VQA 领域先驱学者 [Damien Teney](#) 的高度赞扬和认可。
- [5] **Towards Robust VQA: Making the Most of Biased Samples via Contrastive Learning** *EMNLP'22*
• 【多模态】我们打破了 VQA 模型在独立同分布 (IID) 和 OOD 场景下的 trade-off, 同时兼顾 IID 和 OOD 性能。
- [6] **Check It Again: Progressive Visual Question Answering via Visual Entailment** *ACL'21*
- [7] **Learning Class-Transductive Intent Representations for Zero-shot Intent Detection** *IJCAI'21*
- [8]* **Object Attribute Matters in Visual Question Answering** *AAAI'24*
- [9]* **Schema Item Matters in Knowledge Base Question Answering** *IJCNN'23*
——Under Review Papers——
- [10]* **Towards One-to-many Visual Question Answering** *CVPR'24*
- [11]* **Towards Unified Interactive Visual Grounding in The Wild** *ICRA'24*
• 【大模型 & 多模态】我们提出一种兼顾视觉定位和多轮对话的多模态 LLM, 可通过与人类多轮交互来消除歧义和识别意图, 并准确圈出用户的目标对象 (visual grounding)。该工作有着丰富的应用场景, 如家用机器人。
——Co-author Papers——
- [11] **Cross-modality Multiple Relations Learning for Knowledge-Based Visual Question Answering** *TOMM'23*
- [12] **Visual Dialog for Spotting the Differences between Pairs of Similar Images** *MM'22*
- [13] **A Multi-channel Neural Network for Imbalanced Emotion Recognition** *ICTAI'19*

研究项目 (多个 SM 项目的主要负责人, 不再展示)

Alpaca-CoT (<https://github.com/PhoebusSi/alpaca-CoT/>) 2.1k stars
• 【大模型】Owner, 发起人和主要开发者。我们搭建了方便研究人员上手的大模型轻量级指令微调平台, 将各类指令微调数据、十数种 LLM 和一系列参数效率方法 (e.g., LoRA、P-tuning) 的接口实现统一。此外, 我们还率先将 CoT 数据扩展到 LLaMA 中以提高其推理能力。我们构建了目前最大的指令微调数据集 (月下载量 2.9k)。目前 huggingface 网站中有 50+ 相关大模型是借助我们平台实现的, 这些模型月下载量达 40k+。我们基于该平台对大模型在中文场景下的微调也做出了透彻的实验分析。曾受 [机器之心](#) 的邀请直播分享了该项目。

Table-based Fact Verification with LLMs 论文撰写中
• 【大模型】首个实现表格事实验证的 LLM, 构造了“复杂问题拆解-子问题答案生成-事实验证”的 pipeline。

实习经历

微软亚研 | MSRA, NLC 组, 科研实习生, 负责翻译大模型 MoE、数据蒸馏的实现。【大模型】 2021.10-2022.5

腾讯微信 | WeChat AI, 模式识别中心, 科研实习生, 负责 VQA 鲁棒性研究。【多模态】 2021.3-2021.10