Huiyi Chen

Github | Homepage | Google Scholar | hche0278@student.monash.edu

Education

Southeast University 09/2023 - present
Mphil. in Artificial Intelligence. Supervisor: Xu Yang
WAM: 86.71/100. GPA:3.82/4.0

Monash University 09/2023 - present
M.S. in Information Technology Systems.
WAM: 83.63/100 (High Distinction). GPA:3.875/4.0

Nanchang Hangkong University 09/2019 - 06/2023
B.S. in Network Engineering.
WAM: 90.08/100. Rank: 1/71. GPA:4.01/5.0

Research Interests

In-context learning (ICL), Large Vision-Language Model (LVLM), Multimodal Retrieval-Augmented Generation (MRAG).

Publications and Preprints

* Refers to the authors having the equal contribution, and should be considered as co-first authors.

1. Li Li*, Jiawei Peng*, **Huiyi Chen***, Chongyang Gao, Xu Yang. How to Configure Good In-Context Sequence for Visual Question Answering (CVPR 2024), Citations: 36.

   The paper explores how to configure effective in-context sequences for Visual Question Answering (VQA) tasks to enhance the in-context learning (ICL) capabilities of Large Visual Language Models (LVLM). The paper elaborates on the role of different in-context configuration in LVLM and designs new configuration methods, providing valuable insights for optimizing LVLM's ICL performance in VQA tasks.

2. **Huiyi Chen**, Jiawei Peng, Kaihua Tang, Xin Geng, Xu Yang. Enhancing Multimodal In-Context Learning for Image Classification through Coreset Optimization (Preprinted, submitted to ACM MM 04/2025).

   The paper proposes a novel coreset construction framework, KeCO, for image classification tasks, aimed at enhancing the in-context learning capabilities of Large Vision-Language Models. KeCO leverages untapped data from the support set to aggregate category-relevant information into the coreset via feature-level updates. Notably, KeCO achieves strong performance in a simulated online scenario, demonstrating its practical applicability.

Work in progress

1. **LMM-R1-v2**

   This project is a continuation of our lab's earlier work, LMM-R1: Empowering 3B LMMs with Strong Reasoning Abilities, which has garnered significant attention with 751 stars on GitHub. I currently actively engaged in the development and improvement of the second version.

Selected Project Experience

1. **PixTag : An AWS-powered Serverless Image Storage System with Advanced Tagging Capabilities** (Project Link)

   This project aims to build an AWS-based online system that allows users to store and retrieve images based on auto-generated and manual tags. The focus of this project is to design a serverless application that enables clients to upload their images to public cloud storage. Upon image upload, the application automatically tags the image with the objects detected in it, such as person, car, etc. Later on, users can also modify and add new tags to the images. They can also query images based on the objects present in them.

2. **BookRecord: An Android App dedicated to tracking reading for those who cherish paper-based books.** (Project Link)

   This project aims to design for readers of physical books, offering features to track reading progress, log notes, and analyze reading habits. Users can register accounts, import books, update reading status, and view their reading statistics in a comprehensive analysis page. The app leverages Google Firebase for data storage and a public REST API for book information retrieval.

AWARDS & Others

Scholarships

Southeast University Second-Class Academic Scholarship (Top 10%), China *Oct. 2023*

Nanchang Hangkong University Outstanding Student Award (Top 1%), China *2020 - 2023*

Nanchang Hangkong University First-class Academic Scholarship (Top 1%), China *2020 - 2023*

Competitions

Blue Bridge Cup National Software and Information Technology Professional Talent Programming Competition — 2021, Provincial Second Prize

China College Student Computer Design Competition — 2022, Provincial Second Prize

Language

TOEFL Score: 97 (R)27 (L)27 (S)23 (W)20 — Feb. 2025

Others

1. Since 2020, I have been documenting my learning journey on my CSDN blog, covering topics such as algorithms and deep learning. To date, I have published 290 posts, which have collectively garnered over 361,000 views and attracted 917 followers.
2. Since my undergraduate studies, I have consistently practiced algorithm problems on LeetCode website, practicing my coding skills.