Yanqi Dai

yanqidai@ruc.edu.cn | (+86) 13315036605 | github.com/YanqiDai

Beijing, China | 2000-09

Personal Statement

I am a PhD candidate in the Gaoling School of Artificial Intelligence at Renmin University of China, supervised by Prof. Zhiwu Lu. My research interests lie in **large multimodal models, multimodal role-playing agents, and multi-task learning**. Through my research, my goal is to enhance the multi-task capabilities of large multimodal models and apply multimodal role-playing in mental health.

Education

Renmin University of China, PhD in Artificial intelligence — Sep 2022 – Jun 2027

- GPA: 3.82 / 4.0

Dalian University of Technology, BS in Software Engineering — Sep 2018 – Jun 2022

- GPA: 4.26 / 5.0 (top 1% overall)

Research Experience

(ICLR 2025) MMRole: A Comprehensive Framework for Developing and Evaluating Multimodal Role-Playing Agents – First Author — Mar 2024 – Jul 2024

- Thesis: https://arxiv.org/abs/2408.04203
- We introduce the concept of Multimodal Role-Playing Agents (MRPAs), and propose a comprehensive framework, MMRole, for their development and evaluation. This framework comprises a personalized multimodal dataset and a robust evaluation approach.
- Moreover, we develop the first specialized MRPA, MMRole-Agent.
- Extensive evaluation results show the improved performance of MMRole-Agent and highlight the primary challenges in developing MRPAs, emphasizing the need for enhanced multimodal understanding and role-playing consistency.

CoTBal: Comprehensive Task Balancing for Multi-Task Visual Instruction Tuning – First Author — Oct 2023 – Feb 2024

- Thesis: https://arxiv.org/abs/2403.04343
- We introduce a novel Comprehensive Task Balancing (CoTBal) algorithm for multi-task visual instruction tuning, firstly exploring multi-task optimization in this domain.
- Specifically, we consider two crucial dimensions: Inter-Task Contribution and Intra-Task Difficulty. By quantifying these with performance-based metrics, comprehensive task balancing is thus enabled by assigning more weights to tasks that offer substantial contributions to others, receive minimal contributions from others, and have great intra-task difficulties.
- Extensive experiments show that CoTBal leads to superior and more balanced overall performance in multi-task visual instruction tuning.

(UAI 2023) Improvable Gap Balancing for Multi-Task Learning – First Author — Oct 2022 – Feb 2023

- Thesis: https://arxiv.org/abs/2307.15429
- We propose two novel improvable gap balancing (IGB) algorithms for multi-task learning (MTL): one takes a simple heuristic, and the other (for the first time) deploys deep reinforcement learning for MTL.
- Particularly, instead of directly balancing the losses in MTL, both algorithms choose to dynamically assign task weights for improvable gap balancing, where the improvable gap per task is defined as the distance between the current training progress and desired final training progress.
- Extensive experiments show that our IGB algorithms lead to the best results in MTL via loss balancing and achieve further improvements when combined with gradient balancing.

Internship Experience

Algorithm Intern, Metabrain AGI – Beijing Jun 2023 – Present

- Technical Report: "Awaker2.5-VL: Stably Scaling MLLMs with Parameter-Efficient Mixture of Experts", Co-First Author.
- As a core technical member of the start-up team, I have been deeply involved in developing the ChatImg series (Mar 2023 - Mar 2024) and the Awaker series (Apr 2024 - present) LMMs. Notably, the Awaker series employs the LoRA MoE architecture and has the ability to update independently. Our work has garnered continuous attention and interviews from CCTV.
- Main Responsibilities:
  1. Collection and processing of Chinese pre-training and fine-tuning datasets.
  2. Structural design and code implementation of the ChatImg1.0 model.
  3. Performance evaluation and API deployment of the ChatImg series models.
  4. Structural design, code implementation, model training, and API deployment of the Awaker1.0 model.
  5. Performance evaluation of the Awaker2.0-2.5 models.
  6. Open-source release of the weights and code of the Awaker2.5 model.

Honors & Awards

- Scientific Research Fund, Renmin University of China 2023
- Excellent Scholarship for Postgraduate Study, Renmin University of China 2023
- First-Class Academic Scholarship for Postgraduate Study, Renmin University of China 2023
- Outstanding Graduate of Liaoning Province 2022
- Third Prize of Imagine Cup China Microsoft 2021
- Toly Bread Alumni Scholarship, Dalian University of Technology (awarded to 30 students per year) 2019, 2020
- First Prize of National Mathematics Competition for College Students 2019

Skill List

- Languages: English (CET-6)
- Programming Languages: Python, C, C++, Java
- Certifications: Microsoft Certified Azure AI Fundamentals
- Other Skills: PyTorch Deep Learning Framework, Prompt Engineering, LMM Pre-Training and Fine-Tuning