Qian Yang

qian.yang@mila.quebec Google Scholar Homepage GitHub

EDUCATION

Mila - Quebec AI Institute & Université de Montréal 2023.09 – Present
Ph.D. in Computer Science

- Research topics: Multi-modal Large Language Models, Efficient Training
- Supervisor: Prof. Aishwarya Agrawal
- CGPA: 4.3/4.3
- Anticipated Graduation Date: 07/2027

Harbin Institute of Technology, Shenzhen 2020.09 – 2023.03
MSc in Computer Science and Technology

- Research topics: Multi-modal Learning, Explainable Question Answering
- Supervisor: Prof. Baotian Hu
- Thesis: Fine-grained Alignment for Explainable Multi-modal Inference

University of Electronic Science and Technology of China 2016.09 – 2020.06
BEng in Computer Science and Technology

- CGPA: 3.73/4.0 (top 10%)
- Thesis: Event Extraction based Text Summarization

PUBLICATIONS

- Le Zhang, **Qian Yang**, Aishwarya Agrawal. Assessing and Learning Alignment of Unimodal Vision and Language Models. *CVPR 2025* (***Highlight***).

  – Evaluates unimodal alignment and introduces SAIL, an efficient framework that matches CLIP's zero-shot performance using only 6% paired data and 5 hours of single-GPU training.

- **Qian Yang**, Weixiang Yan, Aishwarya Agrawal. Enhancing Multi-Agent Multi-Modal Collaboration with Fine-Grained Reward Modeling. *NeurIPS 2024 Workshop on AFM.*

  – Uses Process-Reward Modeling (PRM) for zero-shot multi-agent task decomposition, yielding +5.5% gain without extra tuning.

- Yuchen Tian, Weixiang Yan, **Qian Yang**, ..., Dawn Song. CodeHalu: Investigating Code Hallucinations in LLMs via Execution-based Verification. In *Proceedings of The 39th Annual AAAI Conference on Artificial Intelligence*, (AAAI) 2025.

  –Introduces a 4-tier taxonomy for code hallucinations and the CodeHalu benchmark (8.8k samples); evaluates 17 LLMs via dynamic execution-based detection.

- **Qian Yang**, Weixiang Yan, Aishwarya Agrawal. Decompose and Compare Consistency: Measuring VLMs' Answer Reliability via Task-Decomposition Consistency Comparison. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, (EMNLP) 2024.

  – Measure VLM reliability by comparing direct answers with decomposed sub-question reasoning; achieves superior accuracy correlation across 6 tasks without overconfidence bias.

- Le Zhang, Yihong Wu, **Qian Yang**, Jianyun Nie. Exploring the Best Practices of Query Expansion with Large Language Models. *Findings of the Association for Computational Linguistics: EMNLP*, 2024.

- **Qian Yang**, Qian Chen, Wen Wang, Baotian Hu, Min Zhang. Enhancing Multi-modal and Multi-hop Question Answering via Structured Knowledge and Unified Retrieval-Generation. In *Proceedings of the 31st ACM International Conference on Multimedia*, pages 5223-5234, (ACM MM) 2023.
- **Qian Yang**, Yunxin Li, ..., Min Zhang. Chunk-aware Alignment and Lexical Constraint for Visual Entailment with Natural Language Explanations. In *Proceedings of the 30th ACM International Conference on Multimedia*, pages 3587-3597, (ACM MM) 2022.
- Yunxin Li, **Qian Yang**, Qingcai Chen, ..., Lin Ma. Fast and Robust Online Handwritten Chinese Character Recognition with Deep Spatial & Contextual Information Fusion Network. *IEEE Transactions on Multimedia*, vol. 25, pp. 2140-2152, 2022.

Preprints

- **Qian Yang***, Shivam Chandhok*, ..., Leonid Sigal, Aishwarya Agrawal. Learning What Matters: Prioritized Concept Learning via Relative Error-driven Sample Selection (Under Review)

  – Introduces an efficient VLM instruction-tuning framework using dynamic sample selection based on skill-learning progress. Achieves SOTA results without upfront annotations, significantly reducing data and compute costs.

ACADEMIC INTERNSHIPS

Alibaba DAMO Academy, Hangzhou, China *Advisor: Dr. Wen Wang, Dr. Qian Chen*
Enhancing Multi-modal Multi-hop QA with Structured Knowledge *2022.05 – 2022.10*

- Designed an entity-centered fusion model to align cross-modal information using structured knowledge for facilitating connections between different modalities, along with a unified retrieval-generation method to integrate intermediate retrieval results for answer generation; the paper is published in *ACM Multimedia 2023*.

AWARDS AND SCHOLARSHIPS

Women in Machine Learning, MILA (7,000 CAD)	2025
Student Success Scholarship, University of Montreal (5,000 CAD)	2025
Professor Cho Diversity Scholarship, MILA (1,500 CAD)	2025
DIRO Excellence Scholarship, University of Montreal (5,000 CAD)	2024 – 2025
National Encouragement Scholarship (Top 10%, 5,000 RMB)	2019
The First Prize Scholarship, UESTC (Top 20%, 1,000 RMB)	2016 – 2020

PROFESSIONAL ACTIVITIES

WORKSHOP ORGANIZER

- Vision Language Models For All: Building Geo-Diverse and Culturally Aware Vision-Language Models, CVPR 2025 workshop

CONFERENCE REVIEWER

- Outstanding Reviewer of CVPR 2025
- Reviewer of CVPR 2024, ECCV 2024, AAAI 2024, ACM Multimedia 2023, 2024, COLING 2022

TEACHING ROLES

- IFT 6135 - Representation Learning (Autumn 2024), University of Montreal
- Mathematical Logic (Spring 2021), Harbin Institute of Technology, Shenzhen
- Algorithms (Autumn 2020), Harbin Institute of Technology, Shenzhen

TECHNICAL SKILLS

- Programming Languages: Python, C/C++, MATLAB, SQL
- Deep Learning Frameworks: PyTorch, TensorFlow
- Natural Languages: Mandarin (native), English (TOEFL: 99/120, R:26, L:25, S:22, W:26)