Yuchang Su

D.O.B: 12/17/2002

☐ +86-152-1012-3517 • ✉ suyc21@mails.tsinghua.edu.cn

🌐 suyccc.github.io • ⌨ suyccc

Education

Tsinghua University — **Beijing, China**
Bachelor of Engineering in Computer Science and Technology — *Sept. 2021 – Jul. 2025 (expected)*
Overall GPA: 3.91/4.00

Stanford University — **Medical AI and Computer Vision Lab**
Undergraduate Visiting Research Intern — *Jun. 2024 – Sept. 2024*

Standardized Exams

TOEFL: 110 (R29, L30, S23, W28)

GRE: 334 (V164, Q170, AW3.5)

Research Interests

Multimodal Learning, Computer Vision, AI4Biology, AI4Health

Publications

2024: **Yuchang Su**, Yuhui Zhang, Yiming Liu, Serena Yeung-Levy, *Converting Open-ended Questions to Multiple-choice Questions Simplifies Biomedical Vision-Language Model Evaluation*. **Machine Learning for Health 2024**. 🔗

2025: Yuhui Zhang*, **Yuchang Su***, Yiming Liu, Xiaohan Wang, James Burgess, Elaine Sui, Chenyu Wang, Josiah Aklilu, Alejandro Lozano, Anjiang Wei, Ludwig Schmidt, Serena Yeung-Levy, *Automated Generation of Challenging Multiple-Choice Questions for Vision Language Model Evaluation*. **CVPR 2025 (Under Review)**. 🔗

2025: James Burgess, Jeffrey J Nirschl, Laura Bravo-Sánchez, Alejandro Lozano, Sanket Rajan Gupte, Jesus G. Galaz-Montoya, Yuhui Zhang, **Yuchang Su**, . . . , Emma Lundberg, Serena Yeung-Levy, *MicroVQA: A Multimodal Reasoning Benchmark for Microscopy-Based Scientific Research*. **CVPR 2025 (Under Review)**. 📄

2025: **Yuchang Su**, Renping Zhou, Siyu Huang, Xingjiang Li, Tianyang Wang, Ziyue Wang, Min Xu, *Multimodal Generalized Category Discovery*. **AAAI 2025 (Under Review)**. arXiv:2409.1624. 🔗

2024: Yuhui Zhang, Alyssa Unell, Xiaohan Wang, Dhruba Ghosh, **Yuchang Su**, Ludwig Schmidt, Serena Yeung-Levy, *Why are Visually-Grounded Language Models Bad at Image Classification?* **NeurIPS 2024**. 🔗

Research Experiences

Stanford University — **Advisor: Serena Yeung-Levy**
MARVL Lab — *Apr. 2024 – Present*

- ○ **Project 1: Analysis of VLM Classification Ability**
  - Developed and tested fine-tuning strategies for CLIP, including linear probing and full fine-tuning, to analyze performance on image classification tasks.
  - Conducted systematic evaluations on multiple datasets, revealing key insights into how tuning methods affect model generalization.
  - Co-authored a paper, accepted by NeurIPS 2024.

- ○ **Project 2: Automated Generation of Challenging Multiple-Choice Questions for VLM Evaluation**
  - Identified challenges in evaluating Vision-Language Models (VLMs) with open-ended questions, such as scoring inconsistencies and difficulty capturing semantic equivalence.
  - Designed AutoConverter, a multi-agent system (Generator, Reviewer, Refiner) to convert open-ended visual questions into multiple-choice formats with high difficulty and quality.
  - Built VMCBench, a dataset of 10,000 multiple-choice questions from widely used datasets, and evaluated 20 VLMs, setting a new benchmark for standardized evaluation.
  - Wrote a paper as the co-first author, submitted to CVPR 2025.

- ○ **Project 3: Simplifying Biomedical Vision-Language Model Evaluation**
  - Analyzing flaws of Medical open-ended VQAs, focusing on inconsistencies in medical reasoning and limitations of traditional metrics like BLEU in capturing medical semantics.
  - Enhanced the AutoConverter system for medical use and converted 3 popular medical VQA datasets into multi-choice format.
  - Conducted experiments on 18 models, showing high performance correlation between multiple-choice and open-ended evaluations, providing a more consistent assessment framework.
  - Wrote a paper as the first author, accepted by ML4H 2024.

- ○ **Project 4: Cell-level Phenotype & Molecular Retrieval**
  - Predicted cellular phenotypic features and corresponding molecular perturbations from cell painting images, bridging molecular and cellular data.
  - Utilized SigLIP loss for contrastive learning, aligning molecular structures with cellular features in a unified latent space.
  - Incorporated unsupervised data to improve generalization across diverse biological scenarios, aiding in robust therapeutic design.

Carnegie Mellon University **Advisor: Min Xu**
Xu Lab *Apr. 2023 – Sep. 2024*

- ○ **MultiModal Generalized Category Discovery**
  - Extended Generalized Category Discovery (GCD) to a multimodal context, enabling classification of both known and novel categories across diverse data sources.
  - Contributed to developing MM-GCD, a framework aligning feature and output spaces through contrastive learning and distillation, achieving state-of-the-art results on UPMC-Food101 and N24News datasets.
  - Showed through experiments that proper modality alignment significantly reduces decision boundary variance and improves accuracy, while misalignment caused an 18.8% performance drop.
  - Wrote a paper as the first author, submitted to AAAI 2025.

Technical Skills

Programming: C++, Python, Rust, Typescript, System Verilog, MySQL

Tools: PyTorch, NumPy, Scikit-learn

Honors and Activities

2024: Comprehensive Excellence Scholarship, Tsinghua University

2023: Academic Excellence Scholarship, Tsinghua University