Tony Xia

Stanford University, Stanford, California
424.465.4996 | summertony717@gmail.com
TonyXia2001.github.io | linkedin.com/in/summertony717

Keywords: Computer Vision & Graphics, Algorithm Design

Education

Stanford University   Sept 2023—Jun 2025

- **M.S.** in Computer Science with a Specialization in Visual Computing

University of California, Los Angeles   Sept 2019—Jun 2023

- **B.S.** in Computer Science & Applied Mathematics, Magna Cum Laude (GPA: 3.9/4.0)

Publications

MathVista| ICLR 2024   Apr—Oct 2023

- Strategically filtered and pruned problems from **31 Visual Question Answering (VQA) datasets**, ensuring consistency, fairness, and ease-of-use.
- Constructed a **diverse multi-modal benchmark** for large foundation model evaluation.
- Conducted rigorous experiments on **11** prominent **foundation models** such as **Bard**, **GPT4V**.
- Facilitated cutting edge foundation model research through providing a much-needed dataset.

ScienceQA | NeurIPS 2023   Jan—Sept 2022

- Compiled ScienceQA, **the first large-scale VQA dataset** that covers diverse science questions.
- Conducted an evaluation of state-of-the-art **multi-modal neural systems** on the dataset, using Chain-of-Thoughts prompting, resulting in a notable **4% improvement** in performance.
- Devised an efficient algorithm utilizing hash functions to eliminate duplicate entries from the raw image data, **ensuring data integrity** and **enhancing dataset quality**.
- Dataset used by leading research teams at **Google**, **Amazon** and many other companies in the evaluation and development of **novel large language models**.

IconQA | NeurIPS 2022   Jan—Sept 2021

- Collected and curated IconQA, a unique machine learning dataset featuring **over 600k icon images and 100k QA pairs**, utilizing selenium and pandas.
- Conducted an extensive human performance study on Amazon Mechanical Turk over **5 days**.

Relevant Experiences

Student Researcher | Stanford Vision Lab   Oct 2023—Present

- Exploring different research topics for the next year of research.
- Studying **3D visual concept understanding and generation** using both neural and physical approaches.

Officer | ACM ICPC at UCLA   Mar 2020—May 2023

- Organized the **nation's biggest student-run coding contest**, CodeSprintLA, drawing participation from a diverse array of **over 500 contestants** worldwide each year.
- Developed a problem testing tool utilizing React, hosted on AWS EC2, enabling **seamless remote problem testing** for fellow organizers.
- Organized and conducted **over 50 workshops** about data structures and algorithms, preparing them for coding contests as well as technical interviews.

Student Researcher | Prof. Bolei Zhou   Jun 2022—Jan 2023

- Studied pretrained **diffusion probabilistic model**'s behavior under guidance
- Developed methods that help improve feature disentanglement in **guided generation**
- Explored options to allow for more granular control on the image editing task with pretrained diffusion models

Technical Skills

- Python, C, C++, Java, JS, HTML, CSS, Swift, R
- Git, Docker, Linux CLI
- Node, React, Express, MongoDB, Django
- Octave, MATLAB, TensorFlow, PyTorch