Siqiao Huang

HomePage | Github | E-mail | Blog

Education

- **IIIS (Yao Class), Tsinghua University** *2023 - 2027 (expected)*
  B.S. in Computer Science; Beijing, China
  - GPA: 3.92/4.00; Rank: 12/93
  - Selected Courses: Natural Language Processing(A+), Algebra and Computation(A+), Fundamentals of Programming(A+), Basic Principles of Marxism(A+).

Publications

* EQUAL CONTRIBUTIONS, † CORRESPONDING AUTHOR

[1] Bohan Lyu*, **Siqiao Huang***, Zichen Liang*.
SURGE: On the Potential of Large Language Models as General-Purpose Surrogate Code Executors. *Submitted to KDD Datasets & Benchmarks*, 2025.

[2] Shaofeng Yin*, Jialong Wu*, **Siqiao Huang**, Xingjian Su, Xu He, Jianye Hao, and Mingsheng Long†.
Trajectory World Models for Heterogeneous Environments. *Submitted to ICML*, 2025.

Research Experience

- **Billiardbot: Real-World Billiard through VLM Planning and World Model Prediction** *Feb 2025 -*
  Advisor: Prof. Huazhe Xu | Tsinghua University *Aiming for NeurIPS 25*
  - Try to tackle the problem of long-horizon planning with embodied agents
  - Built a realistic physics simulator for the game of billiard, as well as evolving it to a benchmark for dynamics-model prediction and long-horizon planning.
  - Combine the world knowledge embedded in VLMs with domain-specific physics from learned world models to obtain human-level billiard playing with embodied agents.

- **Grounding Video Diffusion Models to Foundation World Models** *Feb 2025 -*
  Advisor: Prof. Mingsheng Long | Tsinghua University *Aiming for NeurIPS 25*
  - Try to answer the question: Can we utilize the pretrained VDMs to build Foundation World Models?
  - While Video Diffusion Models offer high fidelity, it builds on inter-token connections across whole sequence, limiting it's application in predictions where causality plays a huge role.
  - Propose a novel structure to transform pretrained VDMs to action-conditioned auto-regressive World Models.

- **SURGE: LLMs as General-Purpose Surrogate Code Executors** *Feb 2025*
  Advisor: Self-Advised | Tsinghua University
  - Try to answer the question: Can current LLMs serve as General-Purpose Surrogate Code Executors?
  - Curated a **holistic benchmark** to and evaluated multiple open- source and proprietary LLMs' performance
  - Analyze the behavior of LLMs as surrogate models to provide empirical insight.

- **Trajectory World Models for Heterogeneous Environments** *Jul 2024 - Feb 2025*
  Advisor: Prof. Mingsheng Long | Tsinghua University
  - Try to answer the question: Can we effectively transfer knowledge across **different morphologies** in physical interaction modeling to tackle the out-of-distribution challenges in offline reinforcement learning?
  - Pre-train on **data with distinct properties**: Exploratory, Experience replay and Expert Demostration.
  - Demonstrates the **dynamics transfer benefits** in some state-based control environments.

Honors and Awards

- **Comprehensive Excellence Award** *Nov 2024*
  Tsinghua University, University Scholarship
- **Outstanding Sports Scholarship** *Nov 2024*
  Tsinghua University, University Scholarship

Selected Projects

- **A Survey on k-means Clustering Algorithms: Theoretical Analysis & Performance Comparison** *Jan 2025*
  Mostly Theoretical, Tools: Python, Pytorch [🌐] [📄]
  - Elucidated the computational complexity and convergence properties of K-means clustering algorithms and its variants.

- **DreamFactory : Grounding Language Models to World Models** *Nov 2024- Jan 2025*
  Tools: Python, Pytorch [🐙] [📄]
  - Investigated the feasibility of utilizing language models as text-based world models.
  - Proposed a novel architecture to address the self-refutation issue of LLMs and testified it's effectiveness through empirical studies.

- **ManiGen: Generative Simulation Pipeline with Maniskill2** *Oct 2024- Dec 2024*
  Tools: Python, Pytorch, XML [🌐] [🐙] [📄]
  - Developed a generative simulation pipeline using ManiSkill to automate task creation.
  - Utilizes the power of LLMs to propose tasks, generate scenes, and produce task-specific code for rewards, parameters, and metrics.

- **Course Sharing Platform** *Jul 2024*
  Tools: React, Scala, PostgreSQL, HTML, CSS, JavaScript [🐙] [📹]
  - Designed and implemented a PostgreSQL-based course sharing platform using Scala for backend and React for frontend
  - Utilized Stable Diffusion 2 and Llama 2 API to enhance users experiences

- **CAD Escape Game** *Dec 2023- Apr 2024*
  Tools: C#, Unity Engine [🐙] [📄]
  - Developed a 2D Stickman vs CAD-themed game using Unity.
  - Won 2nd prize in Software Design Contest of Tsinghua Univerity (2024).

Skills

- **Programming Languages:** Python, C/C++, C#, Scala, React, PostGreSQL, Swift, Unity Engine.
- **Professional Software:** Pytorch, JAX.
- **Language:** TOEFL: 117/120 (On first trial, Speaking: 30/30). CET-4: 688/710, CET-6: 685/710.

Misc

- **Hobbies:** Basketball, Singing, Piano and Chinese Flute.
- **Groups:** I am a member of the IIIS basketball team and a member of Tsinghua University Chorus.
- In high school, I was quite into Physics & Chemistry, and participated in Olympiad in Physics and Olympiad in Chemistry.