

# Sirui Tao

San Diego, CA, USA | Home Page: [dylantao.github.io](https://dylantao.github.io) | [s1tao@ucsd.edu](mailto:s1tao@ucsd.edu)

## EDUCATION

### University of California, San Diego

2023 - 2025	M.S. Computer Science and Engineering - HCI & Graphics	(Advised by Prof. Steven Dow and Prof. Tzu-Mao Li)
2019 - 2023	B.S. Data Science & B.S. Probability and Statistics - minor in Economics	(Advised by Prof. Judith Fan)

## PUBLICATION

2025	*Wang, Z., *Wang, C., Yoshino, T., <b>Tao, S.</b> , Fu Z., **Li, T.-M. (2025). HotSpot: Screened Poisson Equation for Sign Distance Function Optimization. In submission to Conference on Computer Vision and Pattern Recognition 2025.
2024	* <b>Tao, S.</b> , Liang, I., Peng, C., Wang, Z., **Palani, S., and **Dow, S. (2024). DesignWeaver: Dimensional Scaffolding for Text-to-Image Product Design. In Submission to Conference on Human Factors in Computing Systems 2024.
2023	*Tung, H.-Y., *Ding, M., *Chen, Z., <b>Tao, S.</b> , Lad, V., Bear, D., Gan, C., Tenenbaum, J., Yamins, D., Fan, J. and Smith, K. (2023). Physion++: Evaluating Physical Scene Understanding with Objects Consisting of Different Physical Attributes in Humans and Machines. In Proceedings of the Annual Meeting of the Cognitive Science Society 2023.
2021	*Bear, D., *Wang, E., *Mrowca, D., *Binder, F., Tung, H.-Y., RT, P., Holdaway, C., <b>Tao, S.</b> , Smith, K., Sun, F.-Y., Li, F.-F., Kanwisher, N., Tenenbaum, J., **Yamins, D., and **Fan, J. (2021). Physion: Evaluating physical prediction from vision in humans and machines. In Advances in Neural Information Processing Systems (NeurIPS Datasets & Benchmarks Track) 2021.

## ACADEMIC EXPERIENCE

2024 - Present	<b>UCSD, Protolab - Graduate Researcher</b>	La Jolla, San Diego, CA, USA	<ul style="list-style-type: none"><li>Investigate how to improve generative models with contextual multimodal documents to aid designers and explore how they could help in scenarios of human collaborations.</li></ul>
2023 - Present	<b>UCSD, Tzu-mao Li's Lab - Graduate Researcher</b>	La Jolla, San Diego, CA, USA	<ul style="list-style-type: none"><li>Combining neural methods and traditional graphics pipelines to enable flexible and fast-forward physical interaction prediction; studying inverse physical prediction using diffusion-based models.</li></ul>
2022 - 2023	<b>UCSD, HDSI - Undergraduate Researcher</b>	La Jolla, San Diego, CA, USA	<ul style="list-style-type: none"><li>Develop and implement GraphHSCN, a novel graph neural network architecture capable of capturing long-range interaction information in graph-structured data, exhibiting similar or superior performance compared to established models on selected datasets, with notably increased efficiency. Present the senior capstone and HDSI scholarship poster session - <a href="#">GraphHSCN</a>.</li></ul>
2021 - 2022	<b>UCSD, CogTools Lab - Undergraduate Researcher</b>	La Jolla, San Diego, CA, USA	<ul style="list-style-type: none"><li>Contributed to the development and analysis of Physion under Prof. Judith Fan, a NeurIPS 2021 published research project studying intuitive physics. Focused on data analysis and designing 3D stimuli for non-rigid physical interactions to evaluate the physical understanding of vision models compared to human perception—<a href="#">Physion</a>, <a href="#">Physion++</a>.</li></ul>

## WORKING EXPERIENCE

---

- 2023 **SchedGo - Data Engineer** Remote (Based at La Jolla, San Diego, CA, USA)
- Build the data infrastructure for product analytics and ML algorithm using Firestore;
  - Conduct marketing data analytics to identify key user groups for future promotion efforts;
  - Design a survey model for the front end to gather more direct user feedback.
- 2022 - 2023 **San Diego Supercomputer Center - Deep Learning Engineer Intern** La Jolla, San Diego, CA, USA
- Use Ray Tune to test scalable multi-GPU hyperparameter tuning on HPC infrastructure;
  - Investigate data-centric MLOps solutions to make experimentation more time and resource-efficient.
- 2022 **Tesla - Data Scientist Intern** Fremont, California, USA
- Investigate lossless image compression solutions, saving ~20% of annual image storage;
  - Build and deploy end-to-end process-agnostic advanced statistical process control and quality disposition tools with automatic alert systems; continuously monitor & improve based on user feedback from the process engineering team;
  - Conduct data analyses on various datasets to identify root causes & validate model performance.
- 2021 **Bühler Group - Data Scientist Intern** Wuxi, Jiangsu, China
- Research and initiate the effort to build an Image-Labeling MLOps feature in our B2B product prototype to optimize the model adaptability for defects detection;
  - Prototype an MVP using Figma & Angular within a cross-functional innovation team;
  - Showcase the proposal to the division and get praise from the Director of Innovation;
  - Refine and participate in reviewing new analytics features on our IoT platforms to help our customers with their digital twin adaptation journey with better business insight extraction with descriptive & predictive modeling.
- 2021 **J.D. Power - Analyst Intern** Shanghai, China
- Optimize workflow of the quality analytics team by refining the analytics processes via automating data processing and report pipelines, saving more than 200 working hours/ year for the team;
  - Provide detailed documentation on Git to ensure continuous usage for teammates from non-technical backgrounds.
- 2020 **Bosch - Data Analyst Intern** Wuxi, Jiangsu, China
- Provide effective data analysis and visualization on Tableau for continuous monitoring of data and parameter tuning to gain valuable analytic insights based on the Bosch Industry 4.0 transformation roadmap;
  - Cut down manual inspection and analysis time by providing crucial insight on key bottlenecks and reducing critical bottlenecks diagnosing time by 80% in manufacturing lines with insights on potential optimal parameters;
  - Present dashboards to the MAS (Manufacturing Analytics Solution) Director at the monthly briefing.

## PROJECT EXPERIENCE

---

- 2023 **Tree-of-thoughts Context-aware Encoding for LLMs** La Jolla, San Diego, CA, USA
- Lead the development of a context-aware encoding algorithm for Large Language Models (LLMs), improving knowledge retrieval accuracy and contextual understanding without requiring model fine-tuning.
  - Develop a hierarchical "context tree" structure to compress entire documents into a single prompt, enhancing LLMs' response relevance and creativity to user queries.

## TEACHING EXPERIENCE

---

2024 Summer	<b>UCSD CSE 175 - Teaching Assistant</b>	La Jolla, San Diego, CA, USA
2024 Spring, Fall	<b>UCSD DSGN 1 - Teaching Assistant</b>	La Jolla, San Diego, CA, USA
2024 Winter	<b>UCSD COGS 9 - Teaching Assistant</b>	La Jolla, San Diego, CA, USA

## OTHER SKILLS

---

Software	Python, R, SQL, Angular, D3.js, PyTorch, Clickhouse, FireStore, DevOps, Agile
Data Science	Scalable Analytics, Data Visualization, MLOps, Recommender System, Data Mining, Database
Mathematics	Probabilistic Reasoning, Descriptive and Inferential Statistics, Stochastic Process
AI/ ML	LSTM, GAN, GNN, NeRF, 3D Gaussian Splatting, Reinforcement Learning, LLM

## HONORS & AWARDS

---

2023 - 2024	IGE Shah Fellowship, UCSD
20 Spr, 22 Win	TRELS Research Scholarship, AEP, UCSD
2021 - 2022	Data Science Student Representative, HDSI, UCSD
2021, 2022	HDSI Research Scholarship, HDSI, UCSD
2019 - 2023	Provost Honor, Warren College, UCSD