

KANCHETI SAI SRINIVAS

[personal webpage](#) [google scholar](#)

| cs21rescho1004@iith.ac.in

| (+91)9003107318

RESEARCH INTERESTS

I mainly work on trustworthy machine learning, where the goal is to make ML systems interpretable, fair and robust. My works include

- Visual Reasoning in Multimodal Large Language Models [ongoing]
- Enhancing generalizability of Large Vision-Language models via prompt-tuning [TMLR'25 x2]
- Incorporating external knowledge such as causal priors, or end-user preferences into machine learning systems [TMLR'25, ICML'22]

Keywords: Visual Reasoning, Multimodal LLMs, Vision-Language Models, Algorithmic Recourse, Causality in ML, Trustworthy ML

EDUCATION

Indian Institute of Technology Hyderabad
PhD, Dept. of Computer Science and Engineering
PMRF Scholar December 2020 cycle
Advisor: [Dr. Vineeth N Balasubramanian](#)

Hyderabad, India
Jan 2021 - Present
GPA: 9.47

Indian Institute of Technology Madras
Dual-Degree (BTech & MTech), Dept. of Computer Science and Engineering
Minor in Operations Research

Chennai, India
July 2013 - July 2019
GPA: 8.91

PUBLICATIONS

[2025] **Visual Spatial Reasoning through RL: Successes, Failures, and Lessons** – [K. Sai Srinivas*](#), Aditya S Kanade*, Rohit Sinha, Tanuja ganu, Vineeth N Balasubramanian
Under review at an A* conference [preprint](#)

[2025] **Do SOTA Multimodal Reasoning Models Exhibit Generalized Spatial Intelligence?** – [K. Sai Srinivas*](#), Aditya S Kanade*, Tanuja ganu, Vineeth N Balasubramanian
under review at an A* conference

[2025] **Conformal Prediction in the Age of Multimodal Foundation Models: A Survey** – [K. Sai Srinivas*](#), Christian Stavan*, Kunal Tilaganji*, Sai Mathura Krishnan*, Vineeth N Balasubramanian
Book chapter submission to Alex Gammerman Festschrift

[2025] **Efficient Vocabulary-Free Fine-Grained Visual Recognition in the Age of Multimodal LLMs** – Hari Chandana K, [K. Sai Srinivas](#), Gowtham Reddy A, Vineeth N Balasubramanian
Transactions on Machine Learning Research (TMLR) 2025 [openreview link](#)

[2025] **HARE: Human-in-the-Loop Algorithmic Recourse** – [K. Sai Srinivas](#), Rahul Vigneswaran, Bamdev Mishra, Vineeth N Balasubramanian
Transactions on Machine Learning Research (TMLR) 2025 [openreview link](#)

[2025] **Semantic Alignment for Prompt-Tuning in Vision Language Models** – [K. Sai Srinivas*](#), Hari Chandana K*, Gowtham Reddy A, Vineeth N Balasubramanian
Transactions on Machine Learning Research (TMLR) January 2025 [openreview link](#)

[2024] **Interpretable Model Drift Detection** – Pranoy Panda, [K. Sai Srinivas](#), Vineeth N Balasubramanian, Gaurav Sinha
CODS-COMAD '24: Proceedings of the 7th Joint International Conference on Data Science & Management of Data [paper](#)

[2023] **A Simple Test-time Adaptation Method for Source-free Domain Generalization** – Vaasudev Narayanan, [K. Sai Srinivas](#), Sriranjani Ramakrishnan, Vineeth N. Balasubramanian

ICLR 2023 Workshop: What do we need for successful domain generalization

[link](#)

[2023] **Algorithmic Recourse based on User's Feature-order Preference** – Manan Singh, [K. Sai Srinivas](#), Shivam Gupta, Ganesh Ghalme, Shweta Jain, Narayanan C. Krishnan

Extended Abstract, Young Researcher's Symposium, CODS-COMAD 2023

[paper](#)

[2022] **Matching Learned Causal Effects of Neural Networks with Domain Priors** – [K. Sai Srinivas](#)^{*}, A. Gowtham Reddy^{*}, V. Balasubramanian, Amit Sharma

Proceedings of International Conference on Machine Learning (ICML'22), July 2022; CORE A* conference in ML [paper](#)

^{*} indicates equal authorship

[2021] **Instance-wise Causal Feature Selection for Model Interpretation** – Pranoy Panda, [K. Sai Srinivas](#), V. Balasubramanian

Causality in Vision Workshop @CVPR2021, June 2021

[paper](#)

[2021] **Beyond VQA: Generating multi-word Answers and Rationale to Visual Questions** – [K. Sai Srinivas](#)^{*}, Radhika Dua^{*}, V. Balasubramanian

4th Multimodal Learning and Applications Workshop @CVPR2021, June 2021

[paper](#) [webpage](#)

ONGOING PROJECTS

Visual Interleaving to Enhance Reasoning Capabilities of Multimodal Large Language Models (MLMs) – work done at Microsoft Research India, Bangalore

Chain-of-thought reasoning has enabled MLMs to solve problems in mathematics and general visual reasoning, but falls short for visual spatial tasks. We study various paradigms of CoT reasoning, beginning with text-only reasoning chains within a unified RLVR framework. We then study if interleaving visual tokens can overcome the limitations observed in text-only CoT.

AWARDS & MISCELLANEOUS

- **Visiting Researcher at Microsoft Research India, Bangalore from July 2025**
- Awarded the Microsoft India Research PhD Award 2024 [\[link\]](#)
- Attended CODS-COMAD Jan'2024 at IIIT Bangalore to present our work
- Attended Google Research Week 2023
- Attended ACML 2023 in Bangalore
- Awarded the Prime Minister's Research Fellowship (PMRF) 2021
- Achieved a rank of 304 out of 120,000 attendees in JEE Advanced 2013

SERVICES

- Teaching Assistant for Deep Learning for Vision (Fall 2019, Spring 2021), Deep Learning (Spring 2022), Foundations of Machine Learning (Fall 2022), Advanced Topics In Machine Learning (Spring 2024)
- Reviewer for TMLR, CVPR 2025 (top reviewer), AISTATS 2025, AAAI (2024,2025), NeurIPS 2024, ECCV 2024, and subreviewer for ICML 2024,2023, IJCAI 2023, ICLR 2023, ICLR 2022, AISTATS 2022
- Taught a course titled 'Programming Foundations of Machine Learning' at BVRITH, Spring 2023, Fall 2023
- Volunteered for NeurIPS 2020, ICML 202 virtual conferences

RELEVANT COURSEWORK

- | | |
|---|---|
| • Advanced Topics in Machine Learning | • Probabilistic Models for Machine Learning |
| • Causal Inference and Learning | • Stochastic Processes in Machine Learning |
| • Advanced Data-structures and Algorithms | • Computer Vision, IITM |
| • Convex Optimization – Theory | • Artificial Neural Networks, IITM |

Libraries Pytorch, NumPy, SciPy, Git, Pandas, Matlab

OBJECTIVE

My research efforts are focused on developing algorithms and frameworks to ensure ML models deployed in practice are reliable, interpretable and trustworthy. My current interests are in developing visual reasoning capabilities of Multimodal Large Language Models to enable interfacing with end-users to perform complex visual tasks. As a researcher, I am capable of working individually or in a team. My objective is to work in an environment that will provide opportunities for learning and growth.