Shantanu Jaiswal

+1-412-657-3655 | shantanu12jswl@gmail.com | shantanuj.github.io | Google scholar page link | Github page link

Education

Carnegie Mellon University (CMU) — Pittsburgh, USA
Master of Science in Robotics (Research); Advisor: Deepak Pathak — *Aug 2024 – Present*
- **GPA**: 4.17/4.0

Nanyang Technological University (NTU Singapore) — Singapore
Bachelor of Engineering in Computer Engineering, Specialization in Intelligent Systems — *Aug 2014 – Jun 2018*
- **GPA**: 4.64/5.00 (Honours highest distinction)

Research Interests

Compositional and Long-Context Reasoning; Multimodal Generative Modeling; Embodied and Agentic AI

Publications

Conference papers and preprints

[1] **Shantanu Jaiswal**, Mihir Prabhudesai, Nikash Bhardwaj, Zheyang Qin, Amir Zadeh, Chuan Li, Katerina Fragkiadaki, and Deepak Pathak. "Iterative Refinement Improves Compositional Image Generation". In: ***Under review*** *at CVPR 2026*. [Preprint] [Code (soon)].

[2] **Shantanu Jaiswal**, Debaditya Roy, Basura Fernando, and Cheston Tan. "Learning to Reason Iteratively and Parallelly for Complex Visual Reasoning Scenarios". In: *Conference on Neural Information Processing Systems (NeurIPS) 2024*. [Paper] [Code].

[3] **Shantanu Jaiswal**, Basura Fernando, and Cheston Tan. "TDAM: Top-Down Attention Module for Contextually-Guided Feature Selection in CNNs". In: *European Conference on Computer Vision (ECCV) 2022*. [Paper] [Code].

[4] Ishaan Rawal, Alexander Matyasko, **Shantanu Jaiswal**, Basura Fernando, and Cheston Tan. "Dissecting Multimodality in VideoQA Transformer Models by Impairing Modality Fusion". In: *International Conference on Machine Learning (ICML) 2024*. [Paper] [Code].

[5] Haoran Huan, Mihir Prabhudesai, Mengning Wu, **Shantanu Jaiswal**, and Deepak Pathak. "Can LLMs Lie? Investigation Beyond Hallucination". In: *arXiv preprint arXiv:2509.03518* (2025). [Paper] [Code].

[6] **Shantanu Jaiswal**, Dongkyu Choi, and Basura Fernando. "What do CNNs gain by imitating the visual development of primate infants?" In: *31st British Machine Vision Conference (BMVC) 2020*. [Paper] [Code] [Abstract (Cogsci 2020)].

Other workshop/symposium/short papers and technical reports

[7] Aishik Nagar, **Shantanu Jaiswal**, and Cheston Tan. "Dissecting Zero-Shot Visual Reasoning Capabilities in Vision and Language Models". In: *The Second Tiny Papers Track at ICLR 2024 (Notable)*. [Paper].

[8] Cheston Tan and **Shantanu Jaiswal**. "The Path to AGI Goes through Embodiment". In: *Proceedings of the AAAI Symposium Series. Vol. 1. No. 1. 2023*. [Paper].

[9] Ishaan Rawal, **Shantanu Jaiswal**, Basura Fernando, and Cheston Tan. "Are VideoQA models truly multimodal?" In: *NeurIPS 2023 XAI in Action Workshop*. [Paper].

[10] **Shantanu Jaiswal**, Liu Yan, Dongkyu Choi, and Kenneth Kwok. "A Probabilistic-Logic based Commonsense Representation Framework for Modelling Inferences with Multiple Antecedents and Varying Likelihoods". In: *arXiv 2022*. [Tech. Report].

[11] Annamalai Narayanan, Mahinthan Chandramohan, Rajasekar Venkatesan, Lihui Chen, Yang Liu, and **Shantanu Jaiswal**. "graph2vec: Learning Distributed Representations of Graphs". In: *Proceedings of the 13th International Workshop on Mining and Learning with Graphs (MLG) 2017*. [Paper] [Code].

Research Experience

Graduate Research Assistant – CMU Robotics Institute Nov 2024 – Present
Advisor: Prof. Deepak Pathak *Pittsburgh, USA*

- Investigating reinforcement-learning methods to improve coding and reasoning abilities for long-context tasks.
- Co-designed a inference-time iterative refinement method to improve compositional image generation abilities of state-of-art text-to-image models (paper: [1])
- Contributed to mechanistic interpretability techniques analyzing lying behavior and deception capabilities of LLMs (paper: [4]).

Senior Research Engineer I – A*STAR Center for Frontier AI Research Jan 2019 – July 2024
Advisors: Dr. Cheston Tan and Dr. Basura Fernando; Topic: Cognitively-inspired computer vision *Singapore*

- Investigated design of more effective architectural formulations and training approaches for scene understanding and reasoning tasks (related papers: [2], [3], [6]).
- Contributed to probing techniques and benchmarking studies for vision-language models to systematically examine their reasoning capabilities and analyze potential multimodal biases (related papers: [5], [7], [9]).
- Led development of a probabilistic commonsense knowledge representation framework (tech report: [10]) as part of an industrial grant (from 2019 to 2022) on automated knowledge extraction from aerospace documents.

Undergraduate Research Assistant – NTU School of EEE Sep 2016 – Aug 2017
Advisor: Prof. Lihui Chen; Topic: Deep learning for graph representation learning *Singapore*

- Implemented deep learning approaches for graph representation learning and aspect-based sentiment analysis.
- Contributed to development of the *graph2vec* framework and evaluation of baseline graph representation learning approaches (related paper: [11]).

Scholarships, awards and honours

Graduate Research Funding (full tuition coverage and stipend), CMU Robotics Institute	**2025**
NeurIPS Scholar Award (accomodation and registration grant)	**2024**
President's Research Scholar, Nanyang Technological University	**2017**
Ideasinc Collab4Good Seed Fund (declined), Nanyang Technopreneurship Center	**2016**
Most Innovative Prize, NTU Hackathon on Digital Economy and Services	**2016**
Ministry of Education (MOE) Tuition Grant (merit-based; international student category), Govt. of Singapore	**2014**
Scholarship for Higher Education (SHE-INSPIRE) for meritorious academic performance (top 1%ile overall; perfect scores in Math and Computer Science, Indian School Certificate Board Exams), Govt. of India	**2014**

Selected coursework

CMU: Computer Vision (A+); Visual Learning and Recognition (A); Robot Learning (A+); Math Fund for Robotics (A)
NTU: Machine Learning (A+); Information Retrieval (A+); Neural Networks (A); Living with Mathematics (A+); Math in Real World Applications (A); Engineering Mathematics (A+); Discrete Mathematics (A); Data Structures (A+)

Technical Skills and Languages

Programming languages: Python, C, Java, Matlab, Prolog
Select frameworks: PyTorch, TensorFlow, Needle, ProbLog, Scikit-learn, CoreNLP, Networkx, Pandas, Git, MTurk
Languages: English (native), Hindi

Organizations and Activities

Reviewer for NeurIPS (Main and Datasets track) 2023, 2024; ICLR 2024-26; CVPR 2024-26; ICML 2024	-
Hall Soccer Team Member (trained towards professional soccer in 2nd university year), NTU	Aug 2015 – Dec 2017
Press and Publicity Director, NTU Astronomical Society	Aug 2015 – May 2016
High-school Computer Science Teacher, Shri Ram School Aravali	June 2014 - July 2014
Varsity Soccer (U-17 ASISC North West Regional 2011 winning team), Shri Ram School	Aug 2009 – July 2013
Indian team (3 members) at Pacific Astronomy Summit, Hawaii Space Grant Consortium	Jan 2013 – July 2013