Gabriel Sarch

Pittsburgh, USA
gsarch@andrew.cmu.edu | www.gabesarch.me

Education

Carnegie Mellon University — **Pittsburgh, USA**
Ph.D. in Machine Learning and Neural Computation (Joint Program) — *2020 – Present*
Advisors: Drs. Katerina Fragkiadaki and Michael Tarr

Carnegie Mellon University — **Pittsburgh, USA**
Master of Science in Machine Learning Research — *2023*

University of Rochester — **Rochester, USA**
Bachelor of Science in Biomedical Engineering, magna cum laude — *2016 – 2020*

Selected Grants & Awards

2024: **NeurIPS Spotlight Paper**

○ Awarded to select papers at the NeurIPS conference (top 3%).

2020 – Present: **National Science Foundation Graduate Research Fellowship (NSF-GRFP)**

○ Funding three years of graduate research in machine learning and neuroscience (top 15%).

2023: **Winner of the Embodied AI Workshop Rearrangement Challenge at CVPR 2023**

○ Winning agent submission to move objects in a room to restore them to a given initial configuration.

2022: **Runner-Up of the Amazon Alexa Prize SimBot Embodied Dialogue Challenge**

○ Developed multimodal instruction-following agent as a member of CMU Symbiote Team.

2022: **2020 Biomedical Engineering Award for Academic Excellence (Highest GPA)**

○ Awarded to the highest GPA in the University of Rochester Biomedical Engineering graduating class.

2020: **Tau Beta Pi Engineering Honor Society**

○ Awarded to undergraduate students in the top eighth of their engineering class.

2019: **University of Rochester Center for Visual Science (CVS) Research Fellowship**

○ Fully funded summer of research at University of Rochester.

2018: **National Institutes of Health Ruth L. Kirschstein National Research Service Award**

○ Fully funded summer of research at the National Institutes of Health.

Peer-Reviewed Publications

Sarch, G., Kumaravel, B., Ravi, S., Vineet, V., & Wilson, A. (In Submission). *Multimodal Interactive Contextualized Real World Task Assistance from a Single Demonstration*.

Sarch, G., Jang, L., Tarr, M., Cohen, W., Marino, K., & Fragkiadaki, K. (2024). *VLM Agents Generate Their Own Memories: Distilling Experience into Embodied Programs of Thought*. 38th Advances in Neural Information Processing Systems (NeurIPS). Spotlight.

Sarch, G., Somani, S., Kapoor, R., Jain, A., Tarr, M., & Fragkiadaki, K. (2024). *HELPER-X: A Unified Instructable Embodied Agent to Tackle Four Interactive Vision-Language Domains with Memory-Augmented Language Models*. International Conference on Learning Representations (ICLR) LLM Agents Workshop.

Jain, A., Katara, P., Gkanatsios, N., Harley, A., **Sarch, G.**, Aggarwal, K., Chaudhary, V., Fragkiadaki, K. (2024). *Towards Unified 2D-3D Visual Scene Understanding Foundation Models*. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Spotlight.

Sarch, G., Wu, Y., Tarr, M., & Fragkiadaki, K. (2023). *Open-Ended Instructable Embodied Agents with Memory-Augmented Large Language Models*. Findings of the Association for Computational Linguistics: EMNLP.

Sarch, G., Tarr, M., Wehbe, L., & Fragkiadaki, K. (2023). *Brain Dissection: fMRI-trained Networks Reveal Spatial Selectivity in the Processing of Natural Images*. 37th Advances in Neural Information Processing Systems (NeurIPS).

Sarch, G., Tung, H., Wang, A., Prince, J., Tarr, M. (2023). *3D View Prediction Models of the Dorsal Visual Stream*. Conference on Cognitive Computational Neuroscience (CCN).

Yates, J., Coop, S., **Sarch, G.**, Wu, R., Butts, D., Rucci, M., Mitchell, J. (2023). *Beyond Fixation: Detailed Characterization of Neural Selectivity in Free-Viewing Primates*. Nature Communications.

Sarch, G., Fang, Z., Harley, A. W., Schydlo, P., Tarr, M., Gupta, S., & Fragkiadaki, K. (2022). *TIDEE: Tidying Up Novel Rooms using Visuo-Semantic Common Sense Priors*. European Conference on Computer Vision (ECCV).

Sarch, G.*, Fang, Z.*, Jain, A.*, Harley, A. W., & Fragkiadaki, K. (2021). *Move to See Better: Self-Improving Embodied Object Detection*. British Machine Vision Conference (BMVC). (*Equal contribution)

Presentations

2023: Bucklaew, A., Coop, S. H., **Sarch, G. H.**, Mitchell, J. F. *Laminar and Cell Type Distinctions for Pre-Saccadic Attention in Marmoset MT/MTC*. Journal of Vision. Poster.

2022: Coop, S. H., **Sarch, G. H.**, Bucklaew, A., Yates, J. L., Mitchell, J. F. *Laminar Organization of Pre-Saccadic Attention in Marmoset Area MT*. Journal of Vision. Poster.

2019: **Sarch, G. H.**, Yates, J. L., Coop, S. H., Mitchell, J. F. *Identification of Cortical Layers from Current Source Density (CSD) Analysis and Two Local Field Potential (LFP) Band-Power Measures in Marmoset V1*. Society for Neuroscience. Poster.

2018: **Sarch, G. H.**, Pavindra, P. H., Smith, J. C. *Computational Modeling of Respiratory Neural Circuits*. NIH.

Invited Talks

2024: **Search-based Planning Laboratory**: Task Planning with LLMs. Carnegie Mellon University.

2024: **CMU Catalyst's "LLM Agents Seminar"**: Open-Ended Instructable Embodied Agents with Memory-Augmented Large Language Models. Carnegie Mellon University.

2023: **brAIn Seminar Talk**: "Spatial Processing During Natural Scene Viewing: Insights from Artificial Neural Network Modeling." Carnegie Mellon University.

2023: **Invited Lecture in Biologically Intelligent Exploration (CMU 85-435)**: "How Do You Use Evidence to Make a Decision?" Carnegie Mellon University.

Experience

Stealth Startup
Intern, Technical Staff, AI — *Aug 2024 – Present*
- Working on large-scale multimodal LLM training for autonomous web agents.
- Collaborated with founders to develop and optimize infrastructure supporting model deployment on real websites.

Microsoft Research
Research Intern — *May 2024 – Aug 2024*
Mentors: Andrew Wilson, Bala Kumaravel
- Led project utilizing multimodal cues (eye gaze, hand gestures) for AR vision-language model assistants.

Carnegie Mellon University
Ph.D. Student — *Aug 2020 – Present*
PIs: Katerina Fragkiadaki and Michael Tarr
- Research topics: Large language models, multimodal agents, embodied AI, 3D vision.
- Mentored Master's and Ph.D. students.
- Served on MLD Ph.D. Peers Committee to support incoming MLD Ph.D. students.

Active Vision Laboratory, University of Rochester
Research Assistant — *Aug 2018 – Aug 2020*
PI: Prof. Jude Mitchell
- Studied cortical laminar differences in primate pre-saccadic attention.

Neurobiology Laboratory, National Institutes of Health (NIH)
Internship Program — *May 2018 – Aug 2018*
PI: Dr. Jeffrey Smith
- Studied computational models of the pre-Bötzinger respiratory generator.

Cognitive Neurophysiology Laboratory, University of Rochester
Research Assistant — *Jan 2018 – May 2018*

PI: Prof. Edmund Lalor
- Studied contextualized semantics for speech comprehension decoding in EEG.

Skills

Programming Languages: Python, MATLAB, Unix shell/bash, Git, HTML/CSS

ML/AI Tools & Environments: PyTorch, TensorFlow, PyCortex, AI2-THOR, Habitat AI, CARLA, OpenAI Gym, Web/browser infrastructure

Laboratory Methods: Distributed LLM/VLM training, reinforcement learning, representation learning, computer vision, language model prompting, retrieval-augmented generation, multiple view geometry, search, probabilistic inference, statistical machine learning, density estimation, learning theory, signal processing, electrophysiology

Relevant Coursework: Deep Reinforcement Learning (CMU 10-703), Advanced Machine Learning (CMU 10-715 & 10-716), Intermediate Statistics (CMU 36-705), Graduate Artificial Intelligence (CMU 15-780), Statistical Models of the Brain (CMU 36-759), Cognitive Neuroscience (CMU 85-765 & 03-763)

Teaching Assistantships: Deep Reinforcement Learning (CMU 10-403), Biologically Intelligent Exploration (CMU 85-435)

Mentoring: Lawrence Jang (Master's Student), Sahil Somani (Master's Student), Raghav Kapoor (Master's Student), Snigdha Saha (Master's Student)

Reviewer Service: CVPR (2023, 2024), ICLR (2024), AISTATS (2024), NeurIPS (2023, 2024)