Cem Gokmen

CS PhD Candidate at Stanford, Specializing in Embodied AI and Robotics Stanford, CA 94305
cem@cemgokmen.com | www.cemgokmen.com | linkedin.com/in/cgokmen Updated: October 10, 2024

Research Statement

I develop machine learning algorithms enabling robots to perform complex, long-horizon tasks in real-world environments using techniques like learning from demonstrations, reinforcement learning, sim-to-real transfer, and leveraging off-the-shelf LLM/VLM models. I also have extensive experience building simulation environments and benchmarks to support this research.

Seeking research internships to further drive exciting robotics & embodied AI research.

Education

Stanford University, Stanford, CA Sep. 2022 - Present
Ph.D. in Computer Science, **Advisor:** Prof. Fei-Fei Li

Stanford University, Stanford, CA Sep. 2020 - Jun. 2022
M.Sc. in Computer Science GPA: 4.03

- Google Computer Science Research Mentorship Program (CSRMP) Fellow
- **Select Coursework:** Deep Learning • Principles of Robot Autonomy I • Decision Making Under Uncertainty • Interactive & Embodied Learning • Convolutional Neural Networks for Visual Recognition • Machine Learning with Graphs

Georgia Institute of Technology, Atlanta, GA Aug. 2016 - Dec. 2018
B.Sc. in Computer Science with Undergraduate Research Certification GPA: 3.83

Work Experience

Research Scientist Intern, NVIDIA Research
Hosts: Jim Fan, Yuke Zhu. AI Algorithms Group, Santa Clara, CA. Jun. 2023 - Sep. 2023

- Developed a method leveraging LLMs to generate demos for learning end-to-end transformer policies for long-horizon robotics tasks.

AI Resident, Google [x] / Everyday Robots
Hosts: Mohi Khansari, Daniel Ho. Mountain View, CA. Jun. 2022 - Sep. 2022

- Developed a method to allow robots to estimate their likelihood of success based on previous rollouts of a real-world imitation learning robot policy and decide whether to ask for help.

Software Engineer II, Google
YouTube Premium Team, San Bruno, CA Feb. 2019 - Sep. 2020

- Implemented critical user journeys and led software design, gained expertise on in-app messaging methods, developed features across Python/C++ backends and Android/iOS/Web frontends.

Highlighted Publications

C. Li*, R. Zhang*, J. Wong*, **C. Gokmen***, S. Srivastava*, R. Martin-Martin*, C. Wang*, G. Levine*, *et al.*, "BEHAVIOR-1K: A benchmark for embodied AI with 1,000 everyday activities and realistic simulation," in *6th Annual Conference on Robot Learning*, 2022, **(Co-First Author, CoRL Oral, Best Paper Nominee)**

- We developed **BEHAVIOR-1K**, a human-centered embodied AI benchmark with 1,000 diverse, everyday activities, addressing challenges like long-horizon tasks and complex manipulations in realistic environments. Website: https://behavior.stanford.edu
- BEHAVIOR is an ongoing, four-year project aiming to provide a robotics research direction grounded in human needs, offering unprecedented diversity and realism compared to traditional benchmarks usually focused on limited tasks and environments.
- I developed the process and managed the 10+ person team that acquired and created the 6k-object, 50-scene BEHAVIOR-1K dataset, a massive undertaking involving full cross-integration of the assets with the 1,000 task definitions and 4,000 nouns in the BEHAVIOR-1K knowledgebase. Visit https://behavior.stanford.edu/knowledgebase/ for some examples.
- I co-develop and maintain **OmniGibson** and the older-generation **iGibson**, the high-fidelity simulation environments to simulate the tasks, built on Omniverse Isaac Sim and PyBullet, respectively. Visit https://github.com/StanfordVL/OmniGibson for more info.

C. Gokmen, D. Ho, and M. Khansari, "Asking for help: Failure prediction in behavioral cloning through value approximation," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*, 2023

- We introduced **Behavioral Cloning Value Approximation (BCVA)**, a method to predict failures in robot policies trained using Imitation Learning where no state value estimate is avaiable by default, enabling robots to autonomously ask for help when needed.
- At Everyday Robots, we used our state value approximation to ask for help with 86% precision across over 2000 real-world trials, particularly in complex tasks like latched-door opening.

M. Li, S. Zhao, Q. Wang, K. Wang, Y. Zhou, S. Srivastava, **C. Gokmen**, T. Lee, *et al.*, "Embodied agent interface: Benchmarking LLMs for embodied decision making," in *Neural Information Processing Systems Datasets and Benchmarks Track*, 2024, **(NeurIPS Benchmark Track Oral)**

- We introduced the **Embodied Agent Interface**, a framework to systematically benchmark Large Language Models (LLMs) for decision-making tasks in embodied environments, a recently well-explored area of research that we believe is not benchmarked to an adequate standard.
- This framework allows standardized and fine-grained assessment of LLMs' abilities in handling complex, embodied decision-making scenarios, pinpointing issues like hallucinations and planning errors across 15 LLMs evaluated on benchmarks such as BEHAVIOR and VirtualHome.

Y. Ge*, Y. Tang*, J. Xu*, **C. Gokmen***, C. Li, W. Ai, B. J. Martinez, A. Aydin, *et al.*, "Behavior vision suite: Customizable dataset generation via simulation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2024, **(Co-First Author, CVPR Highlight)**

- We developed the **BEHAVIOR Vision Suite (BVS)**, a toolset designed to generate fully customizable synthetic data for evaluating computer vision models, addressing limitations in real-world datasets like poor labeling and diversity.
- BVS enables controlled experiments by allowing adjustments at multiple levels, including scene, object, and camera parameters, supporting tasks such as model robustness testing, scene understanding, and simulation-to-real transfer for vision tasks like unary and binary state prediction.

Other Publications

T. Dai, J. Wong, Y. Jiang, C. Wang, **C. Gokmen**, R. Zhang, J. Wu, and L. Fei-Fei, "ACDC: Automated creation of digital cousins for robust policy learning," in *8th Annual Conference on Robot Learning*, 2024

C. Gokmen, R. Zhang, S. Srivastava, C. Li, M. Lingelbach, R. Martin-Martin, S. Savarese, J. Wu, *et al.*, "Eye-BEHAVIOR: An eye-tracking dataset for everyday household activities in virtual, interactive, and ecological environments," *Journal of Vision*, Dec. 2022

S. Srivastava, C. Li, M. Lingelbach, R. Martín-Martín, F. Xia, K. E. Vainio, Z. Lian, **C. Gokmen**, *et al.*, "Behavior: Benchmark for everyday household activities in virtual, interactive, and ecological environments," in *Proceedings of the 5th Conference on Robot Learning*, 2022. [Online]. Available: `https://proceedings.mlr.press/v164/srivastava22a.html`

C. Li, F. Xia, R. Martín-Martín, M. Lingelbach, S. Srivastava, B. Shen, K. E. Vainio, **C. Gokmen**, *et al.*, "Igibson 2.0: Object-centric simulation for robot learning of everyday household tasks," in *Proceedings of the 5th Conference on Robot Learning*, 2022. [Online]. Available: `https://proceedings.mlr.press/v164/li22b.html`

S. Cannon, J. J. Daymude, **C. Gokmen**, D. Randall, and A. W. Richa, "A Local Stochastic Algorithm for Separation in Heterogeneous Self-Organizing Particle Systems," in *International Conference on Randomization and Computation (RANDOM)*, 2019. [Online]. Available: `http://drops.dagstuhl.de/opus/volltexte/2019/11269`

Teaching

Head Course Assistant, CS 231N: Deep Learning for Computer Vision
Stanford University — Mar. 2024 - Jun. 2024

- Led team of 22 course assistants, planned development of assignments and exams, mentored student projects, and managed logistics for Stanford CS's most popular, 650-student course.
- See https://cs231n.stanford.edu/2024/reports.html for mentored projects.

Skills

Languages: English (Fluent), Turkish (Native), French (Advanced), Spanish (Beginner).
Programming: Python, Java, C, C++, JavaScript, HTML, CSS.
Research: Reinforcement learning, imitation learning, physics simulation, robotics simulation, motion planning, classical robotics
Frameworks: PyTorch, TensorFlow, Ray, PyBullet, Omniverse, Isaac Sim, CUDA, Docker