

Matei Gardea

matei.gardea@berkeley.edu — mateig.github.io

Education

University of California, Berkeley

Bachelor's + Master's, Electrical Engineering and Computer Sciences

2021–2026

Berkeley, CA

- Hybrid Robotics Lab: humanoid robots, spatial reasoning, video diffusion models. Supervised by Koushil Sreenath.
- Head TA, CS188 (FA24): Intro to AI. Supervised by Pieter Abbeel and Igor Mordatch.

Research

Generative Video for Humanoid Control (GVHC)

2026

Matei Gardea, Koushil Sreenath

- Explored the use of video diffusion models as a source of motion priors for language-conditioned humanoid locomotion.
- Developed a modular monocular-video grounding pipeline combining image-conditioned video generation, temporally consistent video depth estimation, optical-flow-based 3D alignment, and mesh reconstruction.
- In-progress follow-on work includes research into improved robot pose estimation, contact-aware trajectory repair, and physical deployment.

What VLMs Need to See: Input Representations for Spatial Reasoning

2026

Matei Gardea, Ziteng (Ender) Ji, Max Yen

- Conducted a study investigating how various input preprocessing techniques applied to vision-language models (VLMs) affect spatial reasoning, with the goal of improving the performance of downstream VLM-based systems (e.g. VLAs).
- Constructed a controlled synthetic 3D spatial-QA benchmark with roughly 2k scenes, 19k questions, and matched RGB, depth, structured-text, and combined input conditions.
- Evaluated five open VLMs in the 2–4B parameter range; found that auxiliary spatial inputs often reduced overall accuracy, while structured descriptions improved performance by up to 47% on certain tasks, revealing a task-dependent grounding bottleneck for embodied perception.

VLM-Verified Counterfactual Hindsight for Sparse-Reward Manipulation

2026

Daniel Grant, Parshawn Gerafian, Matei Gardea

- Examined whether frozen VLMs could supply useful credit-assignment signal when sparse-reward robotic manipulation provided little or no TD-error signal.
- Designed VLM-guided replay-augmentation experiments; implemented SAC, HER, and PER in MuJoCo + Gymnasium.
- Benchmarked Semantic PER with simulator-verified counterfactual relabeling; reached 0.617 mean success on FetchSlide, +0.43 over HER at 250k steps.

AVIAN: Autonomous Visual Indoor Aerial Navigation

2025

Matei Gardea, Armaan Goklani, Ethan Bensimon

- Studied how far monocular perception, localization, and control could be pushed on a custom-built drone under extreme resource constraints (e.g. no LiDAR, mo-cap, external infrastructure, or onboard GPU).
- Designed and built a sub-250 g quadrotor from scratch (CAD, 3D printing, electronics) to fly a Raspberry Pi + wide-angle camera. Started by building a low-latency high-bandwidth digital streaming module before pivoting to an onboard solution for visual odometry due to thermal + latency issues with video streaming.
- Presented a functional self-contained drone autonomy system capable of visual waypoint navigation in a tight indoor environment.

FABS: Fully Automated Bone Marrow Lesion Segmentation

2020

Matei Gardea, V. Anwari, R. Usman, S. Liu, A. K. Wong

- Investigated automated segmentation of bone marrow lesions in knee MRIs, a challenging osteoarthritis biomarker to detect due to its irregular appearance on high-dynamic-range scans and limited labeled data.
- Designed and trained a fully convolutional neural network inspired by the UNet architecture and coordinated a multi-department annotation effort to create a dataset totaling 16k samples.
- Achieved a 0.709 Dice score; presented my work at the International Workshop on Osteoarthritis Imaging and the Canadian Arthritis Research Conference; won Best Poster at Toronto General Hospital Research Institute Research Day.

Work Experience

Royal Bank of Canada

2024

Research Scientist

Toronto, ON

- Researched transformer/CNN vision-language systems for document parsing and retrieval and built a cloud-based training and evaluation pipeline. Achieved state-of-the-art performance on document reconstruction benchmarks.

Joint Department of Medical Imaging, Toronto General Hospital / UHN

2019–2022

Research Intern

Toronto, ON

- Led medical imaging research under Dr. Andy Kin On Wong, coordinating MRI annotation efforts and developing segmentation models in TensorFlow for osteoarthritis imaging studies.

Skills

Robotics/ML: reinforcement learning, computer vision, 3D perception, SLAM, VLMs, VLAs, video diffusion models, sim-to-real transfer, UAVs, humanoids.

Tools: Python, C/C++, PyTorch, TensorFlow, OpenCV, MuJoCo, Isaac Lab, Blender, Linux, Git, distributed training, BetaFlight.