

Caleb Martin

Robot Learning Researcher · Visuomotor Policies

✉ caleb@arkrobotics.ai | 📞 [4064985019](tel:4064985019)



Summary

Independent robot-learning researcher. Built and trained novel visuomotor policies (Riemannian flow-matching on the $S^3 \times R^4$ product manifold, and a Mamba2 head conditioned via cross-attention FiLM over frozen DINOv3 patches) end-to-end on hardware I owned and operated. Full pipeline from VR teleoperation and a C++20 lock-free data collection service through training, eval, and closed-loop testbench evaluation on a UFactory xArm 7 and a YAM single-arm rig.

Profiles

github

[caleb272](#)

✕ X

[@Digital_DAT](#)

🌐 Portfolio

arkrobotics.ai/projects.html

Experience

Ark Robotics

SF / LA

Founder & Independent Robot Learning Researcher

Aug 2023 - Current

Solo research on visuomotor policies for bimanual manipulation, targeting real industrial deployment. Full research writeup at arkrobotics.ai/projects.html.

- **Riemannian Flow Matching Policy** (Sep 2024 – Jan 2025): Conditional flow matching on the $S^3 \times R^4$ product manifold: rotation as quaternion geodesics on S^3 , translation + gripper Euclidean. NVIDIA Cosmos Causal Video Tokenizer for vision conditioning, transformer decoder cross-attending to the memory as the action head, ODE integration via zuko + a custom fast Euler solver. Added proprioception masking (50%) to address the causal-confusion shortcut. Independently converged to π_0 's flow-matching action-head formulation about a week before their paper. ~90%+ closed-loop pick-and-place on a UFactory xArm 7.
- **Mamba2 Visuomotor Policy** (Sep – Oct 2025): 12-layer Mamba2 stack (hidden 1024, d_state 2048, expand 4) conditioned via per-layer cross-attention FiLM, with gamma, beta generated from learned queries cross-attending to frozen DINOv3-S/16 patch tokens. Residual MLP action head, 30 Hz closed-loop. Documented L1 mode collapse on naturally-clustered demonstrations and a high recovery rate from human intervention.
- **C++20 data collection service**: Single-writer / single-reader lock-free atomic ring buffers (acquire/release ordering, 64-byte aligned, pre-allocated, no allocations on the hot path). FlatBuffer-encoded MCAP output with Foxglove schemas so recordings open directly in Foxglove for debugging. VR-driven recording lifecycle with pre/post-capture buffers.
- **Iris remote teleoperation relay** (Nov 2025): C++ server using rtpengine kernel-mode RTP forwarding (sub-millisecond pass-through) plus a UDP relay for VR controller packets, with an HTTP API for operator-robot session pairings. Feasibility test for remote operators under simulated round-trip lag.
- Built the VR-based teleoperation system and post-processing for both the YAM and the UFactory xArm 7.
- Stack: Python / PyTorch + PyTorch Lightning, CARBS for cost-aware hyperparameter sweeps. Local training on a 3090 Ti / 4090 rig; 8× A100 runs via Lambda Labs.

arkrobotics.ai/projects.html

Transistor Bio

New York, NY

Robotics Engineer

January 2022 - October 2022

Built the data pipeline, camera system, and labeling / experiment admin tool for high-throughput preclinical drug testing on robotic platforms. Trained CV models to track mouse movement for behavioral phenotyping. Company shut down during the 2022 VC downturn.

<https://transistor.bio/>

Heavy-Lift Tree Trimming Drone

Austin, TX

Independent Project

March 2021 - January 2022

Designed and built a 40 lb-capable heavy-lift drone with a 3D-printed frame, onboard arm, and live-video ground control for arborist tree trimming. Mid-flight cutting dynamics proved harder than expected.

TripSuite

San Francisco, CA

Founding Full Stack Software Engineer

October 2022 - July 2023

Founding full-stack engineer. Built the app from zero to one (React, Node.js, Prisma / PostgreSQL); owned designer-mockup → shipped-feature handoff and contributed to architectural decisions.

<https://www.tripsuite.com/>

J.B. Hunt Transportation Services, Inc

Fayetteville, AR

Software Engineer 2

Sep 2020 - Mar 2021

Carrier 360 team ("Uber for freight") on TypeScript / React Native. Pushed for GraphQL adoption and cleaner architecture patterns.

<https://www.jbhunt.com/>

ICatch(Hookd App)

Bozeman, MT

Full Stack Mobile Engineer

May 2019 - October 2020

Architected the Hook'd fishing-competition app (Node.js, GraphQL / Relay, React Native) on a unified codebase running iOS, Android, and web. Used for live tournaments globally. Company shut down early in the pandemic.

<https://appadvice.com/game/app/hookd-fishing-tournaments/1486897806>

ServiceTarget

Bozeman, MT

Junior Full Stack Engineer

April 2017 - May 2019

First full-time engineer out of high school. Helped bootstrap the company from \$60K to \$500K+ ARR with Fortune 1000 customers; end-to-end design and build of the platform with fast release cycles.

<https://servicetarget.com/>

Skills and Technologies

Research areas

imitation learning, flow matching, diffusion policies, state-space models / Mamba2, causal confusion mitigation, action chunking



ML stack

PyTorch, PyTorch Lightning, CARBS, DINOv3, SAM 2.1, NVIDIA Cosmos Tokenizer, zuko, Mamba2, ViT, JEPA



Robotics & systems

VR teleoperation, MCAP / Foxglove, lock-free C++20, Drake kinematics, ROS, ROS2, Isaac Sim, RealSense D405 / D450, OAK-D Pro, rtpengine, CI/CD, Docker



Hardware operated

UFactory xArm 7, YAM single-arm, DH Robotics AG-160-95 gripper



Languages

Python, C/C++, TypeScript, OpenCL, ROS



