

Samuel Boccaro

(646) 574-4702 | shboccaro@gmail.com

EDUCATION

University of Maryland: College Park

Bachelor's of Science in Computer Science & Mathematics GPA: 4.0/4.0

- Select Coursework: Multivariable Calc, Linear Algebra, and Differential Equations I & II, Algorithms, Computer Systems, Programming Languages, Discrete Structures | Activities: BigThink (Audio-ML), Smith Investment Fund

Hunter College High School

GPA: 3.9/4.0 | SAT: 1570 | Regeneron ISEF Finalist, ML Club Co-President

College Park, MD

Expected May 2028

New York, NY

Sep. 2019 – Jun. 2025

EXPERIENCE

Undergraduate Research Assistant

University of Maryland Institute for Health Computing

Nov. 2025 – Present

North Bethesda, MD

- Researched SOTA single-cell RNA-seq modeling literature (Transformers, Diffusion, VAEs, Optimal Transport) and translated it into end-to-end benchmarks (training + prediction), comparing models against the lab's in-house model across standardized dataset splits.
- Implemented LoRA fine-tuning to improve model performance while reducing compute, iterating on rank/target modules/training schedules.
- Stabilized training/inference (formats, shapes, sparse/dense, GPU memory, learning rate) and refactored the model for Hugging Face compatibility (save/load, configs).

Research Intern

Institut Langevin

Jun. 2024 – Aug. 2024

Paris, FR

- Researched potential applications for generative image models in healthcare, focused on improving denoising and segmentation model performance with synthetic data generation
- Built a CNN for denoising + super-resolution of nerve images using fastai and fine-tuned Stable Diffusion 1.5 to generate synthetic corneal nerve images indistinguishable from real data.

PROGRAMMING PROJECTS

Automated 3D Visualization of Cell Behavior | Python, Pytorch, Celpose, Unsupervised Learning, Clustering

- Selected as a Regeneron ISEF 2025 Finalist, representing New York City at the international competition.
- Designed and implemented an AI pipeline to automatically segment cells in microscopy videos and reconstruct 3D meshes, enabling quantitative analysis of cell dynamics.
- Collaborated with Columbia's Kalderon Lab to deploy the software for germarium dataset analysis, producing 3D visualizations that enabled novel biological findings.
- Built an unsupervised clustering framework capable of running spectral, k-means, and HDBSCAN clustering, automatically selecting the best-performing method based on evaluation scores.

Arbitrage Trading Bot | Python, Websockets, REST, NLP, AsyncIO

- Built a Python-based cross-exchange arbitrage bot for Polymarket + Kalshi, pulling and standardizing 50,000 markets (top 25k each via REST) into a consistent format to enable reliable cross-platform comparison and pricing.
- Implemented an NLP matching pipeline that uses a lightweight embedding model for semantic candidate retrieval and Gemini Flash API verification to validate matches.
- Developed a low-latency monitoring + execution loop using websocket streaming for live quotes/order books, computing fee-aware arbitrage edge plus APR and expected profit; delivering 100 ms quote-to-alert latency, and semi-automated order placement gated by explicit user confirmation.
- Collected Polymarket L2 order book data for the top 2,000 markets over 4 months, building a clean historical dataset for backtesting and microstructure research (over 100M updates).

Autograd Engine and Neural Networks in Pure C++ | C++, Calculus, Linear Algebra, Memory Management

- Built a C++ linear algebra library (matrix/vector ops, dynamic memory management) and an MLP framework with backprop, activations, and SGD training.

PUBLICATIONS

4D Imaging of the Germarium Suggests That Follicle Stem Cells and Follicle Cells Self-Organize Around Germline Cysts. Amy Reilein, Taj R Chhabra, **Samuel Boccaro**, Shafrir Pervez, Daniel Kalderon. Abstract accepted for **oral presentation** at the 67th Annual Drosophila Research Conference, March 2026.

Synthetic Procedural Noise and Neural Networks: Enhancing Biomedical Images with Purely Artificial Data Viacheslav Mazlin, Samer Alhaddad, **Samuel Boccaro**. Accepted for **oral presentation** at SPIE Photonics West 2025 and the European Conferences on Biomedical Optics 2025.

AWARDS & SKILLS

Awards: ISEF Air Force Research Laboratory Award, USACO Gold

Skills: Python, C++, Pytorch, JAX, Numpy, Git, HuggingFace, LaTeX