Ryan Kim

224-517-7069 | ryankim17920@gmail.com | linkedin.com/in/ryankim17920 | github.com/RyanKim17920 | ryankim.me

Education

University of Illinois Urbana-Champaign — Urbana, IL
B.S. in Computer Science — *Expected May 2029*

William Fremd High School — Palatine, IL
GPA: 4.0/4.0 (UW), 5.207 (W) | AIME Qualifier (4x) | Simons Fellow — *Aug 2021–May 2026*

Professional Experience

ML Research Contributor — Apr 2026–Present
MedARC — *Remote*
- Ranked **#1** on **NanoPath**, MedARC's **pathology foundation-model competition**; my dual-branch anti-overfitting approach accounts for **47%** of all leaderboard gain over baseline.
- Reproduced **WAIV's robustness work** locally, surpassing its results on **7 of 8 tasks** with a **34%** gain in robustness to **scanner noise**; served local LLMs with **vLLM**.

Pathology AI Research Intern — Oct 2024–Present
AI in Medicine & Computational Biology Lab, Weill Cornell Medicine — *New York, NY*
- Benchmarked **12 pathology foundation models** (UNIv2, Prov-GigaPath) across TCGA and internal cohorts.
- Quantified a **2× rise** in out-of-distribution tumor-purity error (**MAE ∼0.10 → ∼0.20**) — a clinical deployment risk.

Computational Oncology Research Intern — Dec 2024–Jan 2026
Advanced Computing & Oncology Lab, Memorial Sloan Kettering — *New York, NY*
- Engineered a **CPU-only**, locally-deployed clinical LLM pipeline (**LLaMA 3.1-8B**, **Qwen3-4B**, DSPy/MIPROv2) to classify lung cancer CT reports across **370 clinically annotated reports**; co-authored a referral-event analysis abstract accepted at the **ACRO 2026 Summit**.
- Applied Cox regression and survival analysis on **230** patients' clinical and demographic records to identify correlations between percent thymic tissue and NSCLC outcomes.

Simons Summer Research Fellow (Computational Genomics) — June 2025–Present
Yurovsky Lab, Stony Brook University (Simons Foundation) — *Stony Brook, NY*
- Predicted glioblastoma survival by architecting a **DNABERT-based** prognostic model using Cox regression on non-coding regulatory DNA mutations.
- Identified significant age-related survival signals and racial disparities encoded in mutational scores.

Lead Technical Developer — Nov 2025–Apr 2026
MathLinks.org — *Remote*
- Led engineering for a full-stack math competition platform serving **150,000+ users** across **11 partner organizations**; secured **$20,000** in seed funding (Hudson River Trading, Innovate901 1st Place).
- Built React/Node.js infrastructure for high-concurrency contest workflows, including authentication, submissions, and leaderboards.

Full-Stack Developer — Aug 2023–July 2024
Cyberlinc, Inc. — *Remote*
- Engineered a secure crowdfunding platform with Flask and SQLAlchemy, including payment integration.

Technical Projects

PapersToCode | *Python, Full-Stack, Data Pipeline* — Mar 2024–Mar 2026
- Built an automated paper-to-GitHub matching pipeline across **500,000+ ML/AI papers**, identifying papers that still lack public implementations.

micro-DINOv3 | *Pure Python, Computer Vision* — Feb 2026–Apr 2026
- Recreated **DINOv3** in dependency-free Python: hand-written reverse-mode autograd and ViT, EMA teacher-student distillation, DINO + iBOT objectives, KoLeo regularization, axial RoPE with register tokens, and Gram anchoring.

Deep Mixture of Experts via Shallow Embedding | *PyTorch, Computer Vision* — July 2024–Nov 2024
- Reproduced the DeepMoE architecture in PyTorch, recovering its accuracy-vs-compute tradeoff on CIFAR-10; identified a gate-scaling failure at small embedding widths that destabilizes training at the paper's learning rate.

Technical Skills

Languages: Python, C/C++, Java, SQL (PostgreSQL), JavaScript, HTML/CSS, R.
ML/Data Science: PyTorch, TensorFlow, DSPy, HuggingFace, Scikit-learn, Polars, Pandas, NumPy.
Domain Expertise: Computational Pathology, Clinical NLP, Survival Analysis, Computer Vision.
Web/Backend: React, Next.js, Flask, FastAPI, MongoDB.
Infrastructure & Tools: SLURM, vLLM, Docker, Linux, Git.