Wyett "Huaye" Zeng

wyettzeng@g.harvard.edu | 519-729-8107 | wyett-zeng.com | LinkedIn | GitHub | Google Scholar

Highlights

- **500+ citations** for machine learning publications in top venues, including first-author and best paper awards [Google Scholar].
- **Open-Source Contributions**: Co-lead of verl-Tinker, contributed RL, VLM, and LoRA functionalities to verl, veomni, and vLLM.
- Research Topics: Reinforcement Learning, LLM post-training, Reward Modeling, Coding Agents, LoRA, Distributed Training.
- Languages & Tools: PyTorch, Ray, Triton, Cuda, DeepSpeed, vLLM, SGLang, VeOmni, FSDP, Slurm, Docker, Kubernetes.

Work Experience

Research Scientist Intern | ByteDance / TikTok Jun 2026 – Aug 2026 | San Jose, USA

- Optimized Raskel, an internal multi-LoRA training framework, tuning batching and parallelization strategies to improve **MFU by 5%** and optimizing its **Triton multi-LoRA GEMM kernel** to increase H20 throughput by up to **20%** across different workloads.
- Built **verl-Tinker**, an open-source Tinker-compatible API server built on verl, enabling fine-tuning with VeOmni/FSDP actors, vLLM/VeXact rollout, reference engines, and dynamic model offloading for RL training on **120B models** across 64 GPUs [link].
- Implemented importance-sampling and DRO losses in verl; added streaming synchronization for GPT-OSS MoE weights, whose gate and up projections are interleaved across experts, **reducing peak memory** for the affected block to **1%–3% of the original**.
- Added LoRA support for **VLM** training in VeOmni, along with MFU calculation logic for performance analysis and optimization.
- Researched LoRA learning rate estimation and hyperparameter sensitivity, analyzing how gradient-matrix singular values correlate with downstream loss-curve trends across rank, learning rate, alpha, and matrix initialization to predict stable near-optimal configurations under practical search constraints, reducing hyperparameter-search time by **80%** compared with exhaustive sweep.

Machine Learning Engineer Intern | GPTZero (acq. by Superhuman / Grammarly) Sep 2024 – Aug 2025 | Toronto, Canada

- Built the end-to-end **evaluation** and **data pipeline** for **AI Reviewer**, including metric design, iterative dataset curation and annotation, and scalable third-party labeling workflows; led development of the transformer-based reviewer in PyTorch, achieving **88% accuracy** within ±1.0 IELTS band and **97%** within ±1.5 bands, approaching human-grader agreement on standardized scoring.
- This work led to a **pending U.S. Patent**: "Calibrating LLM Reviewer with Per-Criterion, Multi-Stage Human Alignment."

Quantitative Developer Intern | Boosted.ai Jan 2024 – Apr 2024 | Toronto, Canada

- Rewrote the core factor model algorithm, reducing scheduled **inference time by 90%** and saving **500 compute hours** weekly on AWS EC2; used **NumPy**, ClickHouse, and PostgreSQL to update daily economic factors for **20,000+ public securities**.

Research Experience

Swarm Code (In Progress) | Embodied Minds Lab **Harvard University**

- Designed a decentralized **multi-agent coding harness** around a shared decision and dependency graph, enabling agents to coordinate task decomposition, exchange reusable findings, and build on partial solutions through local planning and execution.
- Designing an **agentic multi-turn reinforcement learning** setup to train collaborative coding policies, combining task-level rewards with contribution-level credit to encourage effective coordination and reuse of intermediate solutions across long-horizon tasks.

AceCoder (ACL 2025, First Author) | Tiger Lab | [paper][website][model] **University of Waterloo**

- Developed a fully automated pipeline for large-scale synthesis of (question, test-case) pairs using **vLLM** and **SGLang**, enabling efficient inference generation to create both preference-pair and RLVR datasets, which was included into **OLMo 3**'s training corpus
- Trained Bradley-Terry reward models on the above dataset with LlamaFactory and **DeepSpeed ZeRO** Stage 3, then used the finetuned reward model to rerank best-of-32 candidate responses at test time; achieved average benchmark gains of **10%** for Llama-3.1-8B-Instruct and **5%** for Qwen2.5-Coder-7B-Instruct, **bringing the 7B model to parity with DeepSeek-V2.5-236B**.
- Conducted **R1-style reinforcement learning** with both reward models and test-case pass rewards via **PPO** and **REINFORCE++**, leading to consistent improvements across benchmarks; notably Qwen2.5-Coder-base improved on HumanEval-plus by over **25%**.

MANTIS (TMLR 2024, Outstanding / Best Paper Award) | Tiger Lab | [paper][website][demo] **University of Waterloo**

- Investigated the multi-image reasoning capabilities of LLMs by interleaving image tokens from **vision encoders** such as CLIP and SigLIP with textual instructions, enhancing tasks such as co-reference, comparison, and temporal understanding.
- Fine-tuned the Fuyu model on Mantis-Instruct, achieving a **13%** performance improvement over the **SoTA baseline**, Idefics2-8B.

Education

Harvard University Cambridge, Massachusetts, USA
Master of Science in Data Science (GPA: 3.96/4.0) December 2026 (Expected)

University of Waterloo & Wilfrid Laurier University Waterloo, Ontario, Canada
Bachelor of Computer Science and Bachelor of Business Administration Double Degree April 2025
Laurier Alumni Gold Medalist (Major Average: 94.79/100)