Liu Yuchen *(Lyrica)*

Independent Researcher · 0-to-1 Builder

yliuoe@connect.ust.hk · github.com/kaminoguo

Research

Four papers submitted to NeurIPS 2026 (3 sole-author, 1 co-first). Titles, anonymous repository identifiers, and distinctive method/numerical details are withheld below to preserve double-blind review integrity; full information is available on request after the review period.

Sole author. Mechanistic interpretability of transformer language models — shows that a class of widely-used spectral diagnostics on hidden states is systematically biased by an input-invariant high-variance direction, and proposes a one-line correction; consistent results across multiple open model families and scales.

Sole author. Activation-ablation methodology for Vision-Language-Action (VLA) models — demonstrates that the standard ablation primitive overstates per-layer importance in VLAs, introduces a corrected protocol implemented as forward-hook context managers, and uses it to expose token-role structure in VLA fine-tuning; closed-loop validation in a standard robotic simulator.

Sole author. Cross-view attention analysis in VLAs — quantitative framework relating measurable attention partition to behavioral multi-view fusion, with a training-free diagnostic released as a small reusable library.

Co-first author. Deployment-failure audit suite for VLAs — mutually-exclusive failure-mode taxonomy with calibrated confidence intervals and a CI-style ship-gate, demonstrating shipping-blocking failure modes that look acceptable on aggregate-success metrics alone.

Experience

Founder, OfferI Feb 2025 – Dec 2025 *(closed)*

- AI study-abroad school selection agent for international students.
- Built large-scale distributed crawler system and end-to-end agent workflow.

Projects

Squirrel — AI Memory Layer Dec 2025 – Feb 2026

Memory system for AI coding agents (Claude Code, Cursor, Codex). Rust daemon + MCP server; Python service for LLM-based memory extraction. Organization deleted at 1,000+ GitHub stars.

Xiaoniao — Cross-Platform Translation Tool *(open source)* Aug – Sep 2025

Multi-platform translation tool in Go: Chrome extension + Windows app; three AI provider integrations.

Education

Hong Kong University of Science and Technology Sep 2024 – 2029 *(expected)*

Electronic Engineering & Artificial Intelligence, Dual Major

- Top 3 in COMP6211 (PhD-level AI Startup Course), Summer 2025.

Technical Skills

Languages. Rust, Python, TypeScript, Go (proficient); comfortable picking up new languages quickly.

Research stack.

- *Mechanistic interpretability* for LLMs and VLAs — activation interventions, attention analysis, linear probing, spectral analysis of hidden states.
- *Robotics evaluation* — closed-loop VLA rollouts in MuJoCo-based simulators; deployment audits with failure-mode taxonomies and bootstrap / Wilson confidence intervals.
- *Hands-on models* — mainstream open VLAs (OpenVLA family, π_0, etc.); foundation LLMs and VLMs across the GPT-2 / Pythia / OLMo / Qwen / LLaMA / LLaVA families.
- *Stack* — PyTorch with HuggingFace Transformers / datasets / accelerate.

Hardware. LeKiwi mobile manipulator; dual-arm bimanual setups; embedded systems; mini-PC and SBC development.

Platforms. Native development across Windows, macOS, and Linux (Ubuntu, Kali, Fedora, NixOS, Arch).