Youssef Ahmed

Hangzhou, China | youssef@westlake.edu.cn | +86 198 5814 4250 | LinkedIn | yjoe177.github.io

AI SYSTEMS | EFFICIENT INFERENCE | EMBEDDED ENGINEERING

Education

Westlake University — B.S. in Artificial Intelligence Aug 2025 – Jun 2029

- **Academic record:** 2025–26 GPA **3.787**; full scholarship recipient. Relevant study: algorithms, mathematics, programming, physics and ML.

Technical Skills

Languages / ML: Python, C/C++, PyTorch, TensorFlow, NumPy, pandas, scikit-learn

AI Systems: LLM-agent workflows, tool use, structured generation, automated evaluation, quantization

Embedded / Tools: STM32, ARM Cortex-M, ARM compilers, STM32CubeIDE/CubeMX, Git, Linux/WSL

Research & Engineering Experience

ENCODE Lab, Westlake University — Undergraduate Researcher 2025 – Present

Efficient AI and embedded systems | Advisor: Prof. Huan Wang

- Build embedded AI systems for constrained STM32 hardware, from architecture and quantization through C/C++ inference, profiling and deployment.
- Develop correctness-first testing and benchmarking tools for ARM compilation, integer equivalence, Flash/RAM footprints, inference latency and hardware verification.

Westlake Chemistry 2026 Summer Undergraduate Research Internship Summer 2026

Computational research | Prof. Yiqin Gao's group

- Simulated **862 TIP3P water molecules** with SPONGE; analyzed molecular structure using oxygen–oxygen radial distribution functions.

Selected Technical Projects

Agentic TinyML — Verification-Grounded LLM Engineering 2026 – Present

- Developed an agent-oriented MCU engineering workflow integrating structured code generation, repair, static/semantic checks and ARM build tests.
- Built deterministic test harnesses and auditable evidence records so generated kernels are evaluated by correctness and device-side measurements, not LLM output alone.

TinyMamba — Quantized Keyword Spotting on STM32 2025 – 2026

- Designed and deployed compact sequence-model inference on the **STM32F446RE (168 MHz Cortex-M4)**; integrated log-mel features, integer/C validation and embedded profiling.
- Reached **strong test accuracy** with a **compact parameter count** on a controlled 10-keyword benchmark; validated quantized inference and profiled device memory/latency constraints.

Imbalanced Classification 2026

- Built NumPy optimization baselines on an imbalanced training set; threshold tuning improved same-model validation F1.

Selected Honors & Leadership

2026 National Undergraduate Innovation and Entrepreneurship Training Program: TinyMamba selected at the **national** level with **$1,500** in funding.

Photon Cup, Westlake University (2026): Finalist in a university-level innovation competition among **10 participating projects**.

Ebhar Misr AI & Technology Program: Selected among **6 technology-track teams** from **5,000+ applicants**.

Teaching: Westlake Pre-college Physics Track TA; supported high-school students with SSH circuit experiments and measurement.