SEOYOON YUM

gilbertyum@kaist.ac.kr | GitHub | Website | Google Scholar

EDUCATION

Korea Advanced Institute of Science and Technology (KAIST) *2022 – Present*
B.S. in Computer Science
Selected coursework: Machine Learning, Deep Learning, Operating Systems, Computer Architecture, Systems Programming, Algorithms.

RESEARCH EXPERIENCE

SEAL Lab, KAIST – Research Intern *Sep. 2026 – Present*
Advisor: Prof. Sehoon Kim

- Researching NVFP4 W4A4 quantization and its performance in looped transformer models.
- Optimizing inference kernels to improve inference speedups.

PhaseGate — Self-Initiated Research Project *Jun. 2026 – Sep. 2026*
Independently initiated; subsequently collaborated with Prof. Sehoon Kim (KAIST).

- Built a resource scheduler for on-device LLMs, achieving 2.0× retrieval throughput over a fixed scheduling baseline under the same latency constraint.
- Paper accepted to the NeurIPS 2026 Workshop on On-Device Intelligence.

UC San Diego – Research Intern *Jun. 2026 – Aug. 2026*
Advisor: Prof. Falko Kuester

- Built a zero-shot anomaly detector using a time-series foundation model on real-world operational data.
- Evaluated detection of proxy anomalies and alert-volume tradeoffs; showed that forecast residuals did not reliably predict genuine failures.

SGVR Lab, KAIST – Undergraduate Researcher *Jan. 2026*
Advisor: Prof. Sung-Eui Yoon

- Evaluated vision-language-action (VLA) policies in simulation under 11 meaning-preserving perturbations; average success fell from 60% to 25%, with structural edits causing the largest drop.

CJ AI Center – Research Intern *August 2023*

- Built a real-time brain–computer interface mapping EEG signals to robotic-arm commands using PyTorch.

PUBLICATIONS

[1] **S. Yum** and S. Kim, "PhaseGate: Phase-Aware CPU Retrieval Scheduling for On-Device LLMs on Unified Memory." NeurIPS 2026 Workshop on On-Device Intelligence: Foundation Models under Real-World Constraints.

LEADERSHIP

KAIST American Football Club – Captain *Sep. 2025 – Present*

- Lead team practices, game strategy, and intra-university competitions.

Republic of Korea Army – Squad Leader and English Interpreter *Jan. 2024 – Jul. 2025*

- Led eight soldiers and interpreted between Republic of Korea and U.S. military personnel.

TECHNICAL SKILLS

Programming	Python, C
ML systems	PyTorch, LLM inference, low-precision quantization, GPU kernel optimization, performance profiling
Tools	Linux, Git, Bash
Languages	Korean (native), English (professional; TOEFL 110)