Jiyun Bae

jiyun.bae@postech.ac.kr github.com/jiyunBae007 Google Scholar

Research Interest

My research focuses on understanding and improving the reliability of reasoning in multimodal AI systems. I study how reasoning models behave in the presence of irrelevant information, and how distractors shape their test-time scaling and robustness. My broader interests include visual reasoning, evaluation of reasoning models, and multimodal systems that reason reliably in cluttered, real-world environments.

Education

Ph.D. in Artificial Intelligence Mar. 2026 – Present
Pohang University of Science and Technology (POSTECH)
Advisor: Prof. Jaeho Lee

M.S. in Artificial Intelligence Feb. 2024 – Feb. 2026
Pohang University of Science and Technology (POSTECH)
Advisor: Prof. Jaeho Lee

B.E. in Information Security Mar. 2020 – Feb. 2024
Seoul Women's University

Publications

Understanding the Effects of Distractors on Reasoning Vision-Language Models 2026
***Jiyun Bae**, Hyunjong Ok, Sangwoo Mo, Jaeho Lee; EMNLP 2026*

- Introduced *Idis*, a 277K-example VQA benchmark that systematically varies distractor modality, number, and semantics across perception and reasoning-centric tasks.
- Identified a modality-dependent failure mode in test-time scaling: visual distractors degrade accuracy without lengthening reasoning, whereas textual distractors intensify inverse scaling.
- Showed via attribute-level trace analysis and attention-blocking experiments that distractor attributes verbalized in reasoning traces causally drive incorrect answers, motivating a prompt-based mitigation.

Experience

Ericsson-LG | *Software Developer Intern* Jan. 2022 – Feb. 2022

- Wrote C++ unit tests (Google Test) for the SMF node's HTTP client component in a commercial 5G Core network codebase, contributing code through Gerrit-based review in a Scrum R&D team.

S2W | *NLP Researcher Intern* Jul. 2022 – Sep. 2022

- Built a category-hierarchy crawler (BeautifulSoup) that collected ∼7.9K cybersecurity Wikipedia articles, empirically capping traversal depth to prevent topic drift into irrelevant pages.

Proficiencies

Languages: Korean (Native), English (Professional Working Proficiency)
Machine Learning Libraries: PyTorch (Advanced), TensorFlow (Advanced)

Teaching

Teaching Assistant: CSED554 Deep Learning for Natural Language Processing, Spring 2025

References

Jaeho Lee, Associate Professor at POSTECH
E-mail: jaeho.lee@postech.ac.kr

Sangwoo Mo, Assistant Professor at POSTECH
E-mail: sangwoo.mo@postech.ac.kr