

Doyeon Kim

doyeonkim2002@gmail.com | Seoul, Republic of Korea

 [LinkedIn](#) |  [Github](#) | [Personal Website](#)

EDUCATION

• Sogang University

B.S. in Computer Science and Engineering

◦ GPA: 3.57/4.3

Mar. 2022 – Feb. 2026

Seoul, Republic of Korea

RESEARCH INTERESTS

- **Reasoning in Intelligent Systems:** Developing methods for logical, commonsense, and context-aware reasoning to enable robust decision-making in dynamic environments.
- **Multimodal Understanding:** Integrating vision, language, and audio to build models capable of grounded world understanding.
- **Embodied and Interactive Agents:** Designing agents that learn through real-world interaction and embodiment, enabling effective task execution in physical environments.

PUBLICATIONS

- [1] **LNS: A Lightweight Navigating Searcher for Visually-Rich RAG**
Doyeon Kim, Sangmin Lee, Sein Ha, Hyunbin Park, Du-Seong Chang
In preparation for submission to ACL 2026 [\[PDF\]](#)
- [2] **Quality over Quantity: The Impact of Dataset Refinement on SFT Model Performance**
Doyeon Kim, Sangmin Lee, Dayoung Lee, Sein Ha, Myoungsub Kim, Du-Seong Chang
Korea Software Congress 2025 [\[PDF\]](#)
- [3] **Performance Enhancement of an Agentic RAG-Based Agricultural QA System through Faithfulness Verification with Chain-of-Thought**
Byoungjun Oh, Doyeon Kim, Dayoung Lee, Sein Ha, Siwook Lee, Du-Seong Chang
Korea Computer Congress 2025 [\[PDF\]](#)

RESEARCH EXPERIENCE

- **Undergraduate Research Assistant, Machine Learning Systems (MLSys) Lab** Jul. 2024 – Jul. 2025
Advisor: Prof. Euhyun Moon, Sogang University
 - **Topic: Tensor-Core-Compatible Optimization for CP Decomposition** [\[code\]](#)
 - Investigated CP Decomposition, a tensor factorization method typically solved using the CP-ALS algorithm.
 - Identified a major computational bottleneck in CP-ALS caused by the MTTKRP (Matricized Tensor Times Khatri-Rao Product) operation, which heavily depends on the Khatri-Rao product.
 - Investigated a reformulation that converts the Khatri-Rao product into a matrix multiplication, enabling Tensor Core compatibility.
- **Undergraduate Research Assistant, NLP & Intelligent Spoken Dialogue Interface System (ISDS) Lab** Aug. 2025 – Present
Advisor: Prof. Du-Seong Chang, Sogang University
 - **Topic: Data Quality Optimization and RL-Based Multimodal Search Agents for Agentic RAG** [\[code\]](#)
 - Developed a multimodal search agent for Agentic RAG using an SFT → RL training pipeline built on Qwen2.5-VL-7B-Instruct with a frozen generator architecture.
 - Conducted dataset refinement for multimodal SFT, filtering noisy reasoning trajectories into high-quality samples and demonstrating that data quality outweighs dataset size, achieving a **13.89%** accuracy gain with half the data (Work published in Korea Software Congress 2025).
 - Extended the SFT-initialized model with reinforcement learning, building a multimodal search agent capable of region selection and cropping for improved visual grounding and retrieval quality.
 - Achieved strong answer accuracy improvements under limited data settings, extending ideas from s3 and VRAG-RL, and preparing the work for submission to ACL 2026.

PROJECTS

• **CUDA Matrix Multiplication Optimization** [\[code\]](#)

Tools: C/C++, CUDA

- Learned and applied optimization techniques such as global memory coalescing, shared memory cache-blocking, 1D blocktiling for calculating multiple results per thread, and increasing arithmetic intensity via 2D blocktiling.
- Achieved near-cuBLAS performance.

• **Fact-Verified Agentic RAG for Agricultural QA Systems** [\[code\]](#)

Tools: Python, FAISS, BAAI/bge-m3, GPT-4-Turbo, Upstage Solar-1-Mini, Streamlit

- Developed an Agentic RAG-based agricultural QA system enhanced with two fact-verification techniques: Chain-of-Thought prompting and a post-hoc Fact-Checker algorithm.
- Implemented a baseline RAG QA model using FAISS and the BAAI/bge-m3 embedding model, then integrated CoT reasoning (GPT-4-Turbo) and a fact-checking module (Upstage Solar-1-Mini).
- Developed a Streamlit-based agricultural QA chatbot that provides real-time responses for both crop disease diagnosis and broader agricultural knowledge.
- Experimental results demonstrated notable improvements over the baseline in intent understanding, semantic similarity, and faithfulness:
 - * CoT: Intent Understanding improved by **8.4%**, and Semantic Similarity improved by **7.4%** over the baseline.
 - * Fact-Checker: Faithfulness increased by **0.6%** compared to the baseline.
- Work published in Korea Computer Congress 2025.

SKILLS AND TECHNIQUES

- **Programming Languages:** C, C++, Python, Java
- **GPU Programming:** CUDA

CAMPUS ACTIVITIES

- **Member, CNU Web Development Club**, Sogang University Mar. 2022 – Present

LEADERSHIP EXPERIENCE

- **Team Lead, Industry-Academia Project (CSE4187)** Sep. 2025 – Present
Data Quality Optimization and RL-Based Multimodal Search Agents for Agentic RAG (with DaedongAI Lab)

HONORS AND AWARDS

- National Scholarship 2022
- Academic Support Scholarship 2022

REFERENCES

1. **[Euhyun Moon](#)**
Associate Professor, Sogang University
Email: ehmoon@sogang.ac.kr
2. **[Du-Seong Chang](#)**
Associate Professor, Sogang University
Email: dschang@sogang.ac.kr
3. **[Jungmin So](#)**
Professor, Sogang University
Email: jsol@sogang.ac.kr