AI RESEARCHER / MULTIMODAL INTELLIGENCE

Inha Kang

always curious.

SEOUL, KOREA

Ph.D. Student, KAIST · Research Intern, KRAFTON

rkswlsj13@kaist.ac.kr · 2na-97.github.io ↗ · +82 10 6824 7177
Google Scholar ↗ · GitHub ↗ · LinkedIn ↗ · Project portfolio ↗

Vision-language models
Grounding / Explainable AI

LLMs & agentic systems
Tool use / Evaluation / RL

Physical intelligence
World models / Uncertainty

Understanding failures. Building better AI.

I turn insights from model failures into **better data, representations and learning objectives**. My research connects vision-language reasoning, reliable agentic tool use and physically grounded prediction.

ECCV / ICLR / CVPR
2026 first/co-first-author work

QIFK 2026 Finalist
Qualcomm AI Research

NeurIPS Top Reviewer
2026 main-track reviewing

01 Research & experience

KRAFTON Research Intern — 31 Aug 2026 – Present
LLM foundation models / Agentic tool calling

- Evaluated **BFCL v4** and **tau3-bench**; separated model failures from parser, serving and evaluation-protocol effects.
- Developing **role-aware tool-use data curation** after mid-training/SFT audits, and designing **executable RL tasks** with replay and state-based verification.

KAIST Doctoral Researcher — Aug 2024 – Present
CVML Lab / Advisor: Prof. Hyunjung Shim

- Probe spatial, compositional and language biases in VLMs, then improve **visual grounding, reasoning and adaptation** through representations and learning objectives.
- First/co-first-author work at **ECCV, ICLR and CVPR 2026**; current research in **physical prediction, uncertainty and task-specific world models**.

FaindersAI Deep Learning Researcher — 2022 – 2024
Data-centric AI / Multimodal perception

- Built **VLM-based segmentation-data curation** and real-time retail change detection; researched continual learning and heterogeneous-sensor fusion.

02 Education

KAIST Ph.D. student, Artificial Intelligence — 2024 – 2028 expected
CVML Lab / Prof. Hyunjung Shim / GPA: **4.3/4.3**

KAIST M.S., School of Computing — 2020 – 2022
CGV Lab / Prof. Jinah Park / GPA: 3.88/4.3

GIST B.S., Electronic Engineering & Computer Science — 2016 – 2020
Salutatorian / GPA: 3.94/4.5 / Minors: Business Administration and Sociology

UC Berkeley Visiting study, Computer Science — 2017

03 Selected research

First/co-first-author work. Full publication list on my website. * Equal contribution.

ECCV 2026

Paper ↗

Blind to Position, Biased in Language: Probing Mid-Layer Representational Bias in Vision-Language Encoders for Zero-Shot Language-Grounded Spatial Understanding

N. M. An*, **I. Kang***, M. Lee, H. Shim

Vision-Language Models · Explainable AI · Spatial Reasoning

Probed positional and language biases, then reused mid-layer representations for **training-free grounding and retrieval**.

ICLR 2026

Paper ↗

What "Not" to Detect: Negation-Aware VLMs via Structured Reasoning and Token Merging

I. Kang, Y. Lim, S. Lee, J. Choi, J. Choe, H. Shim

Multimodal Reasoning · Efficient Adaptation · Data-Centric AI

Addressed affirmative bias with structured data, token merging and targeted LoRA; up to **+10.8 NMS-AP** on OVDEval negation.

CVPR 2026

Paper ↗

Real-Time Long Horizon Air Quality Forecasting via Group-Relative Policy Optimization

I. Kang, E. Kim, W. Ryu, J. Shin, S. Yu, Y.-H. Kang, S. Jeong, E. Kim, S. Kim, H. Shim

Foundation Models · Reinforcement Learning · AI for Science

Aligned forecasting with asymmetric decision costs via GRPO; **47.3% lower false-alarm rates** than supervised fine-tuning alone.

PLUME / OPEN DATASET — Paper archive forthcoming

Direct multiday particulate forecasting with physically conditioned correction and ordered uncertainty over East Asia

I. Kang*, W. Ryu*, S.-C. Hong, J.-B. Lee, S. Ban, S. Jeong, Y.-H. Kang, S. Kim, H. Shim

Direct forecasts from issue-time evidence, with composition-guided corrections and an ordered predictive distribution. Public data combines CMAQ reanalysis, gridded observations and model weights.

2016–2024 · Hourly · **0.25°** · Leads to **+120 h**

Dataset on Hugging Face ↗

04 Technical strengths

Multimodal learning & model analysis

Vision-language grounding; representation probing; attention analysis; LoRA; geometric distillation; 3D perception; self-supervised learning.

LLM evaluation & post-training

SFT data curation; multi-turn tool-use trajectories; schema validation; deterministic replay; executable RL tasks; state-based verification; GRPO.

Research engineering

Python; C++; GPU computing; Kokkos; data validation and provenance; reproducible evaluation pipelines.

Also recognized: CVPR Compute Transparency Champion (2026); MICCAI MOOD Challenge, 1st place in 2021 and the 2022 sample-level task.