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

Vision-language reasoning, foundation models and physical prediction. * 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.

Findings of EMNLP 2025 Paper ↗

3D-Aware Vision-Language Models Fine-Tuning with Geometric Distillation

S. Lee*, J. Choi*, **I. Kang**, J. Kim, J. Park, H. Shim

Multimodal AI · 3D Reasoning · Knowledge Distillation

Distilled geometric cues from **3D foundation models** into VLMs to improve spatial reasoning without changing the architecture.

CVPR 2025 Paper ↗

No Thing, Nothing: Highlighting Safety-Critical Classes for Robust LiDAR Semantic Segmentation in Adverse Weather

J. Park, H. Lee, **I. Kang**, H. Shim

3D Perception · Robustness · Domain Generalization

Addressed semantic and local feature corruption to preserve **safety-critical objects** under adverse-weather domain shift.

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 Additional publications

10 publications and challenge papers in total; PLUME is listed separately as current research.

CVPR 2023 Paper ↗

Why is the winner the best?

M. Eisenmann, A. Reinke, V. Weru, ..., **I. Kang**, et al.

Model Evaluation Failure Analysis Data-Centric AI

Contributed to a multi-center study of **80 biomedical-imaging competitions**, examining successful development strategies and the role of failure-case analysis.

Journal of Hazardous Materials 2023 Paper ↗

Three-dimensional label-free visualization of the interactions of PM2.5 with macrophages and epithelial cells using optical diffraction tomography

W. S. Lee*, **I. Kang***, S. J. Yoon, et al.

3D Imaging Quantitative Imaging Biomedical Imaging

Used label-free 3D imaging to quantify cellular $PM_{2.5}$ uptake, refractive index and mass changes.

MICCAI MOOD Challenge 2022 Paper ↗

Joint embedding of 2D and 3D networks for medical image anomaly detection

I. Kang, J. Park

Anomaly Detection Representation Learning 3D Vision

Joint 2D/3D feature learning for brain MRI and abdominal CT; **1st place in the sample-level task**.

KTCP 2022 Paper ↗

End-to-End Vertebra CT Image Segmentation Network with the 3D Surface-Enhanced Module and the Trainable Preprocessing Method

I. Kang, J.-H. Cho, J. Park

3D Vision Segmentation Medical AI

Integrated trainable preprocessing and surface-enhanced features for end-to-end vertebra segmentation.

MICCAI MOOD Challenge 2021 Paper ↗

Self-supervised 3D out-of-distribution detection via pseudoanomaly generation

J.-H. Cho*, **I. Kang***, J. Park

Self-Supervised Learning Synthetic Data OOD Detection

Learned anomaly detection from normal data using synthetic pseudoanomalies; **1st place**.

05 Recognition & research service

Qualcomm Innovation Fellowship Korea – Finalist 2026
QIFK 2026; selected October 29

NeurIPS – Top Reviewer 2026
Main-track reviewing service; recognized October 11

CVPR Compute Transparency Champion Award 2026

MICCAI MOOD Challenge – 1st Place 2021, 2022
2021 challenge; 2022 sample-level task

Best Presentation Award 2022
KTCP

Best Thesis Award 2020
GIST

06 Earlier research & engineering

Machine Learning and Vision Lab Research Intern 2019 – 2020
Gwangju, Korea
Studied scheduled training for long-tailed object detection in autonomous driving.

Aerospace Propulsion Lab Research Intern 2017 – 2018
Gwangju, Korea
Worked on parallel computing and GPU acceleration.

SketchScope: Interactive Volume Visualization ↗ Mar 2021 – Aug 2022
Developed sketch-based volume exploration with hand-gesture and touch interaction across Leap Motion, VR and holographic displays. Extended the platform to HoloLens 2 with QR-code-based volume placement.

ForMixMatch: Virtual Fashion Coordinator ↗ Sep – Dec 2020
Built a VITON-based service for combining upper- and lower-body apparel. Compared four interface layouts in a user study to assess mix-and-match interaction and user confidence.

AI Health Trainer ↗ Apr – Dec 2018
Built real-time mobile posture correction using OpenPose and Realtime PoseNet. Led the GIST representative team for JETS and presented the project at NAVER.

GPU Acceleration with Kokkos ↗ Dec 2017 – Dec 2018
Implemented parallel workloads in C++ with Kokkos; explored architecture-aware memory access, portable execution patterns and GPU work mapping.

07 Technical expertise

Multimodal learning & model analysis
Vision-language grounding; layer-wise representation probing; attention analysis; parameter-efficient adaptation (LoRA); 3D perception; self-supervised learning.

LLM & agentic systems
Mid-training data audits; SFT data curation; multi-turn tool-use trajectories; schema validation; deterministic replay; benchmark and failure analysis.

Physical prediction & learning
Task-specific world models; physics-guided learning; uncertainty estimation; decision-aware reward design; reinforcement learning (GRPO).

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

08 Invited talk & patent

Invited Speaker, Korean Society of Medical Imaging and Information (KOSMII), 2022.

Data Collection Apparatus and Method. Korean Intellectual Property Office; Publication No. 10-2025-0031529.