

Inha Kang

Deep Learning Researcher

✉ rkswlsj13@kaist.ac.kr ☎ +82 10 6824 7177 🔗 Google Scholar 🔗 Portfolio [in LinkedIn](#) [GitHub](#)

Research Interests

Keywords: Computer Vision, Vision-Language Models (VLM), Representation Learning, Multimodal Learning

Compositional Vision-Language Understanding: Researching VLMs capable of understanding complex, compositional language for robust real-world applications.

VLM Explainability & Data-centric AI: Enhancing the trustworthiness and reliability of VLMs by addressing their critical limitations through data-centric AI and reasoning methodologies.

Education

Ph.D. KAIST AI, Seoul, Korea Aug 2024 – Present

- Computer Vision and Machine Learning Lab (Advisor: Prof. Hyunjung Shim)
- **Research Focus:** Vision-Language Models, Multimodal Learning
- GPA: 4.3 / 4.3 (Percentile: 100 / 100)

M.S. KAIST School of Computing, Daejeon, Korea Aug 2020 – Aug 2022

- Computer Graphics and Visualization Lab (Advisor: Prof. Jinah Park)
- **Research Focus:** 3D Medical Imaging, AR/VR
- 2021, 2022 MICCAI MOOD challenge 1st place
- GPA: 3.88 / 4.3 (Percentile: 95.33 / 100)

Visit University of California, Berkeley, Computer Science 2017

B.S. GIST Electronic Engineering and Computer Science, Gwangju, Korea Mar 2016 – Aug 2020

- GPA: 3.94 / 4.5 (Graduated Salutatorian)
- Awards: Received the Best Thesis Award
- Minor in Business Administration and Sociology

Experience

FindersAI, Deep Learning Researcher Seoul, Korea
2022 – 2024

- **Automated Data Engine:** Built a data engine to generate high-quality, error-free segmentation datasets through a novel feature-based filtering pipeline with VLM.
- **Continual Learning & Multi-modal Fusion:** Researched retraining-free model update strategies and developed a tabular model to fuse diverse sensor data for final inference.
- **Real-time Change Detection:** Engineered a system for retail analytics, achieving 2x competitor accuracy.

Machine Learning and Vision Lab (MLV), Research Intern Gwangju, Korea
2019 – 2020

- Designed a scheduled training method for robust long-tailed object detection in autonomous driving.

Aerospace Propulsion Lab (APL), Research Intern Gwangju, Korea
2017 – 2018

- Studied parallel computing and GPU acceleration.

Publications

What “Not” to Detect: Negation-Aware VLMs via Structured Reasoning and Token Merging → [link](#)

ICLR 2026

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

- **Keywords:** Vision-Language Models (VLM), Negation Understanding, Token Merging, Reasoning, Data-centric AI
- **Enhancing VLM Reliability:** Identified and solved the critical “affirmative bias” in VLMs where models ignore negation. Proposed a novel token merging module and a reasoning-based data pipeline, improving detection accuracy by 10.8 points on negation benchmarks.

Real-Time Long Horizon Air Quality Forecasting via Group-Relative Policy Optimization → [link](#)

CVPR 2026

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

- **Keywords:** Air Quality Forecasting, Policy Optimization (GRPO), Foundation Model Fine-Tuning, Real-world Application
- **Reliable AI for Public Health:** Addressed the critical “false alarm” issue in existing global models by applying Reinforcement Learning (GRPO) with asymmetric rewards. Developed a novel regional dataset to enable reliable 5-day real-time forecasts for East Asia, reducing false alarm rates by 47.3%.

No Thing, Nothing: Highlighting Safety-Critical Classes for Robust LiDAR Semantic Segmentation in Adverse Weather → [link](#)

CVPR 2025

J Park, H Lee, I Kang, H Shim

- **Keywords:** LiDAR Semantic Segmentation, Autonomous Driving Safety, Domain Generalization, Robust Perception
- **Robust Perception for Autonomous Driving:** Addressed the safety risks where LiDAR sensors fail to distinguish obstacles from weather noise. Proposed a physics-inspired augmentation strategy that emphasizes object geometry, achieving SoTA performance in detecting safety-critical classes under adverse weather conditions.

3D-Aware Vision-Language Models Fine-Tuning with Geometric Distillation → [link](#)

EMNLP Findings 2025

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

- **Keywords:** Vision-Language Models (VLM), 3D Spatial Understanding, Knowledge Distillation, Zero-Shot Learning
- **Bridging the Gap between 2D and 3D:** Overcame the limitation of existing 2D-based VLMs in understanding 3D geometry and viewpoints. Proposed a novel framework that transfers structural knowledge from 3D models to 2D VLMs, significantly enhancing zero-shot 3D classification performance without massive 3D datasets.

CoPatch: Zero-Shot Referring Image Segmentation by Leveraging Untapped Spatial Knowledge in CLIP → [link](#)

Under Review (2025)

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

- **Keywords:** Zero-Shot Referring Image Segmentation, Vision-Language Models, Training-Free, Spatial Reasoning
- **Unlocking Hidden Spatial Knowledge in CLIP:** Addressed the limitation where existing CLIP-based methods fail to capture complete object shapes due to noisy text-image attention. Proposed a novel training-free framework that leverages intrinsic visual spatial correlations within CLIP, achieving SoTA zero-shot performance on RefCOCO benchmarks.

Why is the winner the best? → [link](#)

CVPR 2023

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

- **Keywords:** Biomedical Image Analysis, Challenge Evaluation, Ranking Stability, Statistical Significance, Meta-Research
- **Ensuring Reliability in AI Benchmarking:** Conducted a large-scale multi-center study to evaluate the validity of rankings in biomedical challenges. Revealed that "winning" algorithms are often not statistically distinguishable from runners-up due to unstable metrics. Provided guidelines for robust and scientifically valid model comparison.

Three-dimensional label-free visualization of the interactions of PM2.5 with macrophages and epithelial cells using optical diffraction tomography → [link](#)

WS Lee*, *I Kang**, SJ Yoon, et al.

- **Keywords:** Optical Diffraction Tomography (ODT), 3D Bio-imaging, Label-free Visualization, Quantitative Phase Imaging
- **Pioneering 3D Label-free Analysis of Nanotoxicity:** Overcame limitations of traditional 2D staining methods (e.g., phototoxicity, lack of depth info) in analyzing fine dust interactions. Applied 3D Optical Diffraction Tomography (ODT) to visualize the engulfment of PM2.5 by immune cells in real-time, enabling precise quantitative analysis of refractive index (RI) and cellular mass changes without labeling.

Joint embedding of 2D and 3D networks for medical image anomaly detection (MOOD Challenge 1st Place) → [link](#)

MICCAI 2022

I Kang, J Park

- **Keywords:** Unsupervised Anomaly Detection, 2D/3D Feature Fusion, Medical Out-of-Distribution (OOD), Challenge Winner
- **Winning Solution for Medical Anomaly Detection:** Achieved **1st Place** in the MICCAI 2022 MOOD Challenge (Sample-level). Addressed the trade-off between local texture details and global volumetric context by jointly embedding 2D and 3D features, successfully identifying subtle anomalies in Brain MRI and Abdominal CT scans where single-view models failed.

End-to-End Vertebra CT Image Segmentation Network with the 3D Surface-Enhanced Module and the Trainable Preprocessing Method → [link](#)

KTCP 2022

I Kang, JH Cho, J Park

- **Keywords:** End-to-End Segmentation, Trainable Preprocessing, 3D Surface Enhancement, Medical CT Analysis
- **Automating Medical Image Pipelines:** Developed an end-to-end network that integrates a "Trainable Preprocessing Module" and a "3D Surface-Enhanced Module," allowing the model to learn optimal image enhancement and boundary detection strategies autonomously.

Self-supervised 3D out-of-distribution detection via pseudoanomaly generation (MOOD Challenge 1st Place) → [link](#)

MICCAI 2021

JH Cho*, *I Kang**, J Park

- **Keywords:** Self-supervised Learning, 3D Anomaly Detection, Pseudo-anomaly Generation, Medical Image Analysis, Challenge Winner
- **Unsupervised 3D Anomaly Detection:** Achieved **1st Place** in the MICCAI 2021 MOOD Challenge. Overcame the scarcity of labeled anomaly data by devising a novel "Pseudo-anomaly Generation" technique. This allowed the model to learn robust defect detection patterns from normal data alone, without requiring expensive medical labeling.

Honors & Awards

- MICCAI Medical Out-Of-Distribution (MOOD) Challenge 1st Place** 2021, 2022
- Won 1st place in the challenge for two consecutive years.
- Best Presentation Award** 2022
- KIISE Transactions on Computing Practices (KTCP)
- Best Thesis Award** 2020
- Awarded upon graduation from GIST.

Invited Talks & Patents

- Invited Speaker** 2022
- Korean Society of Medical Imaging and Information (KOSMII)
- Patents** Aug 2023
- “Data Collection Apparatus and Method”, Korean Intellectual Property Office (KIPO), Pub. No. 10-2025-0031529

Projects

- ODT-based Visualization of PM2.5 Interaction with Cells** Feb 2021 – Dec 2022
- Investigated the cellular interaction of fine dust (PM2.5) using Optical Diffraction Tomography (ODT) and holographic microscopy for 3D label-free visualization.
 - Analyzed Refractive Index (RI) distributions to distinguish uptake patterns between macrophages (active) and epithelial cells (passive) in real-time.
 - Demonstrated ODT as a powerful tool for tracking hard-to-label substances and analyzing biological distribution without phototoxicity.
- SketchScope: Interactive Virtual Science Museum Platform (VR/AR)** cgv.kaist.ac.kr/smc
Mar 2021 – Aug 2022
- Developed “SketchScope,” a multi-modal volume visualization platform allowing users to inspect internal structures via sketching widgets without damaging the object.
 - Implemented bi-directional interactions (hand gestures, touch) across diverse devices including Leap Motion, VR headsets (Oculus), and Holographic Displays.
 - Extended the system to Augmented Reality (HoloLens 2) with QR-code based volume placement, enabling immersive educational experiences for science museums.
- ForMixMatch: Virtual Fashion Coordinator Service** [Project](#)
Sept 2020 – Dec 2020
- Developed an image-based Virtual Try-On (VITON) application to simulate various top and bottom apparel combinations for online shopping.
 - Conducted a user study comparing four different UI layouts (1 vs. n-by-6 vs. n-by-m vs. 6-by-6) to optimize the coordination interface.
 - Verified that adjustable “n by m” mix-match formats significantly improved user confidence in fashion choices compared to single-item views.
- AI Health Trainer** [Project](#)
Apr 2018 – Dec 2018
- Built a mobile fitness application using OpenPose and Realtime PoseNet for real-time posture correction without expensive motion-capture equipment.
 - Implemented personalized features including workout routine generation, dietary recommendations, and a user ranking system to mimic a personal training experience.

- Selected as the university representative for the JETS competition and presented the project at NAVER as a team lead.

GPU Acceleration with KOKKOS

- Studied the C++ Performance Portability Programming Model (KOKKOS) to optimize memory access patterns via architecture-aware arrays.
- Implemented parallel computing tasks using Kokkos to simplify complex performance-optimizing patterns while maintaining advanced features.
- Focused on GPU acceleration and efficient work mapping for high-performance computing tasks.

[Project Presentation](#) 

Dec 2017 – Dec 2018