Jaeho Lee

Homepage | Scholar | jeffy1234@kaist.ac.kr | LinkedIn

Research Interests

My research focuses on how multimodal large language models (MLLMs) understand and reason about the visual world. Recently, I have explored MLLM-based agentic frameworks for tasks such as referring video object segmentation (RVOS) and editable graphic design. I am particularly interested in improving the visual reasoning capabilities of MLLMs, designing effective task-specific agentic pipelines, and developing memory management strategies for agentic systems. Ultimately, I aim to apply these advances to real-world applications, particularly real-time video understanding and embodied agents.

Education

Korea Advanced Institute of Science and Technology (KAIST) Sep. 2025 – Present

M.S/Ph.D Integrated Student in Artificial Intelligence

Advisor: Prof. Seungryong Kim

Yonsei University Mar. 2019 – Aug. 2025

B.S. in Mechanical Engineering

Leave of absence for military service: Jul. 2020 – Apr. 2022

Publications

When to Answer Is Not in the Question: Evidence-Conditioned Answerability for Streaming Video Understanding

Junghyun Park, **Jaeho Lee**, Sangbeom Lim, Junhwan Heo, Woojeong Jin, Eunju Yang, Seungho Jang, Seokju Cho, WonJun Moon, Seungryong Kim

CVPR 2027, under review.

ReAnchor: Reusing Rejected Verifier States for Diffusion Drafting

Sanghyun Lee*, Wonjin Nam*, Seungho Jang, **Jaeho Lee**, Jiho Park, Jaewoon Lee, Seungryong Kim

NAACL 2027, under review [arXiv].

PlayGround: Progressive Layout Generation with Render-Grounded Planning

Jaeho Lee, Junhwan Heo, Junghyun Park, Wonjun Moon, Eunju Yang, Seungho Jang, Seongchan Kim, Seungryong Kim

ICLR 2027, under review [arXiv].

AgentRVOS: Reasoning over Object Tracks for Zero-Shot Referring Video Object Segmentation

Woojeong Jin*, **Jaeho Lee***, Heeseong Shin, Seungho Jang, Junhwan Heo, Seungryong Kim

AAAI 2027, under review [arXiv].

InterRVOS: Interaction-aware Referring Video Object Segmentation

Woojeong Jin, Seongchan Kim, **Jaeho Lee**, Seungryong Kim

CVPR 2026 [arXiv].

MATRIX: Mask Track Alignment for Interaction-aware Video Generation

Siyoon Jin, Seongchan Kim, Dahyun Chung, **Jaeho Lee**, Hyunwook Choi, Jisu Nam, Jiyoung Kim, Seungryong Kim

ICLR 2026 [arXiv].

* Equal contribution.

Academic Services

Reviewer, International Conference on Learning Representations (ICLR 2027).

Honors & Scholarship

Academic Excellence Scholarship, Yonsei University (Aug. 2023; Dec. 2023; Jun. 2024).

Skills

Programming	Python, PyTorch
Research Experience	Dataset Curation & Annotation Pipeline, Benchmark Construction
Languages	Korean (Native), English (Fluent)