Dongjin **Kim**

AI Researcher · Integrated M.S.–Ph.D. Student
KAIST AI, DAVIAN Lab, Seoul, South Korea
+82-10-2128-0852 | dj_kim@kaist.ac.kr | dongjin-kim-2447952a0

Education

Korea Advanced Institute of Science and Technology (KAIST) — *Daejeon, South Korea*
Integrated M.S.–Ph.D. in Artificial Intelligence — *Mar. 2025 – Present*
- DAVIAN Lab, advised by Prof. Jaegul Choo
- Research focus: Vision-Language-Action (VLA), LLM Safety, multi-modal learning

Kyung Hee University — *Suwon, South Korea*
B.S. in Computer Science & Engineering — *Mar. 2019 – Feb. 2025*
- Undergraduate Researcher, Visual AI Lab, advised by Prof. Jung Uk Kim
- Research focus: multi-modal learning, video understanding, sound source localization
- Published 4 papers at top venues (CVPR 2024, ACM MM 2023, CVPR 2025, AAAI 2025)

Research Interests

Primary Vision-Language-Action (VLA), LLM Safety
Secondary Multi-modal learning, Video understanding, Sound source localization

Publications

2026

ExpGuard: LLM Content Moderation in Specialized Domains — *ICLR 2026*
Minseok Choi*, **Dongjin Kim***, Seungbin Yang, Subin Kim, Youngjun Kwak, Juyoung Oh, Jaegul Choo, Jungmin Son (*equal contribution) — *2026*
- Tags: LLM Safety, Content Moderation, Specialized Domains
- arXiv OpenReview Code

LiveWeb-IE: A Benchmark For Online Web Information Extraction — *ICLR 2026*
Seungbin Yang, Jihwan Kim, Jaemin Choi, **Dongjin Kim**, Soyoung Yang, ChaeHun Park, Jaegul Choo — *2026*
- Tags: Benchmark, Information Extraction, Web
- arXiv OpenReview

2025

Object-aware Sound Source Localization via Audio-Visual Scene Understanding — *CVPR 2025*
Sung Jin Um*, **Dongjin Kim***, Sangmin Lee, Jung Uk Kim (*equal contribution) — *2025*
- Tags: Multi-modal, Video Understanding, Sound Source Localization
- arXiv Code

Watch Video, Catch Keyword: Context-aware Keyword Attention for Moment Retrieval and Highlight Detection — *AAAI 2025*
Sung Jin Um, **Dongjin Kim**, Sangmin Lee, Jung Uk Kim — *2025*
- Tags: Multi-modal, Video Understanding, Moment Retrieval, Highlight Detection
- arXiv Code

2024

Learning to Visually Localize Sound Sources from Mixtures without Prior Source Knowledge *CVPR 2024*

Dongjin Kim*, Sung Jin Um*, Sangmin Lee, Jung Uk Kim (*equal contribution) *2024*

- Tags: Multi-modal, Video Understanding, Sound Source Localization
- arXiv Code

2023

Audio-Visual Spatial Integration and Recursive Attention for Robust Sound Source Localization *ACM MM 2023*

Sung Jin Um*, **Dongjin Kim***, Jung Uk Kim (*equal contribution) *2023*

- Tags: Multi-modal, Video Understanding, Sound Source Localization
- arXiv Code

Experience

DAVIAN Lab, KAIST AI *Daejeon, South Korea*

Integrated M.S.–Ph.D. Student *Mar. 2025 – Present*

- Research on VLA, LLM Safety, and multi-modal learning
- Advised by Prof. Jaegul Choo

LETSUR (AI Startup) *Seoul, South Korea*

AI Engineer *Mar. 2024 – Aug. 2024*

- Development of AI-based services utilizing LLMs and RAG

ETRI (Electronics and Telecommunications Research Institute) *Daejeon, South Korea*

AI Research Intern *Jan. 2024 – Feb. 2024*

- Research on Video Moment Retrieval and Highlight Detection

Visual AI Lab, Kyung Hee University *Suwon, South Korea*

Undergraduate Researcher *Nov. 2022 – Feb. 2025*

- Supervised by Prof. Jung Uk Kim
- Published 4 papers at top venues (CVPR 2024, CVPR 2025, AAAI 2025, ACM MM 2023)

Talks & Presentations

Conference Presentations

Object-aware Sound Source Localization via Audio-Visual Scene Understanding *Nashville, TN*

Poster Presentation *CVPR 2025*

- Poster PDF

Learning to Visually Localize Sound Sources from Mixtures without Prior Source Knowledge *Seattle, WA*

Poster Presentation *CVPR 2024*

- Poster PDF

Audio-Visual Spatial Integration and Recursive Attention for Robust Sound Source Localization — *Ottawa, Canada*

Poster Presentation — *ACM MM 2023*

- Poster PDF

Invited Talks

My Experience of CVPR Acceptance as an Undergraduate and Advice for Graduate School — *Seoul, South Korea*

Invited Talk — *2025*

- NexusAI Cohort 2 Final Seminar
- Slides