

JAEYEON KIM

🌐 <https://jaeyeonkim99.github.io> ✉ jaeyeon2@andrew.cmu.edu

Education

Carnegie Mellon University, Language Technologies Institute

Aug. 2025 – Present

PhD Student, School of Computer Science

Advisor: Prof. Carlos Busso and Prof. Shinji Watanabe

Seoul National University, College of Liberal Studies

Mar. 2018 – Aug. 2024

B.S. in Computer Science and Engineering and B.A. in Philosophy

GPA: 4.11/4.30

2-year absence to fulfill military duty (Apr. 2022 - Mar. 2024)

Research Experience

Vision and Learning Lab, Seoul National University

Mar. 2024 – July. 2025

Research Intern, Advisor: Prof. Gunhee Kim

Seoul, South Korea

- Explored audio-multimodal learning.

MAUM AI Inc. (Previously MindsLab Inc.)

Apr. 2022 – Jun. 2024

AI Scientist (Apr. 2022 – Jun. 2023), Senior AI Scientist (Jul. 2023 – Jun. 2024),

Seongnam, South Korea

- Alternative military service as industrial technical personnel (Apr. 2022 - Mar. 2024).
- Developed and maintained automatic speech recognition systems for both Korean and English.
- Led research projects within the audio team, regarding text-to-speech, automatic speech recognition, and other audio-related tasks.

Music and Audio Research Group, Seoul National University

Jul. 2021 – Mar. 2022

Research Intern, Advisor: Prof. Kyogu Lee

Gwanggyo, South Korea

- Explored self-supervised speech representations and spoken language understanding.

Publications (*Equal Contributions)

- [1] **Jaeyeon Kim***, Heeseung Yun*, Gunhee Kim. “Towards Scene-Aware Video-to-Spatial Audio Generation”, *Conditionally Accepted at IJCV*, 2025
- [2] Heeseung Yun, Joonil Na, **Jaeyeon Kim**, Calvin Murdock, Gunhee Kim. “Gaze Beyond the Frame: Forecasting Egocentric 3D Visual Span”, in *NeurIPS (Spotlight)*, 2025
- [3] **Jaeyeon Kim**, Heeseung Yun, Sang Hoon Woo, Chao-Han Huck Yang, Gunhee Kim. “WoW-Bench: Evaluating Fine-Grained Acoustic Perception in Audio-Language Models via Marine Mammal Vocalizations”, *arXiv:2508.20976*, 2025 [Paper] [Project Page]
- [4] Chao-Han Huck Yang, Sreyan Ghosh, Qing Wang, **Jaeyeon Kim**, Hengyi Hong, Sonal Kumar, Guirui Zhong, Zhifeng Kong, S Sakshi, Vaibhavi Lokegaonkar, Oriol Nieto, Ramani Duraiswami, Dinesh Manocha, Gunhee Kim, Jun Du, Rafael Valle, Bryan Catanzaro. “Multi-Domain Audio Question Answering Toward Acoustic Content Reasoning in The DCASE 2025 Challenge”, *arXiv:2505.07365*, 2025 [Paper]
- [5] **Jaeyeon Kim**, Heeseung Yun, Gunhee Kim. “ViSAGE: Video-to-Spatial Audio Generation”, in *ICLR*, 2025 [Paper] [Project Page]
- [6] **Jaeyeon Kim**, Minjeong Jeon, Jaeyoon Jung, Sang Hoon Woo, Jinjoo Lee. “EnCLAP++: Analyzing the EnCLAP Framework for Optimizing Automated Audio Captioning Performance”, in *DCASE2024 Workshop*, 2024 [Paper]
- [7] **Jaeyeon Kim**, Jaeyoon Jung, Sang Hoon Woo, Minjeong Jeon, Jinjoo Lee. “Expanding on EnCLAP with Auxiliary Retrieval Model for Automated Audio Captioning”, *DCASE2024 Challenge Technical Report (Ranked 2nd in Task 6 Automated Audio Captioning)*, 2024 [Paper]
- [8] **Jaeyeon Kim**, Injune Hwang, Kyogu Lee. “Learning Semantic Information from Raw Audio Signal Using Both Contextual and Phonetic Representations”, in *ICASSP (Oral)*, 2024 [Paper]
- [9] **Jaeyeon Kim**, Jaeyoon Jung, Jinjoo Lee, Sang Hoon Woo. “EnCLAP: Combining Neural Audio Codec and Audio-Text Joint Embedding for Automated Audio Captioning”, in *ICASSP*, 2024 [Paper] [Project Page]
- [10] Junhyeok Lee, Wonbin Jung, Hyunjae Cho, **Jaeyeon Kim**, Jaehwan Kim. “PITS: Variational Pitch Inference Without Fundamental Frequency for End-to-End Pitch-Controllable TTS”, *ICML Workshop on SPIGM*, 2023 [Paper] [Project Page]

Scholarships and Awards

Summa Cum Laude <i>College of Liberal Studies, Seoul National University</i>	Aug. 2024
Best Thesis Award <i>Department of Computer Science and Engineering, Seoul National University</i> <i>With paper "EnCLAP: Utilizing Both Sequence-Level Representations and Time Step-Level Representations for Automated Audio Captioning"</i>	Aug. 2024
Bronze Award in College of Liberal Studies Academic Conference <i>College of Liberal Studies, Seoul National University</i> <i>With paper "Situated Intelligence and Practical Knowledge of Artificial Intelligence"</i>	Nov. 2021
Scholarship of 100 years of Humanities <i>Korea Student Aid Foundation</i> <i>4-year full-tuition scholarship with living expense support</i>	Mar. 2018 – Feb. 2022
Talent Award of Korea <i>Minister of Education</i> <i>To 50 outstanding high-school students nationwide recognized as future leaders across diverse fields.</i>	Dec. 2017

Academic Service

Reviewer

- ICASSP (2026)

Organizer

- DCASE2025 Challenge Task5 Audio Question Answering [\[Link\]](#)

Projects

Automatic Speech Recognition Based on Conformer Transducer <i>Led the development of new Korean and English automatic speech recognition systems at MAUM AI Inc.</i> <i>Key responsibilities and achievements include:</i> <ul style="list-style-type: none">• Implemented modeling and training codes, and scalable serving system based on gRPC• Developed streaming recognition API using voice activity detection and asyncio.• Progressively improved encoder architectures and decoding algorithms based on latest research findings.• Curated Korean datasets and conducted thorough training for real-world deployment.• Achieved 40% relative improvements of character error rate compared to the previous system.• Achieved 50% and 75% reduction in inference speed and memory consumption compared to the initial version.• Deployed in 4+ commercial services, actively maintained.	Apr. 2022 – Feb. 2024
--	-----------------------

Additional Information

Language. Korean-Native, English-Fluent (TOEFL: 108)

Philosophy. I am interested in engaging in philosophical discussions on AI and deriving insights for AI research from philosophical perspectives. My focus lies particularly within the domains of philosophy of mind, philosophy of language, and Kantian ethics. You can explore my writings on philosophical discussions [\[here\]](#) (in Korean, titles also in English).