

SUNGNYUN KIM

Graduate School of AI, KAIST
Seoul 02455, Republic of Korea

Email: ksn4397@kaist.ac.kr
Website: <https://sungnyun.github.io/>
[Google Scholar](#) [GitHub](#) [LinkedIn](#) [Twitter](#)

RESEARCH INTERESTS

Audio-Visual-Language Multimodality; Spatial/Social Intelligence of MLLMs; Interactive Voice Agents and Full-Duplex Modeling; Efficiency in Data Synthesis and Model Training

EDUCATION

Korea Advanced Institute of Science and Technology (KAIST)	Seoul, South Korea
Ph.D. in Graduate School of Artificial Intelligence; Advisor: Prof. Se-Young Yun	Feb 2026
M.S. in Graduate School of Artificial Intelligence	Aug 2021
B.S. in Chemical and Biomolecular Engineering (<i>Magna Cum Laude</i>)	Sep 2019
B.S. in Industrial and Systems Engineering; double major	Sep 2019

WORK EXPERIENCE

Postdoctoral Fellow	Mar 2026 – Present
<i>InnoCORE-LLM Research Center, KAIST</i> , Daejeon, Republic of Korea	
PI: Prof. Se-Young Yun & Prof. Tae-Hyun Oh, KAIST	
Managing and directing research projects on audio-visual understanding, generation, and evaluation	
Applied Scientist Intern	Jul 2023 – Nov 2023
<i>AWS AI Labs</i> , Pasadena, CA, USA	
Work with Haofu Liao and Srikar Appalaraju on open-world document intelligence with (V)LLMs; EMNLP 2024	
Research Intern	Jan 2023 – Jun 2023
<i>NAVER Webtoon AI</i> , Pangyo, Republic of Korea	
Work with Junsoo Lee on composable, multimodal text-to-image diffusion models; ESWA 2025	

AWARDS AND HONORS

- Postdoctoral Fellowship, InnoCORE-LLM Research Center, 2026
- Outstanding Reviewer (top 5%), CVPR 2026
- 1st Place Award, AI Frontier Challenge, KAIA 2025 (Prize: \$5,000)
- Top Reviewer (top 1.9%), ICML 2025 (Prize: \$900-value registration support)
- Best Student Paper Award (top 5 papers), ICASSP 2024 (Prize: \$1,000)
- Invited Talk, VESSL AI NeurIPS 2022 Preview
- Outstanding Paper Award, Korea Artificial Intelligence Association (KAIA) 2021 (Prize: \$500)
- Dean's List (top 3%), Faculty of Engineering Dept., KAIST, Spring 2017

SELECTED PUBLICATIONS

Full publication list available on [Google Scholar](#). *: equal contribution, †: corresponding authorship.

Topics: **LLM** (Large Language Models) · **MM** (Multimodal) · **SP** (Speech Processing) · **SI** (Spatial/Social Intelligence) · **REP** (Representation Learning)

- | | |
|---------------|---|
| MM; SI | K. Lee, Sungnyun Kim, S. J. Schlecht, T-H. Oh, and V. Välimäki. Perception-Inspired Bayesian Causal Fusion for Audiovisual Source Localization . Preprint, 2026. [pdf] |
| MM; SI | H. Kim, J. Kim, Sungnyun Kim, and T-H. Oh. Towards Reliable Evaluation of 3D Talking Head Emotion with MLLM Judges . Preprint, 2026. |
| MM; SI | K. Kim, Sungnyun Kim [†] , and T-H. Oh [†] . FloorSAV: Elucidating Spatial Audio-Visual Context with 2D Floormap for AV-LLMs . NeurIPS Workshop on Real-Time Multimodal Conversational AI, 2026. |
| SP; SI | Sungnyun Kim [†] , S. Cho, J. Oh, and S-Y. Yun. Conversation Is a Two-Body Problem: Dyadic Evaluation of Full-Duplex Dialogue Models . NeurIPS Workshop on Real-Time Multimodal Conversational AI, 2026. |
| MM | J. Ko*, Sungnyun Kim*, S. Cho, and S-Y. Yun. Flex-Judge: Text-Only Reasoning Unleashes Zero-Shot Multimodal Evaluators . Advances in Neural Information Processing Systems (NeurIPS), 2025. [pdf] [code] [project] |
| MM | Sungnyun Kim, J. Lee, K. Hong, D. Kim, and N. Ahn. DiffBlender: Composable and Versatile Multimodal Text-to-Image Diffusion Models . Expert Systems with Applications (ESWA), vol. 297, 2025. [pdf] [code] [project] |

- MM; SP** Sungnyun Kim, K. Jang, S. Bae, S. Cho, and S-Y. Yun. **MoHAVE: Mixture of Hierarchical Audio-Visual Experts for Robust Speech Recognition**. Proceedings of the 42nd International Conference on Machine Learning (ICML), 2025. [\[pdf\]](#)
- MM; SP** Sungnyun Kim, S. Cho, S. Bae, K. Jang, and S-Y. Yun. **Multi-Task Corrupted Prediction for Learning Robust Audio-Visual Speech Representation**. International Conference on Learning Representations (ICLR), 2025. [\[pdf\]](#) [\[code\]](#)
- MM; SP** Sungnyun Kim*, K. Jang*, S. Bae, H. Kim, and S-Y. Yun. **Learning Video Temporal Dynamics with Cross-Modal Attention for Robust Audio-Visual Speech Recognition**. IEEE Spoken Language Technology Workshop (SLT), 2024. [\[pdf\]](#) [\[code\]](#)
- LLM** S. Bae*, Y. Kim*, R. Bayat*, Sungnyun Kim, T. Schuster, A. Fisch, H. Harutyunyan, Z. Ji, J. Ha, A. Courville, and S-Y. Yun. **Mixture-of-RecurSIONs: Learning Dynamic Recursive Depths for Adaptive Token-Level Computation**. Advances in Neural Information Processing Systems (NeurIPS), 2025. [\[pdf\]](#) [\[code\]](#)
- LLM** J. Ko, T. Chen, Sungnyun Kim, T. Ding, L. Liang, I. Zharkov, and S-Y. Yun. **DistiLLM-2: A Contrastive Approach Boosts the Distillation of LLMs**. Proceedings of the 42nd International Conference on Machine Learning (ICML), 2025. ([Oral Presentation](#)) [\[pdf\]](#) [\[code\]](#)
- LLM** Sungnyun Kim*[†], H. Liao*, S. Appalaraju, P. Tang, Z. Tu, R. K. Satzoda, R. Manmatha, V. Mahadevan, and S. Soatto. **DockD: Knowledge Distillation from LLMs for Open-World Document Understanding Models**. Conference on Empirical Methods in Natural Language Processing (EMNLP), 2024. [\[pdf\]](#) [\[video\]](#)
- LLM** J. Ko, Sungnyun Kim, T. Chen, and S-Y. Yun. **DistiLLM: Towards Streamlined Distillation for Large Language Models**. Proceedings of the 41st International Conference on Machine Learning (ICML), 2024. [\[pdf\]](#) [\[code\]](#)
- SP; REP** K. Jang, Sungnyun Kim, and H. Kim. **STaR: Distilling Speech Temporal Relation for Lightweight Speech Self-Supervised Learning Models**. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2024. ([Best Student Paper](#)) [\[pdf\]](#) [\[code\]](#)
- SP; REP** K. Jang*, Sungnyun Kim*, S-Y. Yun, and H. Kim. **Recycle-and-Distill: Universal Compression Strategy for Transformer-based Speech SSL Models with Attention Map Reusing and Masking Distillation**. Proceedings of Interspeech, 2023. [\[pdf\]](#) [\[code\]](#)
- REP** S. Bae*, J-W. Kim*, W. Cho, H. Baek, S. Son, B. Lee, C. Ha, K. Tae, Sungnyun Kim[†], and S-Y. Yun[†]. **Patch-Mix Contrastive Learning with Audio Spectrogram Transformer on Respiratory Sound Classification**. Proceedings of Interspeech, 2023. [\[pdf\]](#) [\[code\]](#)
- REP** Sungnyun Kim*, S. Bae*, and S-Y. Yun. **Coreset Sampling from Open-Set for Fine-Grained Self-Supervised Learning**. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023. [\[pdf\]](#) [\[code\]](#) [\[video\]](#)
- REP** J. Oh*, Sungnyun Kim*, N. Ho*, J-H. Kim, H. Song, and S-Y. Yun. **Understanding Cross-Domain Few-Shot Learning Based on Domain Similarity and Few-Shot Difficulty**. Advances in Neural Information Processing Systems (NeurIPS), 2022. [\[pdf\]](#) [\[code\]](#)
- REP** Sungnyun Kim*, G. Lee*, S. Bae*, and S-Y. Yun. **MixCo: Mix-up Contrastive Learning for Visual Representation**. NeurIPS 1st Workshop on Self-Supervised Learning, 2020. [\[pdf\]](#) [\[code\]](#)

RESEARCH PROJECTS

-
- Audio-Visual Talking Avatar Generation and Evaluation** Mar 2026 – Present
NC AI
- Leading a research group developing and evaluating audio-driven talking-avatar generation models
- Distributed Learning for AI Collaboration** Mar 2024 – Dec 2025
Electronics and Telecommunications Research Institute (ETRI)
- Led a research group developing collaborative federated learning algorithms
- Infectious Disease Spread Modeling and Detection** Jul 2021 – Dec 2022
Institute for Security Convergence Research (ISCR)
- Led a research group building an epidemic detection system and unifying it with a spread forecasting model into a single prevention framework
 - Published at ICML 2022 Workshop on Healthcare AI and COVID-19
- Data-efficient Unsupervised Representation Learning** Mar 2020 – Sep 2021
SK Hynix
- Developed a novel self-supervised contrastive learning algorithm leveraging Mix-up for data-efficient representation learning

- Published at NeurIPS 2020 Workshop on Self-Supervised Learning and AAAI 2023

Uncertainty Modeling and Calibration for Safe AI

Apr 2020 – Dec 2020

Venture Research Program for Graduate and PhD students, KAIST

- Developed an uncertainty-aware semi-supervised semantic segmentation model for safety-critical applications

ACADEMIC ACTIVITIES

Reviewer and Chair Service

- Serving as Area Chair at ICLR 2027
- Serving as a peer reviewer in the following conferences and journals: NeurIPS, ICML, ICLR, CVPR, ECCV, ICCV, EMNLP, AAAI, Interspeech, IEEE TNNLS, IEEE TIP, IEEE Access, ACM Computing Surveys, IEEE/ACM TASLP

Advisory Collaboration with Modulabs

- Proposed a novel patchified self-supervised learning algorithm for respiratory sound representation
- Corresponding-authored paper published at Interspeech 2023

Teaching Assistant

- KAIST AI502 Deep Learning Spring 2020
- KAIST AI505 Optimization for AI Fall 2020, Fall 2021, Fall 2022
- KAIST AI603 Machine Learning Theory Spring 2021, Spring 2022
- LG Vision Intelligence Course Fall 2021
- LG-KAIST Advanced AI Course Fall 2020