

Jeong Hun Yeo

Postdoctoral Researcher, School of Electrical Engineering, KAIST

jh.y@ieee.org | [Google Scholar](#) | [LinkedIn](#) | [ORCID](#)

EXPERIENCE

Postdoctoral Researcher, KAIST

Mar. 2026 – Present

Integrated Vision Language Lab, School of Electrical Engineering

– Jang Young Sil Postdoctoral Fellow, focusing on multimodal AI

EDUCATION

Ph.D. in Electrical Engineering, KAIST

Sep. 2022 – Feb. 2026

Advisor: Prof. Yong Man Ro

M.S. in Electrical Engineering, KAIST

Sep. 2020 – Aug. 2022

Advisor: Prof. Yong Man Ro

B.S. in Electrical Engineering, KAIST

Feb. 2014 – Aug. 2020

PUBLICATIONS

* denotes equal contribution.

2026

- [1] **Diffusion Large Language Models for Visual Speech Recognition** EMNLP 2026
Jeong Hun Yeo, Chae Won Kim, Hyeongseop Rha, Yong Man Ro
- [2] **Decoding Strategies for Diffusion-Based ASR: A Systematic Evaluation of Confidence-Based Thresholding** EMNLP 2026
Jeong Hun Yeo, Minsu Kim, Hyeongseop Rha, Yong Man Ro
- [3] **GCAgent: Long-Video Understanding via Schematic and Narrative Episodic Memory** IEEE TMM 2026
Jeong Hun Yeo*, Sangyun Chung*, Sungjune Park, Dae Hoe Kim, Jinyoung Moon, Yong Man Ro
- [4] **Emotion-Coherent Reasoning for Multimodal LLMs via Emotional Rationale Verifier** AAAI 2026
Hyeongseop Rha, Jeong Hun Yeo, Yeonju Kim, Yong Man Ro

2025

- [5] **Zero-AVSR: Zero-Shot Audio-Visual Speech Recognition with LLMs by Learning Language-Agnostic Speech Representations** ICCV 2025
Jeong Hun Yeo*, Minsu Kim*, Chae Won Kim, Stavros Petridis, Yong Man Ro
- [6] **MMS-LLaMA: Efficient LLM-based Audio-Visual Speech Recognition with Minimal Multimodal Speech Tokens** Findings of ACL 2025
Jeong Hun Yeo*, Hyeongseop Rha*, Se Jin Park, Yong Man Ro
- [7] **Personalized Lip Reading: Adapting to Your Unique Lip Movements with Vision and Language** AAAI 2025
Jeong Hun Yeo, Chae Won Kim, Hyunjun Kim, Hyeongseop Rha, Seunghee Han, Wen-Huang Cheng, Yong Man Ro

2024

- [8] **Where Visual Speech Meets Language: VSP-LLM Framework for Efficient and Context-Aware Visual Speech Processing** Findings of EMNLP 2024
Jeong Hun Yeo*, Seunghee Han*, Minsu Kim, Yong Man Ro
- [9] **Efficient Training for Multilingual Visual Speech Recognition: Pre-training with Discretized Visual Speech Representation** ACM MM 2024
Minsu Kim*, Jeong Hun Yeo*, Se Jin Park, Hyeongseop Rha, Yong Man Ro
- [10] **Let's Go Real Talk: Spoken Dialogue Model for Face-to-Face Conversation** ACL 2024 (Oral)
Se Jin Park*, Chae Won Kim*, Hyeongseop Rha, Minsu Kim, Joanna Hong, Jeong Hun Yeo, Yong Man Ro

- [11] **AKVSR: Audio Knowledge Empowered Visual Speech Recognition by Compressing Audio Knowledge of a Pretrained Model** IEEE TMM 2024
Jeong Hun Yeo, Minsu Kim, Jeongsoo Choi, Dae Hoe Kim, Yong Man Ro
- [12] **Visual Speech Recognition for Languages with Limited Labeled Data using Automatic Labels from Whisper** ICASSP 2024 (Oral)
Jeong Hun Yeo*, Minsu Kim*, Shinji Watanabe, Yong Man Ro
- [13] **Towards Practical and Efficient Image-to-Speech Captioning with Vision-Language Pre-training and Multi-modal Tokens** ICASSP 2024
Minsu Kim, Jeongsoo Choi, Soumi Maiti, Jeong Hun Yeo, Shinji Watanabe, Yong Man Ro
- 2023**
- [14] **Lip Reading for Low-resource Languages by Learning and Combining General Speech Knowledge and Language-specific Knowledge** ICCV 2023
Minsu Kim*, Jeong Hun Yeo*, Jeongsoo Choi, Yong Man Ro
- [15] **Multi-Temporal Lip-Audio Memory for Visual Speech Recognition** ICASSP 2023
Jeong Hun Yeo, Minsu Kim, Yong Man Ro
- 2022**
- [16] **Distinguishing Homophenes Using Multi-head Visual-Audio Memory for Lip Reading** AAAI 2022
Minsu Kim, Jeong Hun Yeo, Yong Man Ro

PREPRINTS

- [17] **Learning What to Attend First: Modality-Importance-Guided Reasoning for Reliable Multimodal Emotion Understanding** arXiv:2512.02699
Hyeongseop Rha*, Jeong Hun Yeo*, Se Jin Park, Yong Man Ro
- [18] **Towards Inclusive Communication: A Unified Framework for Generating Spoken Language from Sign, Lip, and Audio** arXiv:2508.20476
Jeong Hun Yeo, Hyeongseop Rha, Sungjune Park, Junil Won, Yong Man Ro

AWARDS & HONORS

Jang Young Sil Postdoctoral Fellowship, KAIST 2026
KAIST Outstanding Paper Award for Graduate Students 2024

PROFESSIONAL SERVICE

Journal Reviewer

- IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)
- International Journal of Computer Vision (IJCV)
- IEEE Transactions on Image Processing (TIP)
- IEEE Transactions on Multimedia (TMM)
- IEEE/ACM Transactions on Audio, Speech, and Language Processing (TASLP)

Conference Reviewer

- IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) 2026
- IEEE/CVF International Conference on Computer Vision (ICCV) 2025
- European Conference on Computer Vision (ECCV) 2026
- Conference on Neural Information Processing Systems (NeurIPS) 2026
- AAAI Conference on Artificial Intelligence (AAAI) 2026
- ACL Rolling Review (ARR) 2025, 2026
- IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2025, 2026
- IEEE International Conference on Image Processing (ICIP) 2024, 2025

Invited Talk

- ETRI, “Efficient LLM-Based Audio-Visual Speech Recognition” Jun. 2025

Teaching

- Teaching Assistant, *Multimedia Processing and Learning*, KAIST Fall 2023