Junseok Ahn

+82-10-6671-6505 | junseok@mmai.io | junseok520.github.io

Google Scholar | LinkedIn | GitHub |

Daejeon - 34184, South Korea

Objective

A passionate and driven AI researcher seeking a challenging Research Scientist or Applied Scientist position. Aiming to leverage deep expertise in multimodal deep learning, particularly in generative models for audio-visual synthesis, to contribute to cutting-edge projects in interactive AI, virtual avatars, and human-computer interaction. TEST CV333

Education

- **Korea Advanced Institute of Science and Technology (KAIST)** *Mar 2023 - Present*
  Integrated MS/PhD in Electrical Engineering Daejeon, South Korea
  - Advisor: Joon Son Chung
  - GPA: 3.98/4.3
- **Korea Advanced Institute of Science and Technology (KAIST)** *Mar 2019 - Feb 2023*
  B.S. in Electrical Engineering Daejeon, South Korea
  - Graduated Summa Cum Laude (Second in class, Top 1% of department)
  - GPA: 4.17/4.3

Publications

C=Conference, J=Journal, P=Patent, S=In Submission, T=Thesis

[S.1] Ji-Hoon Kim*, **Junseok Ahn***, Doyeop Kwak, Joon Son Chung, Shinji Watanabe (2025). **TAVID: Text-driven Audio-Visual Interactive Dialogue Generation**. Manuscript submitted for publication in *Proceedings of the 2025 CVPR*.

[J.1] Youngjoon Jang, Jeongsoo Choi, **Junseok Ahn**, Joon Son Chung (2025). **Deep Understanding of Sign Language for Sign to Subtitle Alignment**. Accepted with minor revision in *Transactions on Multimedia*.

[C.4] Jaemin Jung*, **Junseok Ahn***, Chaeyoung Jung, Tan Dat Nguyen, Youngjoon Jang, Joon Son Chung (2025). **VoiceDiT: Dual-Condition Diffusion Transformer for Environment-Aware Speech Synthesis**. In *Proceedings of the 2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE. April 6-11, 2025, Hyderabad, India.

[C.3] **Junseok Ahn**, Youkyum Kim, Yeunju Choi, Doyeop Kwak, Ji-Hoon Kim, Seongkyu Mun, Joon Son Chung (2024). **VoxSim: A perceptual voice similarity dataset**. In *Proceedings of the 2024 Interspeech*, pp. 2580-2584. IEEE. 1-5 September 2024, Kos Island, Greece. (Nominated as Best Student Paper List)

[C.2] Youngjoon Jang*, Ji-Hoon Kim*, **Junseok Ahn**, Doyeop Kwak, Hong-Sun Yang, Yoon-Cheol Ju, Il-Hwan Kim, Byeong-Yeol Kim, Joon Son Chung (2024). **Faces that Speak: Jointly Synthesising Talking Face and Speech from Text**. In *Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 8818-8828. IEEE. 16-22 June 2024, Seattle, WA, USA.

[C.1] **Junseok Ahn***, Youngjoon Jang*, Joon Son Chung (2024). **Slowfast Network for Continuous Sign Language Recognition**. In *Proceedings of the 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3920-3924. IEEE. 14-19 April 2024, Seoul, Republic of Korea.

* denotes equal contribution.

Experience

- **KAIST Multimodal AI Lab [🌐]** *Sep 2021 - Feb 2022*
  Undergraduate Researcher participated in KAIST Undergraduate Research Program (URP) Daejeon, Republic of Korea
  - *Thesis: Development of speaker recognition model in which non-speaker information is separated*
  - Developed a speaker recognition framework using adversarial training (GRL) to disentangle language information from speaker embeddings.
  - Achieved up to 50% speaker recognition EER improvement on a custom bilingual test set (VoxCeleb1-B) over baseline models.
  - Trained a prerequisite ECAPA-TDNN language ID model, achieving 6.7% EER on the VoxLingua107 dataset to generate pseudo-labels.
  - Verified the effective removal of language-specific information from speaker embeddings, enhancing model robustness in cross-lingual scenarios.

- **Samsung Electronics System LSI [🌐]** *Sep 2021 - Feb 2022*
  Student Researcher (Intern) Republic of Korea
  - Developed a 5G NR PUSCH (Uplink) MIMO receiver simulator in MATLAB over a 24-week project.
  - Implemented and benchmarked receiver algorithms (Zero Forcing, LMMSE, ML), analyzing BER performance under various Eb/No conditions.
  - Analyzed and simulated 5G NR PHY layer (3GPP) standards, including OFDM characteristics, PAPR reduction techniques, and DMRS estimation.

Skills

- **Programming Languages:** Python, Pytorch, MATLAB, LaTex
- **Machine Learning:** Hugging Face (Transformers, Diffusers, Accelerate), OpenCV, FFmpeg, Librosa

Honors and Awards

- **Best Teaching Assistant Award** *Mar 2025*
  KAIST EE Department
  - Selected as best teaching assistant in 2024 Fall semester
  - Teaching course: EE488 Deep Learning for Visual Understanding
- **Dean's List** *Sep 2019, Sep 2020, Sep 2021*
  KAIST
  - Awarded Dean's List on Spring 2019, Spring 2020, Spring 2021 semesters

Teaching Experience

- **EE738 Speech Recognition System** *Spring 2023, Spring 2024*
  Teaching Assistant
  - Head TA in Spring 2024
- **EE488 Deep Learning in Visual Understanding** *Fall 2023, Fall 2024*
  Teaching Assistant
  - Head TA in Fall 2024
  - Got Best TA Award in Fall 2024

Additional Information

Languages: Korean (native), English (fluent)

References

1. **Joon Son Chung**
   Associate Professor, Electrical Engineering
   Korea Advanced Institute of Science and Technology (KAIST)
   Email: joonsonchung@gmail.com
   Relationship: PhD advisor at KAIST