Junseok Oh (오준석)

(+82) 010-6640-0875 | ohjs@sogang.ac.kr | https://june-oh.github.io/june-oh/

EDUCATION

Sogang University Seoul, South Korea
Ph.D. Candidate in Computer Science and Engineering (*Advisor: Prof. Ji-Hwan Kim)* *Mar 2022 – Present*

Sogang University Seoul, South Korea
Master of Engineering in Computer Science and Engineering (*Advisor: Prof. Ji-Hwan Kim)* *Sep 2017 – Aug 2019*
- *Thesis: "Korean Real-Time Automatic Transcription System Using Weakly Labeled Corpus"*

Sogang University Seoul, South Korea
Bachelor of Engineering in Computer Science and Engineering *Mar 2010 – Aug 2017*

Uijeongbu High School Uijeongbu, Gyeonggi-do, South Korea
High School Diploma *Mar 2007 – Feb 2010*

RESEARCH EXPERIENCE

Development of Dialog-based Multi-modal Explainable AI Technology *Apr 2022 – Present*
Funding Agency: Ministry of Science and ICT (MSIT) / IITP
- Developed an AI-based framework for dysarthria (speech disorder) severity classification, providing multi-modal explanations to support diagnostic decision-making.
- Implemented speech-based explainable diagnostic modules that analyze acoustic and linguistic features to provide interpretable feedback on severity levels.
- Researched methods to bridge the gap between complex AI model outputs and user-understandable explanations in a multi-modal dialog environment.

End-to-End Korean Speech Recognition Project *2024 – April 2025*
Industry Partner: Smilegate
- Developed a universal Korean ASR system using a hybrid FastConformer RNN-Transducer (RNNT) + CTC model.
- A cache-aware streaming method was implemented to improve recognition accuracy and reduce latency.
- Implemented context biasing to adapt the ASR decoder for domain-specific vocabulary, such as gaming terminology, thereby improving recognition accuracy and robustness in specialized contexts.

End-to-End Speech Recognition for Telephony *Apr 2024 – Dec 2024*
Industry Partner: LOTTE INNOVATE COMPANY.
- Developed streaming and non-streaming Korean ASR pipelines optimized for 8 kHz telephony data using a FastConformer-CTC architecture.
- Implemented context-biasing modules to dynamically adapt the decoder's vocabulary for domain shift, boosting word-level accuracy.

Automated Korean Speaking Assessment (2024) *May 2024 – Dec 2024*

Funding Agency: Ministry of Culture, Sports and Tourism of South Korea

- Developed a multi-task learning framework using Wav2Vec to jointly model pronunciation, fluency, and content for L2-Korean assessment.
- Integrated Conformer-CTC ASR outputs with a large language model (LLM) generate automated, multi-aspect scoring.

Intelligent Analysis and Classification-based Content Rating Technology to Address Uncontrolled Distribution of Harmful Media *2022 – 2024*

Funding Agency: Ministry of Science and ICT of South Korea

- Led the audio analytics submodule within an automated video content rating framework, designing sound-event detection models for key multimedia events.
- Fine-tuned OpenAI's Whisper ASR model on domain-specific video corpora to enhance transcription robustness in diverse acoustic environments.

Automated Korean Speaking Assessment (2023) *May 2023 – Dec 2023*

Funding Agency: Ministry of Culture, Sports and Tourism of South Korea

- Built an end-to-end evaluation pipeline for L2 Korean speakers by combining Conformer-CTC ASR outputs with BERT-based semantic scoring.
- Developed algorithms to quantify pronunciation accuracy, speech rate, and syntactic correctness from ASR transcripts.

Development of a Video Story Understanding-based QA System to Pass the Video Turing Test

Sep *2017 – Dec 2019*

Funding Agency: Ministry of Science and ICT of South Korea

- Modified Kaldi's sentence-level decoder to achieve sub-1.0 RT for real-time video QA applications as part of a four-member research team.
- Collected and curated domain-specific audio/text corpora from target video series to optimize acoustic and language models.
- Conducted acoustic model retraining and language model updates using the new corpora, improving QA accuracy on complex video narratives.
-

PUBLICATIONS

International Journal

J. Oh, J. Nam, and J.-H. Kim, "HiTCA: Fusing Hierarchical Text and Contextual Audio for Accurate VCR," EURASIP Journal on Audio, Speech, and Music Processing, *under review.* [SCIE]

S. Ma, **J. Oh**, M. Kim, and J.-H. Kim, "Survey on Deep Learning-based Speech Technologies in Voice Chatbot Systems," KSII Transactions on Internet & Information Systems (TIIS), vol. 19, no. 5, pp. 1406-1440, 2025. [SCIE].

J. Oh, E. Cho, and J.-H. Kim, "Integration of WFST language model in pre-trained Korean E2E ASR model," KSII Transactions on Internet and Information Systems (TIIS), vol. 18, no. 6, pp. 1692–1705, 2024. [SCIE].

S. Seo, **J. Oh**, E. Cho, H. Park, G. Kim, and J.-H. Kim, "TP-MobNet: A Two-pass Mobile Network for Low-complexity Classification of Acoustic Scene," Computers, Materials & Continua, vol. 73, no. 2, 2022. [SCIE].

M. Lim, D. Lee, H. Park, Y. Kang, **J. Oh**, J.-S. Park, G.-J. Jang, and J.-H. Kim, "Convolutional neural network based audio event classification," KSII Transactions on Internet and Information Systems (TIIS), vol. 12, no. 6, pp. 2748–2760, Jun. 2018. [SCIE]

International Conference

J. Oh, J.-H. Kim, “Adapter-Only Bridging of Frozen Speech Encoder and Frozen LLM for ASR,” in Proc. Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD), 2026. [**Accepted, Oral Presentation**]

J. Oh, and J.-H. Kim, "SEAM: Bridging the Temporal-Semantic Granularity Gap for LLM-based Speech Recognition," in Findings of the Association for Computational Linguistics: EACL 2026, pp. 2135–2144, 2026.

J. Oh, H. Park, and J.-H. Kim, “Speech Intelligibility Prediction of Dysarthria Using Deep Convolutional Networks,” in *Proc.* APIC-IST 2023, pp. 236–237, 2023.

M. Kim, **J. Oh**, and J.-H. Kim, “Automated Dysarthria Severity Classification Using Diadochokinetic test and Speech Intelligibility Based on LightGBM,” in *Proc.* APIC-IST 2023, pp. 12–13, 2023.

S. Seo, M. Lim, D. Lee, H. Park, **J. Oh**, D. J. Rim, and J.-H. Kim, “Environmental noise robustness for Korean fricatives using speech enhancement generative adversarial networks,” in *Proc.* IEEE Int. Conf. Big Data and Smart Computing (BigComp), pp. 1–4, 2019.

S. Seo, D. J. Rim, M. Lim, D. Lee, H. Park, **J. Oh**, C. Kim, and J.-H. Kim, “Shortcut connections based deep speaker embeddings for end-to-end speaker verification system,” in *Proc.* Interspeech, pp. 17, 2019.

Domestic Journal

이정필, 장재후, 김지현, 김민섭, 김성준, 김민서, 김하영, **오준석**, 정원, 김장연 외, “음성에 기반한 마비말장애 진단과 설명이 가능한 시스템,” 정보과학회지, 42(4): 45-56, 2024. [KSCI/KCI]

H. Park, Y. Kang, M. Lim, D. Lee, **J. Oh**, and J.-H. Kim, “LFMMI-based acoustic modeling by using external knowledge,” The Journal of the Acoustical Society of Korea, vol. 38, no. 5, pp. 607–613, 2019. [KSCI/KCI]

AWARDS

2023 년 한국어 AI 경진대회

주관: 한국지능정보사회진흥원(NIA)

- Track2-1, 상담 음성인식 : 오준석, 김민서, 남주형- team ‘상담 ONE’ , 장려상

2022 한국어 인공지능 경진대회

주관: 한국지능정보사회진흥원(NIA)

- 기업현안(회의음성): 오준석, 김하영 - team 'SGCSE', 최우수상/네이버 대표(1 위)

2021 숫자가 포함된 패턴발화 음성 데이터셋을 활용한 음절인식률 측정 알고리즘 개발 대회

주관: KT alpha

- 박호성, 오준석, 조은수-team ‘검은사케동’, 최우수상 수상(1 위)

PATENTS

KR 10-2699607 (B1) — Corpus Construction Service Provision Server and Method (코퍼스 구축 서비스 제공 서버 및 방법).

- Inventors: Ji-Hwan Kim; Junseok Oh; Hosung Park.
- Assignee: Sogang University Industry–Academic Cooperation Foundation.
- Filed: Apr 29, 2020 (KR 10-2020-0052570); Published: Nov 8, 2021 (KR 10-2021-0133667); Granted: Aug 22, 2024; Gazette: Aug 26, 2024.

CERTIFICATES

NVIDIA Deep Learning Institute Certificate, Building Conversational AI Applications, [2022]
- NVIDIA

TECHNICAL SKILLS

Programming Languages & Scripting
- Python

Deep Learning Frameworks
- Pytorch
- NVIDIA NeMo
- Hugging Face Transformers, PEFT

E2E ASR Toolkits
- Kaldi

Other Tools & Libraries
- KenLM