Yoori Oh

⌂ yoori000.github.io | in Yoori Oh | ✉ yoori0203@gmail.com

Research Interests

Multi-modal Learning, Video Understanding, Audio-Language Understanding, Audio-Visual Speech Processing

My research focuses on **multi-modal learning**, aiming to achieve a **deep understanding** of various modalities to build more comprehensive and intelligent systems. I am expanding my foundational work in **audio-language understanding** to the broader challenges of **video understanding**, alongside advanced **audio-visual speech synthesis**. I address fundamental challenges in imbalanced multi-modal data, designing optimized cross-modal embeddings to enhance both deep comprehension and expressive, controllable synthesis.

Education

Seoul National University   Seoul, Korea
Ph.D in Intelligence and Information   Sep. 2020 – Aug. 2025
Advisor: Prof. Kyogu Lee

Sookmyung Women's University   Seoul, Korea
M.S in Computer Science   Mar. 2015 – Feb. 2017
Advisor: Prof. Yoonhee Kim

Sookmyung Women's University   Seoul, Korea
B.S in Mathematics and Computer Science   Mar. 2010 – Feb. 2015

Work Experiences

Seoul National University   Seoul, Korea
Postdoctoral Researcher   Sep. 2025 – current

Korea Institute of Patent Information   Daejeon, Korea
Researcher   Aug. 2017 – Aug. 2020

Korea Institute of Patent Information   Daejeon, Korea
Intern   Apr. 2017 – Aug. 2017

Publications

[International Conferences](#)
(* Equal contribution; † Co-corresponding authors)

TripleSumm: Adaptive Triple-Modality Fusion for Video Summarization
Sumin Kim*, Hyemin Jeong*, Mingu Kang*, Yejin Kim, **Yoori Oh**† and Joonseok Lee†
International Conference on Learning Representations (ICLR), 2026

Sounding Highlights: Dual-Pathway Audio Encoders for Audio-Visual Video Highlight Detection

Seohyun Joo and **Yoori Oh**

IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2026

LipSody: Lip-to-Speech Synthesis with Enhanced Prosody Consistency

Jaejun Lee, **Yoori Oh** and Kyogu Lee

IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2026

Understanding Audio-Text Retrieval Through Singular Value Decomposition

Yoori Oh, Yoseob Han, Joonhyeon Bae, Jaeheon Sim and Kyogu Lee

ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR), 2025

Speaking Without Sound : Multi-speaker Silent Speech Voicing with Facial Inputs Only

Jaejun Lee, **Yoori Oh** and Kyogu Lee

IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2025

Distance Sampling-based Paraphraser Leveraging ChatGPT for Text Data Manipulation

Yoori Oh*, Yoseob Han* and Kyogu Lee

ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR), 2024

Hear Your Face : Face-based voice conversion with F0 estimation

Jaejun Lee, **Yoori Oh**, Injune Hwang and Kyogu Lee

INTERSPEECH, 2024

Semi-supervised Learning for Continuous Emotional Intensity Controllable Speech Synthesis with Disentangled Representations

Yoori Oh, Juheon Lee, Yoseob Han and Kyogu Lee

INTERSPEECH, 2023

Exploring Train and Test-Time Augmentations for Audio-Language Learning

Eungbeom Kim*, Jinhee Kim*, **Yoori Oh**, Kyungsu Kim, Minju Park, Jaeheon Sim, Jinwoo Lee and Kyogu Lee

arxiv, 2022

A Hybrid Cloud Resource Clustering Method Using Analysis of Application Characteristic
Yoori Oh and Yoonhee Kim
IEEE 2nd International Workshops on Foundations and Applications of Self* Systems (FAS*W), 2017

A SLA-based Spark Cluster Scaling Method in Cloud Environment
Yoori Oh, Jieun Choi, Eunjung Song, Moonji Kim and Yoonhee Kim
18th Asia-Pacific Network Operations and Management Symposium (APNOMS), 2016

International Journals
A Resource Recommendation Method based on Dynamic Cluster Analysis of Application Characteristics
Yoori Oh and Yoonhee Kim
Cluster Computing, 2019

Patents
QUEUE-BASED CONTRASTIVE LEARNING METHOD FOR AUDIO-TEXT RETRIEVAL AND COMPUTER DEVICE THEREOF
Jaeheon Sim, Kyogu Lee, **Yoori Oh**, Yoseob Han and Joonhyun Bae
KR, application number : 10-2025-0194175, 2025. 12

METHOD AND APPARATTUS FOR AUGMENTING AUDIO-TEXT MULTIMODAL DATA IN TRAIN AND TEST TIME
Kyogu Lee, Jinhee Kim, Eungbeom Kim, **Yoori Oh**, Jaeheon Sim, Jinwoo Lee, Kyungsu Kim and Minju Park
PCT, application number : PCT/KR2023/017075, 2023. 10

METHOD AND APPARATTUS FOR AUGMENTING AUDIO-TEXT MULTIMODAL DATA IN TRAIN AND TEST TIME
Kyogu Lee, Jinhee Kim, Eungbeom Kim, **Yoori Oh**, Jaeheon Sim, Jinwoo Lee, Kyungsu Kim and Minju Park
KR, application number : 10-2023-0146812, 2023. 10

METHOD AND APPARATUS FOR STRUCTURING PATENT CLAIMS
Youngmin Ahn, Hojin Park, Suhyun Oh, Jonghyun Chae, Bong-Gun Lee, Yoonho Jung, Jaeok Min and **Yoori Oh**
KR, registration number : 10-2181677, 2020. 11

MOBILE TERMINAL CONTROL METHOD FOR PERSONALIZED DIETARY COACHING
Yoonhee Kim, Jieun Choi, **Yoori Oh** and Youngran Jeon
KR, registration number : 10-2016263, 2019. 8

Honors and Scholarships	Teaching assistance scholarship (Seoul National University)	2024
	Brain Korea 21 Plus scholarship	2020 - 2023
	Bronze Award (12th open source software developer competition)	2018. 11
	Best Presentation Award (KIISE, KSC 2016)	2017. 01
	Bronze Best Paper Award (IEEE Student Paper Contest)	2016. 12
	Grand Prize (2015 Convergence App Contest)	2015. 12
	Best Paper Award (KSII 2015)	2015. 10

Teaching Experience

Teaching assistant, Seoul National University　　Spring 2024
Course title : From Sound to Meaning (Prof. Kyogu Lee)

Academic Services

Reviewer
- IEEE Transactions on Audio, Speech and Language Processing (TASLP)
- AAAI Conference on Artificial Intelligence (AAAI)