

Soo-Whan Chung

RESEARCH SCIENTIST – CLOVA SPEECH, NAVER CLOUD

NAVER 1784, 95, Jeongjail-ro, Bundang-gu, Seongnam-si, Gyeonggi Province, South Korea

✉ soowhan.chung@gmail.com | 🏠 soowhanchung.github.io | 📺 soowhan-chung | 📞 kevin6293 | 🎓 Google Scholar

“We will do something wonderful!”

Education

Doctor of Philosophy

Mar. 2016 - Feb. 2021

DEPARTMENT OF ELECTRICAL & ELECTRONIC ENGINEERING, YONSEI UNIVERSITY

Seoul, S.Korea

DIGITAL SIGNAL PROCESSING & ARTIFICIAL INTELLIGENCE (DSP&AI) LAB. (SUPERVISOR. HONG-GOO KANG)

Bachelor of Engineering

Mar. 2012 - Feb. 2016

DEPARTMENT OF ELECTRICAL & ELECTRONIC ENGINEERING, YONSEI UNIVERSITY

Seoul, S.Korea

FULL SCHOLARSHIP – THE NATIONAL SCHOLARSHIP FOR SCIENCE AND ENGINEERING (KOSAF)

Interests

Speech signal processing

SPEECH ENHANCEMENT/SEPARATION/RESTORATION/GENERATION, SPEAKER RECOGNITION, SPEECH RECOGNITION, KEYWORD SPOTTING

Multi-modal Learning

AUDIO-VISUAL, AUDIO-TEXT, CROSS-MODAL LEARNING

Career

NAVER Cloud

S. Korea

RESEARCH SCIENTIST @ SPEECH (PREV. NAVER CORP.)

Nov. 2020 - Present

- Machine / Deep Learning
- Speech Signal Processing
- Multi-modal Learning

NVIDIA

USA

DEEP LEARNING INSTITUTE CERTIFIED INSTRUCTOR

- Qualify for NVIDIA DLI University Ambassadorship
- Instruct deep learning courses using DLI program toolkit provided by NVIDIA corporation

NAVER

S. Korea

INTERNSHIP RESEARCH SCIENTIST

Jul. 2018 - Dec. 2018

- Collaborate research projects with NAVER Clova Speech (Advisor: Joon Son Chung)
- Conduct researches on audio-visual signal processing and speech signal processing

Publication

Dissertation

Self-supervised Cross-modal Representation for Visual-Guided Speech Separation

2021

ISNI : 0000 0004 9330 4678 [ENG/KOR]

Ph.D.

GRADUATE SCHOOL, YONSEI UNIVERSITY (SUPERVISOR. HONG-GOO KANG)

Journal

Perfect Match: Self-Supervised Embeddings for Cross-modal Retrieval

2020

Soo-Whan Chung, JOON SON CHUNG, HONG-GOO KANG

IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING, VOL. 14, NO. 3, PP.568–576, MAR., 2020

A study on speech disentanglement framework based on adversarial learning for speaker recognition

2020

YOOHWAN KWON, Soo-Whan Chung, HONG-GOO KANG

THE JOURNAL OF THE ACOUSTICAL SOCIETY OF KOREA, VOL. 39, NO. 5, PP.447-453, SEP., 2020

Generic uniform search grid generation algorithm for far-field source localization

JEESEOK LEE, **Soo-Whan Chung**, MIN-SEOK CHOI, HONG-GOO KANG

THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA, 143 (1), EL37-EL42, JAN., 2018

2018

Conference

Listen through the Sound: Generative Speech Restoration Leveraging Acoustic Context Representation

Soo-Whan Chung, MIN-SEOK CHOI

INTERSPEECH (TO BE PRESENT)

2025

Speak in the Scene: Diffusion-based Acoustic Scene Transfer toward Immersive Speech Generation

MISEUL KIM, **Soo-Whan Chung**, YOUNA JI, HONG-GOO KANG, MIN-SEOK CHOI

INTERSPEECH

2024

HD-DEMUCS: General Speech Restoration with Heterogeneous Decoders

DOYEON KIM, **Soo-Whan Chung**, HYEWON HAN, YOUNA JI, HONG-GOO KANG

INTERSPEECH

2023

MF-PAM: Accurate Pitch Estimation through Periodicity Analysis and Multi-level Feature Fusion

WOO-JIN CHUNG, DOYEON KIM, **Soo-Whan Chung**, HONG-GOO KANG

INTERSPEECH

2023

Imaginary Voice: Face-styled Diffusion Model for Text-to-Speech

JIYOUNG LEE, JOON SON CHUNG, **Soo-Whan Chung**

IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING (ICASSP)

2023

MoLE : Mixture of Language Experts for Multi-Lingual Automatic Speech Recognition

YOOHWAN KWON, **Soo-Whan Chung**

IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING (ICASSP)

2023

An Empirical Study on Speech Restoration Guided by Self-supervised Speech Representation

JAEUK BYUN, YOUNA JI, **Soo-Whan Chung**, SOYEON CHOE, MIN-SEOK CHOI

IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING (ICASSP)

2023

Diffusion-based Generative Speech Source Separation

ROBIN SCHEIBLER, YOUNA JI, **Soo-Whan Chung**, JAEUK BYUN, SOYEON CHOE, MIN-SEOK CHOI

IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING (ICASSP)

2023

Learning Audio-Text Agreement for Open-vocabulary Keyword Spotting

HYEON-KYEONG SHIN, HYEWON HAN, DOYEON KIM, **Soo-Whan Chung**, HONG-GOO KANG

INTERSPEECH – 🏆 BEST STUDENT PAPER AWARD FINALIST

2022

SASV 2022: The First Spoofing-Aware Speaker Verification Challenge

JEE-WEON JUNG, HEMLATA TAK, HYE-JIN SHIM, HEE-SOO HEO, BONG-JIN LEE, **Soo-Whan Chung**, HA-JIN YU, NICHOLAS EVANS, TOMI KINNUNEN

INTERSPEECH

2022

Baseline Systems for the First Spoofing-Aware Speaker Verification Challenge: Score and Embedding Fusion

HYE-JIN SHIM, HEMLATA TAK, XUECHEN LIU, HEE-SOO HEO, JEE-WEON JUNG, JOON SON CHUNG, **Soo-Whan Chung**, HA-JIN YU, BONG-JIN LEE, MASSIMILIANO TODISCO, HÉCTOR DELGADO, KONG AIK LEE, MD SAHIDULLAH, TOMI KINNUNEN AND NICHOLAS EVANS

ODYSSEY

2022

Phase Continuity: Learning Derivatives of Phase Spectrum from Speech Enhancement

DOYEON KIM, HYEWON HAN, HYEON-KYEONG SHIN, **Soo-Whan Chung**, HONG-GOO KANG

IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING (ICASSP)

2022

Looking into Your Speech: Learning Cross-modal Affinity for Audio-visual Speech Separation

JIYOUNG LEE*, **Soo-Whan Chung***, SUNOK KIM, HONG-GOO KANG, KWANGHOON SOHN

IEEE/CVF CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION (CVPR)

2021

Look Who's Talking: Active Speaker Detection in the Wild

YOU JIN KIM, HEE-SOO HEO, SOYEON CHOE, **Soo-Whan Chung**, YOOHWAN KWON, BONG-JIN LEE, YOUNGKI KWON, JOON SON CHUNG

INTERSPEECH

2021

End-to-end Lip Synchronisation based on Pattern Classification

YOU JIN KIM, HEE-SOO HEO, **Soo-Whan Chung**, BONG-JIN LEE

IEEE SPOKEN LANGUAGE TECHNOLOGY WORKSHOP (SLT)

2021

FaceFilter: Audio-visual speech separation using still images Soo-Whan Chung , SOYEON CHOE, JOON SON CHUNG, HONG-GOO KANG INTERSPEECH – 🏆 BEST STUDENT PAPER AWARD	2020
Seeing voices and hearing voices: Learning discriminative embeddings using cross-modal self-supervision Soo-Whan Chung , HONG-GOO KANG, JOON SON CHUNG INTERSPEECH	2020
MIRNet: Learning multiple identities representations in overlapped speech HYEWON HAN, Soo-Whan Chung , HONG-GOO KANG INTERSPEECH	2020
Intra-class variation reduction of speaker representation in disentanglement framework YOOHWAN KWON, Soo-Whan Chung , HONG-GOO KANG INTERSPEECH	2020
Perfect match: Improved Cross-modal Embeddings for Audio-visual Synchronisation Soo-Whan Chung , JOON SON CHUNG, HONG-GOO KANG IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING (ICASSP)	2019
Gradient-based Active Learning Query Strategy for End-to-end Speech Recognition YUAN YANG, Soo-Whan Chung , HONG-GOO KANG IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING (ICASSP)	2019
A study on search grid points for data-driven 3-D beamsteering JEE SOK LEE, Soo-Whan Chung , MIN-SEOK CHOI, HONG-GOO KANG HANDS-FREE SPEECH COMMUNICATIONS AND MICROPHONE ARRAYS (HSCMA)	2017

Research Projects

Naver / Naver Cloud

Generative Audio/Speech

MAIN ROLE: SCIENTIFIC RESEARCH	2022 - Present
<ul style="list-style-type: none"> Diffusion-based Audio/Speech Processing Token-based Speech Generation 	

Keyword Detection

MAIN ROLE: SCIENTIFIC RESEARCH	2020 - Present
<ul style="list-style-type: none"> Implement open-vocabulary keyword detection system Design acoustic models with small model size and low latency 	

Speech Enhancement/Restoration

MAIN ROLE: SCIENTIFIC RESEARCH	2021 - Present
<ul style="list-style-type: none"> Improve speech quality of recorded speech signals Design the speech enhancement model for the noise reduction and the de-reverberation 	

Academics

Reviewer

ICASSP, INTERSPEECH, MDPI, SPEECH COMMUNICATIONS

Organizer

SPOOFING-AWARE SPEAKER VERIFICATION (SASV) CHALLENGE, 2022

Conference Session Chair

INTERSPEECH (2022), INTERSPEECH (2024)

Skills

Programming	C/C++, Python
Deep learning	PyTorch, Tensorflow
Languages	Korean, English