

# Jee-weon Jung (updated September 29, 2024)

Email: [jeeweonj@ieee.org](mailto:jeeweonj@ieee.org)

Github: <https://github.com/Jungjee>

Google Scholar: <https://scholar.google.com/citations?user=A5OcLdAAAAAJ>

Webpage: <https://jungjee.github.io>

## PROFESSIONAL EXPERIENCE

---

- **Apple** September 2024 -  
*Senior Research Scientist (full time)* *Supervisor: Stephen Shum*
- **Carnegie Mellon University** May 2023 - September 2024  
*Postdoctoral Research Associate (full time)* *PI: Shinji Watanabe*
  - **Speech processing:** extended area of expertise by having 1st author publications on speech enhancement, speech summarization, and automated audio captioning. Co-authored papers on speech recognition, speech foundation model, multi-modal audio-visual recognition and generation, and spoken language understanding.
  - **Papers:** published 30 peer-reviewed papers
- **Naver Clova, Republic of Korea** December 2020 - April 2023  
*Research Scientist (full time)* *Supervisor: Bong-Jin Lee*
  - **Speaker recognition:** developed speaker embedding extractors for ClovaNote service
  - **Sound(audio) event detection:** developed volleyball whistle detection models for Naver volleyball streaming service.
  - **Audio anti-spoofing:** Developed state-of-the-art anti-spoofing systems.
  - **Papers and patents:** published 27 peer-reviewed papers and filed 3 patents

## EDUCATION

---

- **PhD, Computer Science and Engineering** March 2017 - February 2021  
*University of Seoul* *Seoul, Republic of Korea*  
*Thesis: Raw waveform speaker verification using an end-to-end deep neural network architecture and feature map scaling*  
*Supervisor: Ha-Jin Yu*
- **BBA** March 2013 - February 2017  
*University of Seoul* *Seoul, Republic of Korea*
- **BS, Computer Science and Engineering (Double major)** March 2013 - February 2017  
*University of Seoul* *Seoul, Republic of Korea*

## HONORS AND AWARDS

---

- Ranked **1st** place at the **DCASE 2024 Challenge task 6** (automated audio captioning, AAC) - 2024
- MERL award: **Joint CMU-MERL team wins DCASE2023 Challenge on Automated Audio Captioning** - 2023
- Ranked **1st** place at the **DCASE 2023 Challenge task 6-a** (automated audio captioning, AAC) - 2023
- Ranked **3rd** place at the **Third DIHARD Speech Diarisation Challenge** (track 1 core, speaker diarization, SD) - 2021
- Ranked 4th place at the **DCASE 2021 Challenge task 1-a** (acoustic scene classification, ASC) - 2021
- Ranked 6th place at the **DCASE 2019 Challenge task 1-a** (acoustic scene classification, ASC) - 2019
- Samsung Convergence Software Course Scholarship - 2015

## SKILLS

---

- **Language skills:** English, Korean (native), Chinese
- **Programming:** Python, Bash, C
- **Toolkits:** PyTorch, Theano, Keras, Kaldi

## COMMUNITY ENGAGEMENTS

---

### [Organizing special sessions, challenges, and workshops]

- **Spoofing-aware Speaker Verification (SASV) 2022 challenge** and a special session at **Interspeech 2022** (Founding organizer)
- VoxCeleb Speaker Recognition Challenge and satellite workshops at Interspeech **2022** and **2023**
- **Automatic Speaker Verification and Spoofing Countermeasures (ASVspoof) 5 Challenge** and satellite workshop at **Interspeech 2024**

### [Chairing sessions]

- IEEE ICASSP 2022
- ISCA Interspeech 2024
- ISCA Interspeech 2023
- ISCA Interspeech 2022
- ISCA Odyssey 2024

[Reviewing conference, workshop, and journal submissions]

- IEEE ICASSP 2019
- IEEE ICASSP 2020
- IEEE ICASSP 2021
- IEEE ICASSP 2022
- IEEE ICASSP 2023
- IEEE ICASSP 2024 (also meta-reviewer)
- IEEE ICASSP 2025
- IEEE ASRU 2023
- IEEE SLT 2022
- DCASE Workshop 2022
- DCASE Workshop 2024
- ISCA Interspeech 2021
- ISCA Interspeech 2022
- ISCA Interspeech 2023
- ISCA Interspeech 2024
- ISCA Odyssey 2024
- IEEE TASLP
- IEEE SPL
- IEEE Access
- Nature Scientific reports
- Elsevier CSL
- EUSIPCO
- ASVspoof 5 Workshop

[Giving tutorials]

ISCA Interspeech 2023 [[link](#)]

TEACHING AND TALKS

- 
- |   |   |
|---|---|
| • <b>Vulnerability of ASV and research on spoofing-robustness – invited talk</b><br><i>University of Buenos Aires</i>   | August 2024<br>BA, Argentina                |
| • <b>Speaker recognition and diarization – guest lecture</b><br><i>Carnegie Mellon University, 11492/11692/18495: Speech Technology for Conversational AI</i> | Spring 2024<br>PA, United States            |
| • <b>Speaker recognition in speech processing – guest lecture</b><br><i>Carnegie Mellon University, 11492/11692/18495: Speech Processing</i>                  | Spring 2023<br>PA, United States            |
| • <b>Python programming – 3-day course</b><br><i>Korea Meteorological Administration</i>  | September 2019<br>Jeju, Republic of Korea   |
| • <b>Python for deep learning – 1-day course</b><br><i>Korea Meteorological Administration</i>  | May 2019<br>Jeju, Republic of Korea         |
| • <b>Raw waveform speaker verification – invited talk</b><br><i>Naver</i>   | January 2019<br>Kyung-gi, Republic of Korea |

[2024]: 26 publications, including 9 first author publications

- [Paper] Jaesung Huh, Joon Son Chung, Arsha Nagrani, Andrew Brown, **Jee-weon Jung**, Daniel Garcia-Romero, Andrew Zisserman, “The Vox Celeb Speaker Recognition Challenge: A Retrospective,” in **IEEE/ACM Transactions on Audio, Speech, and Language Processing**.
- [Paper] Hye-jin Shim, Md Sahidullah, **Jee-weon Jung**, Shinji Watanabe, Tomi Kinnunen, “Beyond Silence: Bias Analysis through Loss and Asymmetric Approach in Audio Anti-Spoofing,” in **Proc. SynData4GenAI (Interspeech 2024 Satellite)**.
- [Paper] Xin Wang, Hector Delgado, Hemlata Tak, **Jee-weon Jung**, Hye-jin Shim, Massimiliano Todisco, Ivan Kukanov, Xuechen Liu, Md Sahidullah, Tomi Kinnunen, Nicholas Evans, Kong Aik Lee, Junichi Yamagishi, “ASVspoof 5: Crowdsourced Speech Data, Deepfakes, and Adversarial Attacks at Scale,” in **Proc. ASVspoof 2024 Workshop (Interspeech 2024 Satellite)**.
- [Paper] **Jee-weon Jung\***, Xin Wang\*, Nicholas Evans, Shinji Watanabe, Hye-jin Shim, Hemlata Tak, Siddhant Arora, Junichi Yamagishi, Joon Son Chung, “To what extent can ASV systems naturally defend against spoofing attacks?,” in **Proc. Interspeech**.
- [Paper] **Jee-weon Jung\***, Wangyou Zhang\*, Jiatong Shi\*, Zakaria Aldeneh, Takuya Higuchi, Barry-John Theobald, Ahmed Hussen Abdelaziz, Shinji Watanabe, “ESPnet-SPK: full pipeline speaker embedding toolkit with reproducible recipes, self-supervised front-ends, and off-the-shelf models,” in **Proc. Interspeech**.
- [Paper] Kihyun Nam, Hee-Soo Heo, **Jee-weon Jung**, Joon Son Chung, “Disentangled Representation Learning for Environment-agnostic Speaker Recognition,” in **Proc. Interspeech**.
- [Paper] Wangyou Zhang, Kohei Saijo, **Jee-weon Jung**, Chenda Li, Shinji Watanabe, Yanmian Qian, “Beyond Performance Plateaus: A Comprehensive Study on Scalability in Speech Enhancement,” in **Proc. Interspeech**.
- [Paper] Zakaria Aldeneh, Takuya Higuchi, **Jee-weon Jung**, Skyler Seto, Tatiana Likhomanenko, Stephen Shum, Ahmed Hussen Abdelaziz, Shinji Watanabe, Barry-John Theobald, “Can you Remove the Downstream Model for Speaker Recognition with Self-Supervised Speech Features?,” in **Proc. Interspeech**.
- [Paper] Yifan Peng, Jinchuan Tian, William Chen, Siddhant Arora, Brian Yan, Yui Sudo, Muhammad Shakeel, Kwanghee Choi, Jiatong Shi, Xuankai Chang, **Jee-weon Jung**, Shinji Watanabe, “OWSM v3.1: Better and Faster Open Whisper-Style Speech Models based on E-Branchformer,” in **Proc. Interspeech**.
- [Paper] Hye-jin Shim\*, **Jee-weon Jung\***, Tomi Kinnunen, Nicholas Evans, Jean-Francois Bonastre, Itshak Lapidot, “a-DCF: an architecture agnostic metric with application to spoofing-robust speaker verification,” in **Proc. Speaker Odyssey**.
- [Paper] **Jee-weon Jung\***, Roshan Sharma\*, William Chen, Bhiksha Raj, Shinji Watanabe, “AugSumm: towards generalizable speech summarization using synthetic labels from large language models,” in **Proc. ICASSP**.
- [Paper] Wangyou Zhang\*, **Jee-weon Jung\***, Shinji Watanabe, Yanmin Qian, “Improving Design of Input Condition Invariant Speech Enhancement,” in **Proc. ICASSP**.
- [Paper] Kwanghee Choi, **Jee-weon Jung**, Shinji Watanabe, “Understanding Probe Behaviors through Variational Bounds of Mutual Information,” in **Proc. ICASSP**.
- [Paper] Samuele Cornell, **Jee-weon Jung**, Shinji Watanabe, Stefano Squartini, “One model to rule them all? Towards End-to-End Joint Speaker Diarization and Speech Recognition,” in **Proc. ICASSP**.
- [Paper] Doyeop Kwak\*, Jaemin Jung\*, Kihyun Nam, Youngjoon Jang, **Jee-weon Jung**, Shinji Watanabe, Joon Son Chung, “VoxMM: Rich Transcription of Conversations in the Wild,” in **Proc. ICASSP**.
- [Paper] Soumi Maiti, Yifan Peng, Shukjae Choi, **Jee-weon Jung**, Xuankai Chang, Shinji Watanabe, “Voxtlm: unified decoder-only models for consolidating speech recognition, synthesis and speech, text continuation tasks,” in **Proc. ICASSP**.
- [Paper] Shih-Lun Wu, Xuankai Chang, Gordon Wichern, **Jee-weon Jung**, François Germain, Jonathan Le Roux, Shinji Watanabe, “Improving Audio Captioning Models with Fine-grained Audio Features, Text Embedding Supervision, and LLM Mix-up Augmentation,” in **Proc. ICASSP**.
- [Paper] Xuankai Chang, Brian Yan, Kwanghee Choi, **Jee-weon Jung**, Yichen Lu, Soumi Maiti, Roshan Sharma, Jiatong Shi, Jinchuan Tian, Shinji Watanabe, Yuya Fujita, Takashi Maekaku, Pengcheng Guo, Yao-Fei Cheng, Pavel Denisov, Kohei Saijo, Hsiu-Hsuan Wang, “Exploring Speech Recognition, Translation, and Understanding with Discrete Speech Units: A Comparative Study,” in **Proc. ICASSP**.
- [Paper] Siddhant Arora, Hayato Futami, **Jee-weon Jung**, Yifan Peng, Roshan Sharma, Yosuke Kashiwagi, Emiru Tsunoo, Shinji Watanabe, “UniverSLU: Universal Spoken Language Understanding for Diverse Classification and Sequence Generation Tasks with a Single Network,” in **Proc. NAACL**.
- [Paper] Siddhant Arora, Ankita Pasad, Chung-Ming Chien, Jionghao Han, Roshan Sharma, **Jee-weon Jung**, Hira Dhamyal, William Chen, Suwon Shon, Hung-yi Lee, Karen Livescu, Shinji Watanabe, “On the Evaluation of Speech Foundation Models for Spoken Language Understanding,” in **Proc. ACL Findings**.
- [Preprint] Zakaria Aldeneh, Takuya Higuchi, **Jee-weon Jung**, Li-Wei Chen, Stephen Shum, Ahmed Hussen Abdelaziz, Shinji Watanabe, Tatiana Likhomanenko, Barry-John Theobald, “Speaker-IPL: Unsupervised Learning of Speaker Characteristics with i-Vector based Pseudo-Labels,” in **arXiv**.

- [Preprint] Jiatong Shi, Jinchuan Tian, Yihan Wu, **Jee-weon Jung**, Jia Qi Yip, Yoshiki Masuyama, William Chen, Yuning Wu, Yuxun Tang, Massa Baali, Dareen Alharhi, Dong Zhang, Ruifan Deng, Tejes Srivastava, Haibin Wu, Alexander H Liu, Bhiksha Raj, Qin Jin, Ruihua Song, Shinji Watanabe, “ESPnet-Codes: Comprehensive Training and Evaluation of Neural Codecs for Audio, Music, and Speech,” in [arXiv](#).
- [Preprint] **Jee-weon Jung**, Yihan Wu, Xin Wang, Ji-Hoon Kim, Soumi Maiti, Yuta Matsunaga, Hye-jin Shim, Jinchuan Tian, Nicholas Evans, Joon Son Chung, Wangyou Zhang, Seyun Um, Shinnosuke Takamichi, Shinji Watanabe, “SpoofCeleb: Speech Deepfake Detection and SASV In The Wild,” in [arXiv](#).
- [Preprint] **Jee-weon Jung**, Wangyou Zhang, Soumi Maiti, Yihan Wu, Xin Wang, Ji-Hoon Kim, Yuta Matsunaga, Seyun Um, Jinchuan Tian, Hye-jin Shim, Nicholas Evans, Joon Son Chung, Shinnosuke Takamichi, Shinji Watanabe, “Text-To-Speech Synthesis In The Wild,” in [arXiv](#).
- [Preprint] Minsu Kim\*, **Jee-weon Jung\***, Hyeongseop Rha, Soumi Maiti, Siddhant Arora, Xuankai Chang, Shinji Watanabe, Yong Man Ro, “TMT: Tri-Modal Translation between Speech, Image, and Text by Processing Different Modalities as Different Languages,” in [arXiv](#).

- [Tech. Rep] **Jee-weon Jung**, Dong Zhang, Chao-Han Huck Yang, Shih-Lun Wu, David M. Chan, Zhifeng Kong, Ruifan Deng, Yaqian Zhou, Rafael Valle, Shinji Watanabe, “Automatic Audio Captioning with Encoder Fusion, Multi-Layer Aggregation, and Large Language Model Enriched Summarization,” in [https://dcase.community/documents/challenge2024/technical\\_reports/DCASE2024\\_Jung\\_74.t6.pdf](https://dcase.community/documents/challenge2024/technical_reports/DCASE2024_Jung_74.t6.pdf).

[2023]: 14 publications, including 3 first and 3 corresponding author publications

- [Paper] Yifan Peng, Jinchuan Tian, Brian Yan, Dan Berrebbi, Xuankai Chang, Xinjian Li, Jiatong Shi, Siddhant Arora, William Chen, Roshan Sharma, Wangyou Zhang, Yui Sudo, Muhammad Shakeel, **Jee-weon Jung**, Soumi Maiti, Shinji Watanabe, “Reproducing Whisper-Style Training Using an Open-Source Toolkit and Publicly Available Data,” in [Proc. IEEE ASRU](#).
- [Paper] **Jee-weon Jung**, Soonshin Seo, Hee-Soo Heo, Geonmin Kim, You Jin Kim, Young-ki Kwon, Minjae Lee, and Bong-Jin Lee, “Encoder-decoder multimodal speaker change detection,” in [Proc. Interspeech](#).
- [Paper] Sung Hwan Mun\*, Hye-jin Shim\*, Hemlata Tak\*, Xin Wang, Xuechen Liu, Md Sahidullah, Myeonghun Jeong, Min Hyun Han, Massimiliano Todisco, Kong Aik Lee, Junichi Yamagishi, Nicholas Evans, Tomi Kinnunen, Nam Soo Kim, and **Jee-weon Jung**, “Towards single integrated spoofing-aware speaker verification embeddings,” in [Proc. Interspeech](#).
- [Paper] Hye-jin Shim, **Jee-weon Jung**, Tomi Kinnunen, “Multi-Dataset Co-Training with Sharpness-Aware Optimization for Audio Anti-spoofing,” in [Proc. Interspeech](#).
- [Paper] Kihyun Nam, Youkyum Kim, Jaesung Huh, Hee Soo Heo, **Jee-weon Jung**, Joon Son Chung, “Disentangled representation learning for multilingual speaker recognition,” in [Proc. Interspeech](#).
- [Paper] Hee-Soo Heo, **Jee-weon Jung**, Jingu Kang, Youngki Kwon, You Jin Kim, Bong-Jin Lee, Joon Son Chung, “Curriculum learning for self-supervised speaker verification,” in [Proc. Interspeech](#).
- [Paper] **Jee-weon Jung**, Hee-Soo Heo, Bong-Jin Lee, Jaesung Huh, Andrew Brown, Youngki Kwon, Shinji Watanabe, Joon Son Chung, “In search of strong embedding extractors for speaker diarisation,” in [Proc. ICASSP](#).
- [Paper] You Jin Kim\*, Hee-Soo Heo\*, **Jee-weon Jung\***, Youngki Kwon, Bong-Jin Lee, Joon Son Chung, “Advancing the dimensionality reduction of speaker embeddings for speaker diarisation: disentangling noise and informing speech activity,” in [Proc. ICASSP](#).
- [Paper] Hee-Soo Heo, Youngki Kwon, Bong-Jin Lee, You Jin Kim, **Jee-weon Jung**, “High-resolution embedding extractor for speaker diarisation,” in [Proc. ICASSP](#).
- [Paper] Youngki Kwon, Hee-Soo Heo, Bong-Jin Lee, You Jin Kim, **Jee-weon Jung**, “Absolute decision corrupts absolutely: conservative online speaker diarisation,” in [Proc. ICASSP](#).
- [Paper] Shih-Lun Wu, Xuankai Chang, Gordon Wichern, **Jee-weon Jung**, François G Germain, Jonathan Le Roux, Shinji Watanabe, “On the Use of Pretrained Deep Audio Encoders for Automated Audio Captioning Tasks,” in [Proc. FAST-zero](#).
- [Preprint] Jaesung Huh, Andrew Brown, **Jee-weon Jung**, Joon Son Chung, Arsha Nagrani, Daniel Garcia-Romero, Andrew Zisserman, “VoxSRC 2022: The Fourth VoxCeleb Speaker Recognition Challenge,” in [arXiv](#).
- [Tech. Rep] Shih-Lun Wu, Xuankai Chang, Gordon Wichern, **Jee-weon Jung**, François Germain, Jonathan Le Roux, Shinji Watanabe, “BEATs-based Audio Captioning Model with Instructor Embedding Supervision and ChatGPT Mix-up,” in [https://dcase.community/documents/challenge2023/technical\\_reports/DCASE2023.Wu\\_31.t6a.pdf](https://dcase.community/documents/challenge2023/technical_reports/DCASE2023.Wu_31.t6a.pdf).
- [Tech. Rep] Hector Delgado, Nicholas Evans, **Jee-weon Jung**, Tomi Kinnunen, Ivan Kukanov, Kong Aik Lee, Xuechen Liu, Hye-jin Shim, Md Sahidullah, Hemlata Tak, Massimiliano Todisco, Xin Wang, Junichi Yamagishi, “ASVspoof 5 Evaluation Plan,” in [https://www.asvspoof.org/file/ASVspoof5\\_Evaluation\\_Plan.pdf](https://www.asvspoof.org/file/ASVspoof5_Evaluation_Plan.pdf).

[2022]: 10 publications, including 6 first author publications

- [Paper] **Jee-weon Jung**, Hee-Soo Heo, Hemlata Tak, Hye-jin Shim, Joon Son Chung, Bong-Jin Lee, Ha-Jin Yu, Nicholas Evans, “Aasist: Audio anti-spoofing using integrated spectro-temporal graph attention networks,” in [Proc. ICASSP](#).
- [Paper] Hye-jin Shim, **Jee-weon Jung**, Ju-ho Kim, Ha-Jin Yu, “Attentive Max Feature Map and Joint Training for Acoustic Scene Classification,” in [Proc. ICASSP](#).

- [Paper] Yongki Kwon, Hee-Soo Heo, **Jee-weon Jung**, You Jin Kim, Bong-Jin Lee, Joon Son Chung, “Multi-scale speaker embedding-based graph attention networks for speaker diarisation,” in **Proc. ICASSP**.
- [Paper] **Jee-weon Jung**<sup>\*</sup>, You Jin Kim<sup>\*</sup>, Hee-Soo Heo, Bong-Jin Lee, Youngki Kwon, Joon Son Chung, “Pushing the limits of raw waveform speaker verification,” in **Proc. Interspeech**.
- [Paper] **Jee-weon Jung**<sup>\*</sup>, Hemlata Tak<sup>\*</sup>, Hye-jin Shim, Hee-Soo Heo, Bong-Jin Lee, Soo-Whan Chung, Ha-Jin Yu, Nicholas Evans, Tomi Kinnunen, “SASV 2022: The First Spoofing-Aware Speaker Verification Challenge,” in **Proc. Interspeech**.
- [Paper] 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, Nicholas Evans, “Baseline systems for the first spoofing-aware speaker verification challenge: score and embedding fusion,” in **Proc. Speaker Odyssey**.
- [Paper] Hemlata Tak, Massimiliano Todisco, Xin Wang, **Jee-weon Jung**, Junichi Yamagishi, Nicholas Evans, “Automatic speaker verification spoofing and deepfake detection using wav2vec 2.0 and data augmentation,” in **Proc. Speaker Odyssey**.
- [Paper] Sung Hwan Mun<sup>\*</sup>, **Jee-weon Jung**<sup>\*</sup>, Min Hyun Han, Nam Soo Kim, “Frequency and Multi-Scale Selective Kernel Attention for Speaker Verification,” in **Proc. SLT**.
- [Preprint] **Jee-weon Jung**, Hee-Soo Heo, Bong-Jin Lee, Jaesong Lee, Hye-jin Hsim, Youngki Kwon, Joon Son Chung, Shinji Watanabe, “Large-scale learning of generalised representations for speaker recognition,” in **arXiv**.
- [Preprint] **Jee-weon Jung**, Hemlata Tak, Hye-jin Shim, Hee-Soo Heo, Bong-Jin Lee, Soo-Whan Chung, Hong-Goo Kang, Ha-Jin Yu, Nicholas Evans, Tomi Kinnunen, “SASV Challenge 2022: A Spoofing Aware Speaker Verification Challenge Evaluation Plan,” in **arXiv**.

[2021]: 10 publications, including 5 first author publications

- [Paper] **Jee-weon Jung**<sup>\*</sup>, Hee-Soo Heo<sup>\*</sup>, Ha-Jin Yu, Joon Son Chung, “Graph Attention Networks for Speaker Verification,” in **Proc. ICASSP**.
- [Paper] **Jee-weon Jung**, Hye-jin Shim, Ju-ho Kim, Ha-Jin Yu, “DCASENET: An integrated pre-trained deep neural network for detecting and classifying acoustic scenes and events,” in **Proc. ICASSP**.
- [Paper] **Jee-weon Jung**, Hee-Soo Heo, Youngki Kwon, Joon Son Chung, Bong-Jin Lee, “Three-class Overlapped Speech Detection using a Convolutional Recurrent Neural Network,” in **Proc. Interspeech**.
- [Paper] Hye-jin Shim<sup>\*</sup>, **Jee-weon Jung**<sup>\*</sup>, Ju-ho Kim, Ha-jin Yu, “Capturing Discriminative Information Using a Deep Architecture in Acoustic Scene Classification,” in **Applied Sciences**.
- [Paper] Youngki Kwon, **Jee-weon Jung**, Hee-Soo Heo, You Jin Kim, Bong-Jin Lee, Joon Son Chung, “Adapting Speaker Embeddings for Speaker Diarisation,” in **Proc. Interspeech**.
- [Paper] Hemlata Tak, **Jee-weon Jung**, Jose Patino, Massimiliano Todisco, Nicholas Evans, “Graph Attention Networks for Anti-Spoofing,” in **Proc. Interspeech**.
- [Paper] Hemlata Tak, **Jee-weon Jung**, Jose Patino, Madhu Kamble, Massimiliano Todisco, Nicholas Evans, “End-to-End Spectro-Temporal Graph Attention Networks for Speaker Verification Anti-Spoofing and Speech Deepfake Detection,” in **Proc. ASVspoof Workshop 2021 (Interspeech satellite)**.
- [Paper] Ju-ho Kim, Hye-jin Shim, **Jee-weon Jung**, Ha-Jin Yu, “A Supervised Learning Method for Improving the Generalization of Speaker Verification Systems by Learning Metrics from a Mean Teacher,” in **Applied Sciences**.
- [Tech. Rep] Hee-Soo Heo<sup>\*</sup>, **Jee-weon Jung**<sup>\*</sup>, Hye-jin Shim, Bong-Jin Lee, “Clova Submission For The DCASE 2021 Challenge: Acoustic Scene Classification Using Light Architectures and Device Augmentation,” in [http://dcase.community/documents/challenge2021/technical\\_reports/DCASE2021\\_Heo\\_30-t1.pdf](http://dcase.community/documents/challenge2021/technical_reports/DCASE2021_Heo_30-t1.pdf).
- [Tech. Rep] Hee-Soo Heo, **Jee-weon Jung**, Youngki Kwon, J Kim, Jaesung Huh, Joon Son Chung, Bong-Jin Lee, “NAVER Clova Submission To The Third DIHARD Challenge,” in [https://dihardchallenge.github.io/dihard3/system\\_descriptions/dihard3\\_system\\_description\\_team73.pdf](https://dihardchallenge.github.io/dihard3/system_descriptions/dihard3_system_description_team73.pdf).

[2020]: 11 publications, including 9 first author publications

- [Paper] **Jee-weon Jung**, Hee-Soo Heo, Hye-jin Shim, Ha-Jin Yu, “Knowledge Distillation in Acoustic Scene Classification,” in **IEEE Access**.
- [Paper] **Jee-weon Jung**, Hye-jin Shim, Ju-ho Kim, Seung-bin Kim, Ha-Jin Yu, “Acoustic Scene Classification using Audio Tagging,” in **Proc. Interspeech**.
- [Paper] Hye-jin Shim<sup>\*</sup>, **Jee-weon Jung**<sup>\*</sup>, Ju-ho Kim, Ha-Jin Yu, “Integrated Replay Spoofing-Aware Text-Independent Speaker Verification,” in **Applied Sciences**.
- [Paper] Seung-bin Kim<sup>\*</sup>, **Jee-weon Jung**<sup>\*</sup>, Hye-jin Shim, Ju-ho Kim, Ha-Jin Yu, “Segment Aggregation for short utterances speaker verification using raw waveforms,” in **Proc. Interspeech**.
- [Paper] **Jee-weon Jung**<sup>\*</sup>, Seung-bin Kim<sup>\*</sup>, Hye-jin Shim, Ju-ho Kim, Ha-Jin Yu, “Improved rawnet with filter-wise rescaling for text-independent speaker verification using raw waveforms,” in **Proc. Interspeech**.
- [Paper] **Jee-weon Jung**<sup>\*</sup>, Ju-ho Kim<sup>\*</sup>, Hye-jin Shim, Seung-bin Kim, Ha-Jin Yu, “Selective Deep Speaker Embedding Enhancement for Speaker Verification,” in **Proc. Speaker Odyssey**.



- [Paper] Ju-ho Kim\*, **Jee-weon Jung\***, Hye-jin Shim, Ha-Jin Yu, "Audio Tag Representation Guided Dual Attention Network for Acoustic Scene Classification," in **Proc. DCASE Workshop**.
- [Paper] **Jee-weon Jung**, Hye-jin Shim, Ju-ho Kim, Ha-Jin Yu, " $\alpha$ -feature map scaling for raw waveform speaker verification," in **the Journal of the Acoustical Society of Korea**.
- [Paper] Hye-jin Shim, Hee-Soo Heo, **Jee-weon Jung**, Ha-Jin Yu, "Self-supervised pre-training with acoustic configurations for replay spoofing detection," in **Proc. Interspeech**.
- [Preprint] **Jee-weon Jung**, Hye-jin Shim, Hee-Soo Heo, Ha-Jin Yu, "A study on the role of subsidiary information in replay attack spoofing detection," in **arXiv**.
- [Tech. Rep] Hye-jin Shim, Ju-ho Kim, **Jee-weon Jung**, Ha-jin Yu, "Audio tagging and deep architectures for acoustic scene classification: Uos submission for the DCASE 2020 challenge," in [https://dcase.community/documents/challenge2020/technical\\_reports/DCASE2020\\_Shim\\_120.pdf](https://dcase.community/documents/challenge2020/technical_reports/DCASE2020_Shim_120.pdf).

[2019]: 7 publications, including 4 first author publications

- [Paper] **Jee-weon Jung**, Hee-Soo Heo, Ju-ho Kim, Hye-jin Shim, Ha-Jin Yu, "RawNet: Advanced end-to-end deep neural network using raw waveforms for text-independent speaker verification," in **Proc. Interspeech**.
- [Paper] **Jee-weon Jung\***, Hye-jin Shim\*, Hee-Soo Heo, Ha-Jin Yu, "Replay attack detection with complementary high-resolution information using end-to-end DNN for the ASVspoof 2019 Challenge," in **Proc. Interspeech**.
- [Paper] Hee-Soo Heo\*, **Jee-weon Jung\***, Hye-jin Shim, Ha-Jin Yu, "Acoustic scene classification using teacher-student learning with soft-labels," in **Proc. Interspeech**.
- [Paper] Hee-Soo Heo, **Jee-weon Jung**, IL-Ho Yang, Sung-Hyun Yoon, Hye-jin Shim, Ha-Jin Yu, "End-to-end losses based on speaker basis vectors and all-speaker hard negative mining for speaker verification," in **Proc. Interspeech**.
- [Paper] **Jee-weon Jung\***, Hee-Soo Heo\*, Hye-jin Shim, Ha-Jin Yu, "Distilling the Knowledge of Specialist Deep Neural Networks in Acoustic Scene Classification," in **Proc. DCASE Workshop**.
- [Paper] Ju-ho Kim, Hee-Soo Heo, **Jee-weon Jung**, Hye-jin Shim, Seung-Bin Kim, Ha-Jin Yu, "Utilization of age information for speaker verification using multi-task learning deep neural networks," in **the Journal of the Acoustical Society of Korea**.
- [Preprint] Hee-Soo Heo, **Jee-weon Jung**, Hye-jin Shim, IL-Ho Yang, Ha-Jin Yu, "Cosine similarity-based adversarial process," in **arXiv**.

[2018]: 6 publications, including 5 first author publications

- [Paper] **Jee-weon Jung**, Hee-Soo Heo, IL-ho Yang, Hye-jin Shim, Ha-Jin Yu, "A Complete End-to-End Speaker Verification System using Deep Neural Networks: from Raw Signals to Verification Result," in **Proc. ICASSP**.
- [Paper] **Jee-weon Jung**, Hee-Soo Heo, IL-Ho Yang, Hye-Jin Shim, Ha-Jin Yu, "Avoiding Speaker Overfitting in End-to-End DNNs using Raw Waveform for Text-Independent Speaker Verification," in **Proc. Interspeech**.
- [Paper] **Jee-weon Jung\***, Hee-soo Heo\*, Hye-jin Shim, Ha-Jin Yu, "Short utterance compensation in speaker verification via cosine-based teacher-student learning of speaker embeddings," in **Proc. IEEE ASRU**.
- [Paper] **Jee-weon Jung\***, Hee-Soo Heo\*, Hye-jin Shim, Ha-Jin Yu, "DNN based multi-level features ensemble for acoustic scene classification," in **Proc. DCASE Workshop**.
- [Paper] **Jee-weon Jung**, Hee-Soo Heo, Hye-jin Shim, Ha-Jin Yu, "Multi channel far field speaker verification using teacher student deep neural networks," in **the Journal of the Acoustical Society of Korea**.
- [Paper] Hye-jin Shim, **Jee-weon Jung**, Hee-Soo Heo, Sung-Hyun Yoon, Ha-Jin Yu, "Replay spoofing detection system for automatic speaker verification using multi-task learning of noise classes," in **Proc. IEEE TAAI**.

[2017]: 3 publications, including 2 first author publications

- [Paper] **Jee-weon Jung**, Heesoo Heo, Ilho Yang, Sunghyun Yoon, Hyejin Shim, Hajin Yu, "D-vector based speaker verification system using Raw Waveform CNN," in **Proc. ANIT**.
- [Paper] **Jee-weon Jung**, Hee-Soo Heo, IL-Ho Yang, Sung-Hyun Yoon, Hye-Jin Shim, Ha-Jin Yu, "DNN-based audio scene classification for DCASE 2017: dual input features, balancing cost, and stochastic data duplication," in **Proc. DCASE Workshop**.
- [Paper] Hee-soo Heo, **Jee-weon Jung**, IL-ho Yang, Sung-hyun Yoon, Ha-jin Yu, "Joint training of expanded end-to-end DNN for text-dependent speaker verification," in **Proc. Interspeech**.