

Hye-jin Shim

PH.D. · AI RESEARCHER

✉ shimhz6.6@gmail.com | 📞 hye-jin-shim-95273b108/ | 🏠 Hye-jin Shim

Updated date: Jul 27th, 2026

Summary Applied audio ML researcher with 10+ years designing and training deep learning models across speech and ambient audio: acoustic scene/event understanding (five top-10 DCASE Challenge entries), speech quality assessment and enhancement evaluation under real-world channel conditions, and self-supervised audio representation learning. Experienced in deploying under real-time and embedded constraints (real-time speaker recognition for the Korean National Police Agency, embedded ASR for robot platforms) and in model efficiency via knowledge distillation and teacher-student training. Core contributor to widely adopted open-source speech libraries (ESPnet, VERSA) and organizer/primary metric designer for international evaluation challenges (ASVspoof 5, WildSpoof at ICASSP 2026). 45 publications (14 first-author) at Interspeech, ICASSP, and NeurIPS (2025 spotlight), incl. AASIST, a standard audio anti-spoofing architecture. Strong PyTorch engineering on HPC (Slurm, Docker). Currently leading a trustworthiness audit of audio-LLM speech-quality judges: shortcut learning, score calibration, and demographic bias.

Employment

Self-employed (Seattle, USA)

INDEPENDENT RESEARCHER

June. 2025 - Present

- WildSpoof Challenge organization (ICASSP 2026 Grand Challenge); end-to-end design of data, baselines, and evaluation protocol
- Leading a trustworthiness audit of audio-LLM quality judges: shortcut learning, calibration, and demographic bias

Carnegie Mellon University (Pittsburgh, USA)

POST-DOCTORAL RESEARCHER LANGUAGE TECHNOLOGIES INSTITUTE

Feb. 2024 - May. 2025

- PI: Shinji Watanabe
- Led the audio/speech quality assessment project for channel profiling; model development, open-source project contribution, data analysis
- Served as an organizer of ASVspoof 5 challenge; prepared the baselines, analyzed audio deep fake data, invented the primary evaluation metrics
- Mentored projects; audio large model development, geolocation-aware language identification, partial spoof diarization, audio deep fake source tracing

University of Eastern Finland (Joensuu, Finland)

POST-DOCTORAL RESEARCHER SCHOOL OF COMPUTING

Dec. 2022 - Nov. 2023

- PI: Tomi Kinnunen
- Led the SPEECHFAKES project on data bias analysis for audio deep fake detection
- Served as an organizer of ASVspoof 5 challenge; data preparation, planned the challenge pipeline
- Developed model trained with diverse domain datasets using sharpness aware optimization

Education

University of Seoul (Seoul, South Korea)

DOCTOR OF PHILOSOPHY IN COMPUTER SCIENCE AND MACHINE LEARNING

Aug. 2022

- GPA: 4.29 / 4.50
- Thesis: 'Graph Attention Deep Neural Networks for Speaker Verification and Audio Anti-spoofing Integrated System.'
- Supervisor: Ha-Jin Yu

University of Seoul (Seoul, South Korea)

MASTER OF SCIENCE IN COMPUTER SCIENCE

Feb. 2019

- GPA: 4.05 / 4.50
- Thesis: 'Replay attack spoofing detection system for speaker recognition using multi-task deep learning.'
- Supervisor: Ha-Jin Yu

Kwangwoon University (Seoul, South Korea)

BACHELOR OF COMMERCE IN INTERNATIONAL TRADE

Feb. 2017

- GPA: 3.94 / 4.50
- Major courses: mainly focused on Economics, but also including Business, Chinese Economics & Industry
- Summa cum laude - Full Scholarship(2014), Top 15% - Partial Scholarship(2013, 2015-16)
- Exchange Student at Shenzhen University, China (2014)
- Passed courses related to CS:
Web development (HTML), C language, Computer Engineering, Data structure using C++, Algorithm, Linear Algebra, Probability and Statistics
- Undergraduate research intern at University of Seoul starting from Jun 2016
- Served as a volunteer teacher in elementary, middle, and high schools (2 years)
- Served as a volunteer mentor for Chinese exchange student adaptation program (2 years)

Skills

Programming Python, C/C++
Tools PyTorch, ESPnet, Kaldi, TensorFlow/Keras, Docker, Slurm, Git
Languages Korean, English, Chinese, French (Beginner)

Community Engagements

2022- **Conference Review**, INTERSPEECH, ICASSP, ASRU, SLT, Odyssey
2020- **Journal Review**, Digital Signal Processing, Pattern Recognition Letters, Applied Acoustics, Computer Speech & Language, Transactions on Audio, Speech and Language
2025- **Challenge Organizer**, WildSpoof Challenge, ICASSP 2026
2023- **Challenge Organizer**, ASVspoof5 Challenge, INTERSPEECH 2024 Satellite
2022 **Session Chair**, Spoofing-Aware Speaker Verification(Special Session), INTERSPEECH 2022
2022 **Challenge Organizer**, Spoofing-Aware Speaker Verification Challenge, INTERSPEECH 2022

Projects

2024-25 **Acoustic modeling - audio large model development**, National Institute of Advanced Industrial Science and Technology (AIST)
2024-25 **Channel Profiling - Speech Quality Assessment**, Defence Science and Technology Agency (DSTA)
2022-23 **SPEECHFAKES - Audio deepfake bias analysis**, Academy of Finland
2020-22 **Acoustic Scene Classification System Using Audio Tagging Expression Features**, The National Research Foundation of Korea
2020-21 **PTTS generation using call voice**, KT
2017-21 **Real-time Speaker Recognition via Voice-print Analysis**, The Korean National Police Agency
2019-20 **Development of Speaker Recognition Algorithm for Voice Cloning TTS**, KT
2019-20 **Development of Speaker Recognition and Environment Recognition Algorithms**, Insoori Solution
2017-20 **Development of Free-running Speech Recognition Technologies for Embedded Robot System**, The Ministry of Trade, Industry & Energy
2017-20 **Development of Voice Forgery Detection System for Speaker Authentication**, The National Research Foundation of Korea
2018-19 **Development of Speaker Identification and Verification Technology through Voice Analysis**, Samsung Electronics

Challenge Participations

2025 **3rd Place**, ICME 2025 Audio Encoder Capability Challenge
2021 **4th Place**, Detection and Classification of Acoustic Scenes and Events (DCASE) Challenge - Task 1
2020 **10th Place**, Detection and Classification of Acoustic Scenes and Events (DCASE) Challenge - Task 1
2019 **10th Place**, Automatic Speaker Verification Spoofing and countermeasures (ASVspoof) Challenge
2019 **6th Place**, Detection and Classification of Acoustic Scenes and Events (DCASE) Challenge - Task 1
2018 **7th Place**, Detection and Classification of Acoustic Scenes and Events (DCASE) Challenge - Task 1
2017 **8th Place**, Detection and Classification of Acoustic Scenes and Events (DCASE) Challenge - Task 1

Teaching Experience

2024 **Carnegie Mellon University**, Speech Technology for Conversational AI - Speech/Audio Classification
2023 **INTERSPEECH**, Tutorial Presenter: Advances in Audio Anti-spoofing and Deepfake Detection using Graph Neural Networks and Self-supervised Learning
2023 **University of Eastern Finland**, Introduction to Speech and Machine Learning - Audio Anti-spoofing using Graph Neural Networks
2023 **University of Eastern Finland**, Introduction to Speech and Machine Learning - Recent Trends in Audio Anti-spoofing
2023 **University of Eastern Finland**, Special Course in Speech Security and Privacy - Advanced Topics in Audio Anti-spoofing
2023 **University of Eastern Finland**, Special Course in Speech Security and Privacy - Introduction to Graph Neural Networks
2022 **University of Eastern Finland**, Presenter: Exploration in Acoustic Domain Deep-learning Research
2017 **University of Seoul**, Teaching Assistant: C language

< INTERNATIONAL JOURNAL > (9 PAPERS: 3 FIRST-AUTHOR, 6 CO-AUTHORED)

- Jagabandhu Mishra, Manasi Chhibber, **Hye-jin Shim**, Tomi Kinnunen, “Probabilistic Characterization of Spoofed Speech.”, *Computer Speech And Language*, 2025
- Xin Wang, Héctor Delgado, Hemlata Tak, Jeeweon Jung, **Hye-jin Shim**, Massimiliano Todisco, Ivan Kukanov, Xuechen Liu, Md Sahidullah, Tomi Kinnunen, Nicholas Evans, Kong Aik Lee, Junichi Yamagishi, Myeonghun Jeong, Ge Zhu, Yongyi Zang, Soumi Maiti, Florian Lux, Nicolas Muller, Wangyou Zhang, Chengzhe Sun, Shuwei Hou, Siwei Lyu, Sebastien Le Maguer, Cheng Gong, Hanjie Guo, Liping Chen, Vishwanath Singh, “ASVspoof 5: Design, Collection and Validation of Public Resources for Spoofing, Deepfake, and Adversarial Attack Detection Using Crowdsourced Speech.”, *Computer Speech And Language*, 2025
- Md Sahidullah, **Hye-jin Shim**, Rosa González Hautamäki, Tomi Kinnunen, “Shortcut Learning in Classifier Black-Boxes: Applications to Voice Anti-Spoofing and Biometrics”, *IEEE Journal of selected topics in signal processing*, 2025
- 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, Sinnosuke Takamichi, Shinji Watanabe, “SpoofCeleb: Speech Deepfake Detection and SASV In The Wild.”, *IEEE Open Journal of Signal Processing*, 2024
- **Hye-jin Shim**, Jee-weon Jung, Ha-jin Yu, “Which to select?: Analysis of speaker representation with graph attention networks.”, *The Journal of the Acoustical Society of America*, 2024
- **Hye-jin Shim***, Jee-weon Jung*, Ju-ho Kim, Ha-jin Yu, “Capturing Discriminative Information Using a Deep Architecture in Acoustic Scene Classification.”, *Applied Sciences*, 2021
- Ju-ho Kim, **Hye-jin Shim**, Jungwoo Heo, Ha-Jin Yu, “A Supervised Learning Method for Improving the Generalization of Speaker Verification Systems by Learning Metrics from a Mean Teacher.”, *Applied Sciences*, 2021
- Jee-weon Jung, Hee-soo Heo, **Hye-jin Shim**, Ha-Jin Yu, “Knowledge distillation in acoustic scene classification.”, *IEEE Access*, 2020
- **Hye-jin Shim***, Jee-weon Jung*, Ju-ho Kim, Ha-jin Yu, “Integrated Replay Spoofing-Aware Text Independent Speaker Verification.”, *Applied Sciences*, 2020

< INTERNATIONAL CONFERENCE > (43 PAPERS: 11 FIRST-AUTHOR, 32 CO-AUTHORED)

- Zhuoyan Tao, Jiatong Shi, **Hye-jin Shim**, Shinji Watanabe, “ANCHOR: Autoregressive Non-intrusive Chunk-Ordered Refinement for Joint Multi-Resolution Speech Quality Modeling”, *Accepted to Interspeech*, 2026
- Yihan Wu, Jee-weon Jung, **Hye-jin Shim**, Xin Cheng, Xin Wang, “WildSpoof: advancing in-the-wild data in Text-to-Speech generation and Spoofing-aware automatic speaker verification”, *ICASSP*, 2026
- Yihan Wu, Jee-weon Jung, **Hye-jin Shim**, Xin Cheng, Xin Wang, “WildSpoof Challenge Evaluation Plan”, *Preprint*, 2025
- Qingzheng Wang, **Hye-jin Shim**, Jiancheng Sun, Shinji Watanabe, “Geolocation-Aware Robust Spoken Language Identification”, *ASRU*, 2025
- Kyo-Won Koo, Chan-yeong Lim, Jee-weon Jung, **Hye-jin Shim**, Ha-Jin Yu, “Token-based Attractors and Cross-attention in Spoof Diarization”, *ASRU*, 2025
- Jiatong Shi, Yifan Cheng, Bohao Su, **Hye-jin Shim**, Jinchuan Tian, Samuele Cornell, Yiwen Zhao, Siddhant Arora, Megan Choo, Daniel Leong, Shinji Watanabe, “ARECHO: Autoregressive Evaluation via Chain-Based Hypothesis Optimization for Speech Multi-Metric Estimation”, *NeurIPS as spotlight*, 2025
- Anyu Ying, Natarajan Balaji Shankar, Chyi-Jiunn Lin, Mohan Shi, Pu Wang, **Hye-jin Shim**, Siddhant Arora, Hugo Van Hamme, Abeer Alwan, Shinji Watanabe, “Benchmarking Training Paradigms, Dataset Composition, and Model Scaling for Child ASR in ESPnet”, *Interspeech satellites workshop on Child Computer Interaction*, 2025
- Shikhar Bharadwaj, Samuele Cornell, Kwanghee Choi, Satoru Fukayama, **Hye-jin Shim**, Soham Deshmukh, Shinji Watanabe, “OpenBEATS: A Fully Open-Source General-Purpose Audio Encoder”, *WASPAA*, 2025
- 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, Sinnosuke Takamichi, Shinji Watanabe, “Text-To-Speech Synthesis In The Wild”, *INTERSPEECH*, 2025
- Jiatong Shi, **Hye-jin Shim**, Shinji Watanabe, “Uni-VERSA: Versatile Speech Assessment with a Unified Network”, *INTERSPEECH*, 2025
- Jiatong Shi, **Hye-jin Shim**, Jinchuan Tian, Siddhant Arora, Haibin Wu, Darius Petermann, Jia Qi Yip, You Zhang, Yuxun Tang, Wangyou Zhang, Dareen Alharthi, Yichen Huang, Koichi Saito, Jionghao Han, Yiwen Zhao, Chris Donahue, Shinji Watanabe, “VERSA: A Versatile Evaluation Toolkit for Speech, Audio, Music”, *NAACL 2025 Demo Track*, 2025
- Manasi Chhibber, Jagabandhu Mishra, **Hye-jin Shim**, Tomi Kinnunen, “An Explainable Probabilistic Attribute Embedding Approach for Spoofed

- **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”, *Interspeech Satellite Workshop - SynData4GenAI*, 2024
- Jee-weon Jung*, Xin Wang*, Nicholas Evans, Shinji Watanabe, **Hye-jin Shim**, Hemlata Tak, Sidhant Arora, Junichi Yamagishi, Joon Son Chung, “To what extent can ASV systems naturally defend against spoofing attacks?”, *Interspeech*, 2024
- **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.”, *Odyssey*, 2024
- Jee-weon Jung, Hee-Soo Heo, Bong-Jin Lee, Jaesong Lee, **Hye-jin Shim**, Youngki Kwon, Joon Son Chung, Shinji Watanabe, “Large-scale learning of generalised representations for speaker recognition.”, *Preprint*, 2023
- 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 Evaluation Plan.”, *Preprint*, 2023
- **Hye-jin Shim**, Rosa González Hautamäki, Md Sahidullah, Tomi Kinnunen, “How to Construct Perfect and Worse-than-Coin-Flip Spoofing Countermeasures: A Word of Warning on Shortcut Learning.”, *INTERSPEECH*, 2023
- **Hye-jin Shim**, Jee-weon Jung, Tomi Kinnunen, “Multi-Dataset Co-Training with Sharpness-Aware Optimization for Audio Anti-spoofing.”, *INTERSPEECH*, 2023
- 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, Jee-weon Jung, “Towards single integrated spoofing-aware speaker verification embeddings.”, *INTERSPEECH*, 2023
- Jee-weon Jung*, Hemlata Tak*, **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.”, *INTERSPEECH*, 2022
- Ju-ho Kim*, Jungwoo Heo*, **Hye-jin Shim**, Ha-Jin Yu, “Extended U-Net for Speaker Verification in Noisy Environments.”, *INTERSPEECH*, 2022
- **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, Hector Delgado, Kong Aik Lee Md Sahidullah, Tomi Kinnunen, Nicholas Evans, “Baseline Systems for the First Spoofing-Aware Speaker Verification Challenge: Score and Embedding Fusion.”, *Odyssey*, 2022
- **Hye-jin Shim***, Jungwoo Heo*, Jae-han Park, Ga-hui Lee, Ha-jin Yu, “Graph Attentive Feature Aggregation for Text-Independent Speaker Verification.”, *ICASSP*, 2022
- **Hye-jin Shim**, Jee-weon Jung, Ju-ho Kim, Ha-jin Yu, “Attentive Max Feature Map for Acoustic Scene Classification with Joint Learning considering the Abstraction of Classes.”, *ICASSP*, 2022
- Ju-ho Kim, **Hye-jin Shim**, Jungwoo Heo, Ha-Jin Yu, “RawNeXt: Speaker verification system for variable-duration utterances with deep layer aggregation and dynamic scaling policies.”, *ICASSP*, 2022
- 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.”, *ICASSP*, 2022
- 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.”, *Preprint*, 2022
- Jee-weon Jung, **Hye-jin Shim**, Ju-ho Kim, Ha-jin Yu, “DCASENet: An integrated pretrained deep neural network for detecting and classifying acoustic scenes and events.”, *ICASSP*, 2021
- **Hye-jin Shim**, Hee-soo Heo, Jee-weon Jung, Ha-jin Yu, “Self-supervised pre-training with acoustic configurations for replay spoofing detection.”, *INTERSPEECH*, 2020
- Jee-weon Jung, Seung-bin Kim, **Hye-jin Shim**, Ju-ho Kim, Ha-jin Yu, “Improved RawNet with Filter-wise Rescaling for Text-independent Speaker Verification using Raw Waveforms.”, *INTERSPEECH*, 2020
- Jee-weon Jung, **Hye-jin Shim**, Ju-ho Kim, Ha-jin Yu, “Acoustic scene classification using audio tagging.”, *INTERSPEECH*, 2020
- Seung-bin Kim, Jee-weon Jung, **Hye-jin Shim**, Ju-ho Kim, Ha-jin Yu, “Segment Aggregation for short utterances speaker verification using raw waveforms.”, *INTERSPEECH*, 2020
- Jee-weon Jung*, Ju-ho Kim*, **Hye-jin Shim**, Seung-bin Kim, Ha-jin Yu, “Selective Deep Speaker Embedding Enhancement for Speaker Verification.”, *Odyssey*, 2020
- 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.”, *ASRU*, 2019
- 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.”, *INTERSPEECH*, 2019
- Hee-soo Heo, Jee-weon Jung, **Hye-jin Shim**, Ha-jin Yu, “Acoustic scene classification using teacher-student learning with soft-labels.”, *INTERSPEECH*, 2019

- 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.", *INTERSPEECH*, 2019
- Hee-soo Heo, Jee-weon Jung, **Hye-jin Shim**, Ha-jin Yu, "End-to-end losses based on speaker basis vectors and all-speaker hard negative mining for speaker verification.", *INTERSPEECH*, 2019
- 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.", *INTERSPEECH*, 2019
- **Hye-jin Shim**, Jee-weon Jung, Hee-soo Heo, Sung-hyun Yoon, Ha-jin Yu, "Replay attack spoofing detection system using replay noise by multi-task learning.", *TAAI*, 2018
- 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.", *INTERSPEECH*, 2018
- Jee-weon Jung, Hee-soo Heo, IL-Ho Yang, Sung-hyun Yoon, **Hye-jin Shim**, Ha-jin Yu, "D-vector based speaker verification system using Raw Waveform CNN.", *ANIT*, 2017