SEONG-GYUN LEEM

⚲ Dallas, TX ‖ ☎ (+1)469-655-1655 ‖ ✉ sean.sgleem@gmail.com
⌂ https://sgleem.github.io/ in https://www.linkedin.com/in/seong-gyun-leem

RESEARCH INTERESTS

My main research goal is to **advance speech AI and develop empathetic speech intelligence for real-world applications**. To achieve this goal, I focus on addressing practical challenges in implementing spoken language understanding systems in mobile platforms and clinical settings.

EDUCATION

Ph.D. in Electrical and Computer Engineering   Aug 2020 - Oct 2024
The University of Texas at Dallas, Richardson, TX
Advisor: Carlos Busso
Thesis: "Speech emotion recognition in the presence of background noise"

Master's in Computer Science and Engineering   Mar 2018 - Feb 2020
Korea University, Seoul, South Korea
Advisor: Dongsuk Yook
Thesis: "Linear spectral feature transformation for the environment adaptation of speech interface (Korean)"

Bachelor's in Computer Science and Engineering   Mar 2012 - Feb 2018
Korea University, Seoul, South Korea

RESEARCH EXPERIENCES

The University of Texas at Dallas
Richardson, TX, United States
Research Assistant   Aug 2021 - Dec 2024
- Developed a robust fine-tuning framework to adapt speech representation models for emotion recognition tasks, leading to the publication of 15 papers in IEEE journals and top-tier signal processing conferences.
- Designed an innovative smartphone-based social isolation detection framework to assess loneliness and mood in real-world settings, demonstrating expertise in applied machine learning and behavioral analysis.
- Contributed to the creation and curation of a large-scale emotional speech database, MSP-Podcast 2.0, demonstrating expertise in real-world data curation.
- Mentored three undergrad and graduate students, publishing 3 papers in top-tier speech processing conferences.

Robert Bosch LLC
Pittsburgh, PA, United States
Deep Learning Research Intern - Audio AI   May 2024 - Aug 2024
- Designed and implemented a self-supervised learning framework for general-purpose audio representation, currently in the process of patent application.

Dolby Laboratories
San Francisco, CA, United States
Audio research intern   May 2023 - Aug 2023
- Developed a controllable speech-emotion conversion framework leveraging discrete audio tokens and HiFi-GAN, contributing to Dolby's research project.

Korea University
Seoul, South Korea
Research Staff   Mar 2020 - June 2020
Research Assistant   Mar 2018 - Feb 2020
- Designed and implemented a speech anonymization system using advanced voice conversion technologies.
- Designed and implemented flexible keyword spotting systems for voice user interface.
- Published 3 papers in IEEE journals and a top-tier signal processing conference.

TEACHING EXPERIENCES

- EESC 6360 **Digital Signal Processing 1**, The University of Texas at Dallas   Spring 2021
  Assist Dr. Carlos Busso as a teaching assistant

- ENGR 2300 **Linear Algebra for Engineers** at The University of Texas at Dallas,   Spring 2021
  Assist Dr. Mohammed Z Ali as a teaching assistant

- EESC 6366 **Signals and systems**, The University of Texas at Dallas · Fall 2020
  Assist Dr. Carlos Busso as a teaching assistant

- ENGR 2300 **Linear Algebra for Engineers**, The University of Texas at Dallas · Fall 2020
  Assist Dr. Shaheen Ahmed as a teaching assistant

- COSE361 **Artificial Intelligence**, Korea University · Spring 2019
  Assist Dr. Dongsuk Yook as a teaching assistant

- COSE362 **Machine Learning**, Korea University · Fall 2018
  Assist Dr. Dongsuk Yook as a teaching assistant

- COSE362 **Machine Learning**, Korea University · Spring 2018
  Assist Dr. Dongsuk Yook as a teaching assistant

PROFESSIONAL ACTIVITIES

- Reviewer: IEEE Transactions on Pattern Analysis and Machine Intelligence (2022), IEEE/ACM Transactions on Audio, Speech, and Language Processing (2023, 2024), IEEE Transactions on Multimedia (2022), IEEE Transactions on Affective Computing (2024-a, 2024-b, 2024-c), The Computer Journal, IEEE International Conference on Acoustics, Speech and Signal Processing (2024, 2025)

- Judge: ACM UTD Research Symposium (Spring/Fall 2024), *Great Minds in STEM* (GMiS) conference (2024).

AWARDS

- Received *ECE Department Teaching Assistant Awards*, University of Texas at Dallas, in 2021.

PUBLICATIONS

Journal Articles

1. L. Goncalves, **S. Leem**, W. Lin, B. Sisman, and C. Busso *"Versatile audiovisual learning for handling single and multi modalities in emotion regression and classification tasks,"* IEEE Transactions on Affective Computing, vol. to appear, 2024. (arXiv:2305.07216).

2. H. Chou, L. Goncalves, **S. Leem**, A. Salman, C. Lee and C. Busso, *"Minority Views Matter: Evaluating Speech Emotion Classifiers with Human Subjective Annotations by an All-Inclusive Aggregation Rule,"* IEEE Transactions on Affective Computing, 2024, doi: 10.1109/TAFFC.2024.3411290.

3. **S. Leem**, D. Fulford, J.P. Onnela, D. Gard and C. Busso, *"Selective acoustic feature enhancement for speech emotion recognition with noisy speech,"* IEEE/ACM Transactions on Audio, Speech and Language Processing, vol. 32, pp. 917-929, 2024.

4. J. Harvill, **S. Leem**, M. Abdelwahab, R. Lotfian and C. Busso, *"Quantifying Emotional Similarity in Speech,"* in IEEE Transactions on Affective Computing, doi: 10.1109/TAFFC.2021.3127390.

5. I. Yoo, K. Lee, **S. Leem**, H. Oh, B. Ko and D. Yook, *"Speaker Anonymization for Personal Information Protection Using Voice Conversion Techniques."* IEEE Access 8 (2020): 198637-198645.

6. **S. Leem**, I. Yoo and D. Yook, *"Multitask Learning of Deep Neural Network-Based Keyword Spotting for IoT Devices,"* IEEE Transactions on Consumer Electronics, vol. 65, no. 2, pp. 188-194, May 2019.

Conference Proceedings

1. J. Tzeng, **S. Leem**, A. Salman, C. Lee, C. Busso, *"Noise-Robust Speech Emotion Recognition Using Shared Self-Supervised Representations with Integrated Speech Enhancement,"* in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2025), To appear.

2. H.-C. Chou, H. Wu, L. Goncalves, **S. Leem**, A. N. Salman, C. Busso, H.-Y Lee, and C.-C. Lee, "Embracing ambiguity and subjectivity using the all-inclusive aggregation rule for evaluating multi-label speech emotion recognition systems," in IEEE Spoken Language Technology Workshop (SLT 2024), Macao, China, December 2024.

3. **S. Leem**, D. Fulford, J.P. Onnela, D. Gard and C. Busso, *"Keep, Delete, or Substitute: Frame Selection Strategy for Noise-Robust Speech Emotion Recognition,"* in Interspeech 2024, Kos Island, Greece, September 2024, pp. 3734-3738.

4. A. Naini, S. Subramanium, **S. Leem**, and C. Busso, *"Combining relative and absolute learning formulations to predict emotional attributes from speech,"* IEEE Automatic Speech Recognition and Understanding Workshop (ASRU 2023), Taipei, Taiwan, December 2023.

5. L. Martinez-Lucas, A. Salman, **S. Leem**, S. Upadhyay, C. Lee and C. Busso, *"Analyzing the effect of affective priming on emotional annotations,"* in International Conference on Affective Computing and Intelligent Interaction (ACII 2023), Cambridge, MA, USA, September 2023.

6. **S. Leem**, D. Fulford, J.P. Onnela, D. Gard and C. Busso, *"Computation and memory efficient noise adaptation of Wav2Vec2.0 for noisy speech emotion recognition with skip connection adapters,"* in Interspeech 2023, Dublin, Ireland, August 2023, pp. 1888-1892.

7. H. Chou, L. Goncalves, **S. Leem**, C. Lee and C. Busso, *"The importance of calibration: Rethinking confidence and performance of speech multi-label emotion classifiers,"* in Interspeech 2023, Dublin, Ireland, August 2023, pp. 641-645.

8. **S. Leem**, D. Fulford, J.P. Onnela, D. Gard and C. Busso, *"Adapting a self-supervised speech representation for noisy speech emotion recognition by using contrastive teacher-student learning,"* in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2023), Rhodes Island, Greece, 2023.

9. **S. Leem**, D. Fulford, J.P. Onnela, D. Gard and C. Busso, *"Not All Features Are Equal: Selection of Robust Features for Speech Emotion Recognition in Noisy Environments,"* in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2022), Singapore, May 2022.

10. **S. Leem**, D. Fulford, J.P. Onnela, D. Gard and C. Busso, *"Separation of emotional and reconstruction embeddings on ladder network to improve speech emotion recognition robustness in noisy conditions,"* in Interspeech 2021, Brno, Czech Republic, August-September 2021.

11. D. Yook, **S. Leem**, K.Lee and I. Yoo, *"Many-to-Many Voice Conversion Using Cycle-Consistent Variational Autoencoder with Multiple Decoders,"* Proc. Odyssey 2020 The Speaker and Language Recognition Workshop, pp. 215-221, 2020.

ArXiv Papers

1. **S. Leem**, D. Fulford, J.P. Onnela, D. Gard and C. Busso, *"Describe Where You Are: Improving Noise-Robustness for Speech Emotion Recognition with Text Description of the Environment,"* arXiv preprint, arXiv:2407.17716 (2024).

<u>Under review</u>

1. **S. Leem**, D. Fulford, J.P. Onnela, D. Gard and C. Busso, *"Describe Where You Are: Improving Noise-Robustness for Speech Emotion Recognition with Text Description of the Environment,"* IEEE Transactions on Affective Computing, <u>Under Review</u>.

2. L. Martinez-Lucas, A. Salman, **S. Leem**, W. Chien, S. Upadhyay, C. Lee and C. Busso, *"Affective Priming in Emotional Annotations and its Effect on Speech Emotion Recognition,"* IEEE Transactions on Affective Computing, <u>Under Review</u>.