Soonshin Seo

[Personal Details & Contact Information]

- 1993.12.18, Male
- Contact: ssseo@sogang.ac.kr
- Website: https://github.com/sunshines14

[Research Interests]

- Automatic Speech Recognition
- Automatic Speaker Verification
- Acoustic Scene Classification
- Sound Event Detection

[Education]

- **Ph.D. Candidate (2021.09 ~)**
  - Auditory Intelligence Lab., Computer Science and Engineering, Sogang University, South Korea
  - Advisor: Prof. Ji-Hwan Kim
- **Ph.D. Student (Joint Master. & Ph.D. Program) (2018.09 ~ 2021.08)**
  - Auditory Intelligence Lab., Computer Science and Engineering, Sogang University, South Korea
  - Advisor: Prof. Ji-Hwan Kim
  - CGPA (3.60/4.30)
- **Undergraduate Intern (2017.01 ~ 2018.08)**
  - Auditory Intelligence Lab., Computer Science and Engineering, Sogang University, South Korea
  - Advisor: Ji-Hwan Kim
- **B.E. degree (2014.03 ~ 2018.08)**
  - Linguistics & Computer Science and Engineering (Double Major), Hankuk University of Foreign Studies, South Korea
  - CGPA (3.09/4.50)

[Publications]

- **International Journals**
  - Donghyun Lee, Hosung Park, **Soonshin Seo**, Hyunsoo Son, Gyujin Kim, and Ji-Hwan Kim, "Robustness of Differentiable Neural Computer Using Limited Retention Vector-based Memory Deallocation in Language Model", KSII Transactions on Internet and Information Systems, 2021. (SCIE, IF 0.648, Accepted)
  - Donghyun Lee, Hosung Park, **Soonshin Seo**, Changmin Kim, Hyunsoo Son, Gyujin Kim, and Ji-Hwan Kim, "Language Model Using Differentiable Neural Computer Based on Forget Gate-based Memory Deallocation", *Computer, Materials & Continua*, 2021. (SCIE, IF 4.89, Published)
  - **Soonshin Seo**, Ji-Hwan Kim, "Self-Attentive Multi-Layer Aggregation with Feature Recalibration and Normalization for Text-Independent Speaker Verification System", *Electronics*, 9(10), 2020. (SCIE, IF 2.412, Published)

· Hosung Park, **Soonshin Seo**, Changmin Kim, Hyunsoo Son, Ji-Hwan Kim, "Hybrid CTC-attention networkbased end-to-end speech recognition system for Korean language", *Journal of Web Engineering*, 2020. (SCIE, IF 0.396, Under Review)

■ **Domestic Journals**

· **Soonshin Seo**, Ji-Hwan Kim, "Masked Cross Self-Attentive Encoding for Speaker Embedding", *The Journal of the Acoustical Society of Korea*, 39(5), 2020. (SCOPUS, Published)

■ **International Conferences**

· Gyujin Kim, **Soonshin Seo**, Donghyun Lee, Hosung Park, Changmin Kim, Hyunsoo Son, and Ji-Hwan Kim, "Metric Learning-based Multilevel Parameter Adaptation for Converted and Synthesized Speech Spoofing Detection", in *Proceedings of the International Workshop in Smart Info-Media Systems in Asia (SISA)*, 2020. (Oral)

· **Soonshin Seo**, Changmin Kim, Donghyun Lee, Hosung Park, Hyunsoon Son, Gyujin Kim and Ji-Hwan Kim, "Acoustic Scene Classification System in Multi-Device Environment Using Frequency-Tuned Spectrogram and Residual Convolutional Neural Networks", in *Proceedings of the International Conference on Electronics, Electrical Engineering, Computer Science (EEECS)*, 2020. (Oral)

· **Soonshin Seo**, Changmin Kim, and Ji-Hwan Kim, "Multi-Channel Feature using Inter-Class and Inter-Device Standard Deviations for Acoustic Scene Classification", in *Detection and Classification of Acoustic Scenes and Events Challenge (DCASE Challenge)*, 2020. (Technical Report)

· **Soonshin Seo**, Daniel Jun Rim, Minkyu Lim, Donghyun Lee, Hosung Park, and Ji-Hwan Kim, "Robust Speaker Verification System in Vehicle Driving Environment", in *Proceedings of the Seoul International Conference on Speech Sciences (SICSS)*, 2019. (Oral)

· **Soonshin Seo**, Daniel Jun Rim, Minkyu Lim, Donghyun Lee, Hosung Park, Junseok Oh, Changmin Kim, and Ji-Hwan Kim, "Shortcut Connections based Deep Speaker Embeddings for End-to-End Speaker Verification System", in *Proceedings of the Annual Conference of the International Speech Communication Association (INTERSPEECH)*, 2019. (Poster)

· Hosung Park, **Soonshin Seo**, Daniel Jun Rim, Changmin Kim, Hyunsoo Son, Jeong-Sik Park, and Ji-Hwan Kim, "Korean Grapheme Unit-based Speech Recognition Using Attention-CTC Ensemble Network", in *Proceedings of the International Symposium on Multimedia and Communications (ISMAC),* 2019. (Oral)

· **Soonshin Seo**, Minkyu Lim, Donghyun Lee, Hosung Park, Junseok Oh, Daniel Jun Rim, and Ji-Hwan Kim, "Environmental Noise Robustness for Korean Fricatives using Speech Enhancement Generative Adversarial Networks", in *Proceedings of the IEEE International Conference on Big Data and Smart Computing (IEEE BigComp)*, 2019. (Oral)

· Hosung Park, **Soonshin Seo**, Minkyu Lim, Donghyun Lee, Yoseb Kang, Juneseok Oh, and Ji-Hwan Kim, "Implementation of Korean Grapheme-to-Phoneme Rules with Morpheme Analysis", in *Proceedings of the International Conference on Electronics, Electrical Engineering, Computer Science (EEECS)*, 2018. (Oral)

· Hosung Park, **Soonshin Seo**, Minkyu Lim, Donghyun Lee, Yoseb Kang, Juneseok Oh, and Ji-Hwan Kim, "Sequence-to-Sequence Korean Phoneme-to-Text Conversion for Korean Speech Recognition", in *Proceedings of the International Conference on Electronics, Electrical Engineering, Computer Science (EEECS)*, 2018. (Oral)

· **Soonshin Seo**, Hosung Park, Minkyu Lim, Donghyun Lee, and Ji-Hwan Kim, "CMVN based Noise Processing for Unvoiced Sound /ㅅ/ in Korean", in *Proceedings of the Seoul International Conference on Speech Sciences (SICSS)*, 2017. (Poster)

- **Domestic Conferences**

  · Hosung Park, **Soonshin Seo**, Hyunsoo Son, Changmin Kim, and Ji-Hwan Kim, "Self-attentive Layer for Discriminant Vector Training in Low-resource Speech Recognition", in *Proceedings of the Korea Computer Congress*, 2020. (Oral, Best Paper)

  · **Soonshin Seo**, Minkyu Lim, Donghyun Lee, Yoseb Kang, Juneseok Oh, and Ji-Hwan Kim, "Performance Enhancement of Speech Recognition System using Noisy Speech", in *Proceedings of the Korean Society of Speech Sciences*, 2018. (Poster)

  · Soonshin Seo, Hosung Park, Donghyun Lee, Minkyu Lim, Yoseb Kang, and Ji-Hwan Kim, "Implementation of Noisy Speech Generation System for Acoustic Model Performance Improvement in Noisy Environment", in *Proceedings of the Korean Society of Speech Sciences*, 2017. (Poster)

[Research Experience]

- **Development of Data Augmentation Technology by using Heterogeneous Information and Data Fusion (2020.04 ~ on going)**

  · Member of the Speech and Audio Team

  · Supported by the Ministry of Science and ICT

  · Research Details

    · Korean Speech Recognition System in YouTube Traveling Vlog Environments

    · Sound Event Detection System in YouTube Traveling Vlog Environments

- **Development of Human Enhancement Technology for Auditory and Muscle Support (2020.05 ~ 2020.11)**

  · Member of the Speech and Audio Team

  · Supported by the Ministry of Science and ICT

  · Research Details

    · Real Time Sound Event Detection System under Different Mobile Devices

      · Data Collection from YouTube/Preparation/Augmentation

      · Log Mel-Spectrogram, Delta and Delta-Delta features

      · Residual CNN based Classifier using Late Fusion

      · Real-Time Android Test (Google Pixel, LG V50, Samsung Galaxy S7, Apple iPhone SE)

- **Technical development of Korean Speech Recognition System in Vehicle (2018.09 ~ 2019.12.31)**

  · Member of the Speech and Audio Team

  · Supported by the Ministry of Trade, Industry and Energy

  · Research Details

    · Automatic Korean Speech Recognition System in Vehicle Driving

      · TDNN-based Acoustic Model

      · N-Gram-based Language Model Adaptation

        - WFST-based Decoding Networks
    - Korean Speaker Verification System in Vehicle Driving
        - Data Preparation for Vehicle Driving
        - Data Augmentation using Noise Mixing and SpecAugment
        - Voice Activity Detection
        - L2-Normalization
        - Multi-Utterances Enrollment
    - Residual CNN-based Deep Speaker Embedding
        - Additional Identity Mapping
        - Shortcut Connections based multiple pooling
        - Random Masking Method
        - Cross Self-Attention Module

- **Development of QA systems for Video Story Understanding to Pass the Video Turing Test (2018.09 ~ on going)**
    - Member of the Speech and Audio Team
    - Supported by the Ministry of Science and ICT
    - Research Details
        - Korean Real Time Speech Recognition System
            - Online Decoding Parameters Optimization
            - N-Gram-based Language Model Compression
        - Automatic Transcription Generation System
            - Weakly Labeled Data Collected in YouTube
            - DNN-based Forced Alignment
- **Development of Distant Speech Recognition and Multi-Task Dialog Processing Technologies for In-Door Conversational Robots (2017.01 ~ 2020.05)**
    - Member of the Speech and Audio Team
    - Supported by the Ministry of Trade, Industry and Energy
    - Research Details
        - Noisy Speech Generation System for Environmental Noise Data
        - Korean Speech Recognition System for Multi-Channel

[Achievements]

- **Recipient of the 13th Place in Task 1a of the IEEE AASP Challenge on Detection and Classification of Acoustic Scenes and Events (2020.07)**
    - Task 1a : Acoustic Scene Classification with Multiple Devices

- **Recipient of the 5th Place in S/W Implementation & Demo Challenge (2017.12)**
  - From the Korea Software Congress 2017

[Teaching Experience]

- **Speech Recognition Course (Advenced Level) (2018.11)(2019.07)(2019.08)(2020.07)**
  - Teaching Assistant
  - At Samsung Electronics Leadership Center
- **Speech Recognition Course (Advenced Level) (2019.02)**
  - Teaching Assistant
  - At LG Electronics Seocho R&D Campus
- **The 36th Speech Communication and Signal Processing Conference (2019.08)**
  - Tutorial Speaker
  - Subject: Implementation of Speech Recognition System Using Kaldi Toolkit
  - At University of Seoul
- **Introduction to Dialogue-based User Interface (2019 Fall Semester)**
  - Teaching Assistant
  - At Sogang University

[Graduate Coursework]

- **2020 2ND Semester**
  - Pattern Recognition
  - Special Topics on Statistical Signal Processing
  - AI System Architecture
- **2020 1ST Semester**
  - Special Study
  - Analysis and Design of Speech Recognitions Systems
  - Pattern Recognition
  - Topics in Computer Network
- **2019 2ND Semester**
  - Special Study
- **2019 1ST Semester**
  - Special Study
  - Speech Processing
  - Natural Language Processing
- **2018 2ND Semester**

- Artificial Intelligence II
- Intellectual Property and Patent
- Introduction to Dialogue-based User Interface
- Advanced Neural Networks

[Technical Skills]

- Python, Linux, Shell Script, C++
- Kaldi, PyTorch, Keras with Tensorflow 2.0