Li-Wei Chen

 \Box +1 412 636 3999 ☑ lwchen.jeff@gmail.com github.com/b04901014 **☑** Google Scholar

Research Interests

- Speech processing, especially voice conversion (VC), text-to-speech synthesis (TTS), self-supervised representation learning (SSL), and speech emotion recognition (SER)
- NLP tasks including natural language understanding, language modeling and task-oriented conversation
- Design machine learning systems for the above tasks or improve existing algorithms

Education

2022-Present

Ph.D. in Language and Information Technology, Carnegie Mellon University, Pittsburgh

2020-2022 Sept. Aug.

Master of Language Technologies, Carnegie Mellon University, Pittsburgh

Overall GPA - 4.23/4.33

2015-2020 Sept. Jan.

Bachelor of Electronic Engineering, National Taiwan University, Taipei

Overall GPA - 4.10/4.3 - 3.96/4.0

Research Experience

2022–Present Ph.D. Student, Carnegie Mellon University, Pittsburgh

- O Advisor: Alexander Rudnicky, Shinji Watanabe
- Developed an auto-regressive TTS system for real-world spontaneous speech that outperforms existing methods in terms of intelligibility, naturalness, and diversity
- O Proposed a unified system for one-shot voice conversion (VC) on pitch, rhythm, and speaker attributes of speech, outperforming existing works

2024-2024 May. Aug.

Research Intern, Apple Machine Learning Research, Cupertino, CA

- Manager: Zakaria Aldeneh
 - See our paper: Exploring Prediction Targets in Masked Pre-Training for Speech Foundation Models.

2023-2023 May. Aug.

Student Researcher, Google Research, Mountain View, CA

 Manager: Soroosh Mariooryad, RJ Skerry-Ryan Topic: Joint Audio-Visual Synthesis

2020-2022 Aug.

Master's Student, Carnegie Mellon University, Pittsburgh

- Advisor: Alexander Rudnicky
- Investigated the importance of temporal context and different features for speech emotion recognition
- Improved the performance on speech emotion recognition by a large margin by exploring different fine-tuning techniques of self-supervised pretrained models
- Designed a novel architecture with cross-modality attention mechanism to realize fine-grained style control on the transformer-based text-to-speech synthesis (TTS)

2018-2018

Research Intern, Taiwan Al Labs, Taipei

- O Responsible for the TTS part of a smart speaker product, using Tacotron2 for speech synthesize
- Implemented parallel WaveNet as vocoder replacing WaveNet to significantly reduce latency

2017-2019

Undergraduate Student, National Taiwan University, Taipei

- O Advisor: Hung-Yi Lee, Lin-Shan Lee
- Developed an end-to-end spoken term detection system with attention-based CNN
- Proposed a model using GANs to transform impaired speech into normal one while preserving linguistic content
- Achieved better results in terms of Mean Opinion Score compared to existing models

Competition Experience

2021–2022 Alexa Prize Taskbot Competition, Carnegie Mellon University, Pittsburgh May. May.

- Advisor: Alexander Rudnicky
- O Technical team leader of a 8 member team, mainly responsible for the software engineering of the dialogue logic
- Our system advanced to the semi-finals, the report is available here

2021–2021

Zero Resource Speech Challenge 2021, Carnegie Mellon University, Pittsburgh

- O Advisor: Alexander Rudnicky, Shinji Watanabe
- Collaborated with Yahoo! JAPAN to develop a algorithm on learning unsupervised speech representation, achieving top result in one of the metrics
- Paper accepted to Interspeech (see in publications)

Publications

- [1] L. Chen, T. Higuchi. H. Bai. A. H. Abdelaziz, A. Rudnicky, S. Watanabe, T. Likhomanenko, B. Theobald, Z. Aldeneh, "Exploring Prediction Targets in Masked Pre-Training for Speech Foundation Models", in Proceedings of ICASSP 2025, 2025.
- [2] Z. Aldeneh, T. Higuchi, J. Jung, L. Chen, S. Shum, A. H. Abdelaziz, S. Watanabe, T. Likhomanenko, B. Theobald: Unsupervised Learning of Speaker Characteristics with i-Vector based Pseudo-Labels", in Proceedings of ICASSP 2025, 2025.
- [3] T. Chi, T. Fan, L. Chen, A. Rudnicky, P. Ramadge, "Latent Positional Information is in the Self-Attention Variance of Transformer Language Models Without Positional Embeddings", in Proceedings of ACL 2023, 2023.
- [4] L. Chen, S. Watanabe, A. Rudnicky, "A Unified One-Shot Prosody and Speaker Conversion System with Self-Supervised Discrete Speech Units", in Proceedings of ICASSP 2023, 2023.
- [5] L. Chen, A. Rudnicky, "Exploring Wav2vec 2.0 fine-tuning for improved speech emotion recognition", in Proceedings of ICASSP 2023, 2023.
- [6] P. Wu, L. Chen, C. Cho, S. Watanabe, L. Goldstein, A. Black, G. Anumanchipalli, "Speaker-Independent Acoustic-to-Articulatory Speech Inversion", in Proceedings of ICASSP 2023, 2023.
- [7] L. Chen, A. Rudnicky, S. Watanabe, "A Vector Quantized Approach for Text to Speech Synthesis on Real-World Spontaneous Speech", in Proceedings of AAAI 2023, 2023.
- [8] L. Chen, A. Rudnicky, "Fine-grained style control in Transformer-based Text-to-speech Synthesis", in Proceedings of ICASSP 2022, 2022.
- [9] Y. Xia*, L. Chen*, A. Rudnicky, R. M. Stern, "Temporal Context in Speech Emotion Recognition", in Proceedings of the Interspeech 2021, 2021. (* Equal contribution)
- [10] T. Maekaku, X. Chang, Y. Fujita, L. Chen, S. Watanabe, A. Rudnicky, "Speech Representation Learning Combining Conformer CPC with Deep Cluster for the ZeroSpeech Challenge 2021", in Proceedings of the Interspeech 2021, 2021.
- [11] L. Chen, H. Lee, and Y. Tsao, "Generative Adversarial Networks for Unpaired Voice Transformation on Impaired Speech", in *Proceedings of the Interspeech 2019*, 2019.

Teaching Experience

Jan.

2022–2022 **Teaching Assistant**, Carnegie Mellon University, Pittsburgh

- O Course: Multimodal Machine Learning (Course number 11-777, 2022 Spring by Yonatan Bisk)
- O Course Website: https://yonatanbisk.com/teaching/mmml-s22/
- Graded and advised each team regarding their final project, responsible for teams that involves speech modality in their project

2018-2018 Jan. July

Teaching Assistant, National Taiwan University, Taipei

- O Course: Machine Learning and Having it Deep and Structured (2018 Spring by Hung-Yi Lee)
- O Course Website: https://speech.ee.ntu.edu.tw/~hylee/mlds/2018-spring.php
- O Designed and graded assignments of Reinforcement Learning, including Policy Gradient, Deep Qlearning Networks, Actor-Critic, and other improved algorithms of them