

Li-Wei Chen

+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
Sept. **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
Sept. **Ph.D. Student**, *Carnegie Mellon University*, Pittsburgh
- 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
 - Proposed a unified system for one-shot voice conversion (VC) on pitch, rhythm, and speaker attributes of speech, outperforming existing works
- 2020–2022
Sept. 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
Aug. Sept. **Research Intern**, *Taiwan AI Labs*, Taipei
- Responsible for the TTS part of a smart speaker product, using Tacotron2 for speech synthesizer
 - Implemented parallel WaveNet as vocoder replacing WaveNet to significantly reduce latency
- 2017–2019
Sept. Sept. **Undergraduate Student**, *National Taiwan University*, Taipei
- 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
May. May. **Alexa Prize Taskbot Competition**, *Carnegie Mellon University*, Pittsburgh
- Advisor: Alexander Rudnicky
 - 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, see our paper (in publications) for the details of our system
- 2021–2021
Jan. May **Zero Resource Speech Challenge 2021**, *Carnegie Mellon University*, Pittsburgh
- Advisor: Alexander Rudnicky, Shinji Watanabe
 - Collaborated with Yahoo! JAPAN to develop an 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**, S. Watanabe, A. Rudnicky, "A Unified One-Shot Prosody and Speaker Conversion System with Self-Supervised Discrete Speech Units", *arXiv preprint arXiv:2211.06535*, 2022. (Accepted to ICASSP 2023)
- [2] **L. Chen**, A. Rudnicky, "Exploring Wav2vec 2.0 fine-tuning for improved speech emotion recognition", *arXiv preprint arXiv:2110.06309*, 2021. (Accepted to ICASSP 2023)
- [3] P. Wu, **L. Chen**, C. Cho, S. Watanabe, L. Goldstein, A. Black, G. Anumanchipalli, "Speaker-Independent Acoustic-to-Articulatory Speech Inversion", *arXiv preprint arXiv:2302.06774*, 2023. (Accepted to ICASSP 2023)
- [4] **L. Chen**, A. Rudnicky, S. Watanabe, "A Vector Quantized Approach for Text to Speech Synthesis on Real-World Spontaneous Speech", *arXiv preprint arXiv:2302.04215*, 2023. (Accepted to AAI 2023 Main Track)
- [5] **L. Chen**, A. Rudnicky, "Fine-grained style control in Transformer-based Text-to-speech Synthesis", in *Proceedings of ICASSP 2022*, 2022.
- [6] **L. Chen**, T. Chi, D. Shah, C. Gomes, J. Bao, K. Ganesan, P. Joshi, S. Kumar, D. Naik, J. Hagerty, A. Rudnicky, "Tartan: A taskbot that assists with recipes and do-it-yourself projects", *Alexa Prize TaskBot Challenge Proceedings*, 2022.
- [7] Y. Xia*, **L. Chen***, A. Rudnicky, R. M. Stern, "Temporal Context in Speech Emotion Recognition", in *Proceedings of the Interspeech 2021*, 2021. (* Equal contribution)
- [8] 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.
- [9] **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

- 2022–2022 **Teaching Assistant**, *Carnegie Mellon University*, Pittsburgh
Jan. May
- Course: Multimodal Machine Learning (*Course number 11-777, 2022 Spring by Yonatan Bisk*)
 - Course Website: <https://yonatanbisk.com/teaching/mml-s22/>
 - Graded and advised each team regarding their final project, responsible for teams that involves speech modality in their project
- 2018–2018 **Teaching Assistant**, *National Taiwan University*, Taipei
Jan. July
- Course: Machine Learning and Having it Deep and Structured (*2018 Spring by Hung-Yi Lee*)
 - Course Website: <https://speech.ee.ntu.edu.tw/~hylee/mls/2018-spring.php>
 - Designed and graded assignments of Reinforcement Learning, including Policy Gradient, Deep Q-learning Networks, Actor-Critic, and other improved algorithms of them

Computer Skills

- Basic JAVASCRIPT, HTML, MATLAB
- Intermediate C#, L^AT_EX, Linux, Bash, Java
- Advanced PYTHON, C++
- Others
- Speech Toolkits: Kaldi, World, HTK, SPTK
 - Machine Learning Libraries: Pytorch, Tensorflow, Keras

Language Proficiency

- Native Mandarin Chinese, Taiwanese Hokkien
- GRE General Q: 170, V: 161, AWA: 4.0
- TOEFL iBT 106 (30/29/22/25)