Ruchao Fan

Affiliate: Department of Electrical and Computer Engineering, UCLA
Email: fanruchao@g.ucla.edu Personal Page Google Scholar

Research Interests

Speech Processing, Self-supervised and Unsupervised Learning, Domain Adaptation, Data augmentation, End-to-end Speech Recognition (Non-autoregressive transformer, CTC, RNN-T), Children's ASR

Education

University of California, Los Angeles (UCLA) **Los Angeles, U.S.A.**
Ph.D. in Electrical and Computer Engineering *Sept. 2019 - Present*

Beijing University of Posts and Telecommunications (BUPT) **Beijing, China**
M.S. in Information and Communication Engineering *Sept. 2016 - Jun. 2019*
B.Eng. in Communication Engineering *Sept. 2012 - Jun. 2016*

Professional Experience

UCLA - Speech Processing and Auditory Perception Lab **Los Angeles, U.S.A.**
Graduate Student Researcher - Advisor: Prof. Abeer Alwan *Sept. 2019 - Present*
Research on children's ASR as a low-resource task
- Develop data augmentation and model pre-training methods to improve children's ASR
- Develop non-autoregressive models for child ASR to improve decoding efficiency

Microsoft Corporation. - MSR Redmond **Redmond, U.S.A.**
Research Intern - Mentors: Dr. Yiming Wang, Yashesh Gaur, and Jinyu Li *June 2022 - Sept. 2022*
Research on self-supervised learning algorithms
- Develop a novel CTC training mechanism to solve the misalignment problem.
- Proposed methods consistently improve the performance compared to CE training.

Microsoft Corporation. - MSR Redmond **Redmond, U.S.A.**
Research Intern - Mentors: Dr. Guoli Ye and Dr. Jinyu Li *June 2021 - Sept. 2021*
Research on spell correction for end-to-end speech recognition
- Investigated methods of non-autoregressive spell correction for a transformer-transducer
- Proposed a sampling-based error detection method for spell correction

PAII Inc. - US Research Lab **Palo Alto, U.S.A.**
Research Scientist Intern - Mentors: Dr. Wei Chu and Dr. Peng Chang *June 2020 - Sept. 2020*
Research on non-autoregressive transformers for end-to-end speech recognition
- Used CTC alignment as extra information for token-level acoustic embedding extraction
- Proposed an error-based sampling method during inference to improve performance

Sogou Inc. - Voice Interaction Technology Center **Beijing, China**
Research Intern - Mentors: Dr. Pan Zhou and Dr. Wei Chen *Apr. 2018 - Aug. 2019*
Research on attention-based encoder-decoder (AED) end-to-end speech recognition
- Proposed an online AED with 3.5% relative WER degradation compared to an offline AED
- Improved the performance of speech-transformer with parallel-schedule sampling and relative positional encoding

Publications

[16] **R. Fan**, W. Chu, P. Chang, A. Alwan, “A CTC Alignment-based Non-autoregressive Transformer for End-to-end Automatic Speech Recognition”, Accepted to Transactions on Audio, Speech and Language Processing.

[15] **R. Fan**, Y. Wang, Y. Gaur, and J. Li, “CTCBERT: Advancing Hidden-unit BERT with CTC Objectives”, Accepted to ICASSP 2023.

[14] **R. Fan**, G. Ye, Y. Gaur, and J. Li, “Acoustic-aware Non-autoregressive Spell Correction with Mask Sample Decoding”, arXiv preprint arXiv:2210.08665.

[13] **R. Fan**, Y. Zhu, J. Wang, and A. Alwan, "Towards Better Domain Adaptation for Self-Supervised Models: A Case Study of Child ASR," in IEEE Journal of Selected Topics in Signal Processing, vol. 16, no. 6, pp. 1242-1252, Oct. 2022, doi: 10.1109/JSTSP.2022.3200910..

[12] **R. Fan**, and A. Alwan, “DRAFT: A Novel Framework to Reduce Domain Shifting in Self-supervised Learning and Its Application to Children’s ASR”, Proc. Interspeech 2022, pp. 4900-4904.

[11] A. Johnson, **R. Fan**, R. Morris, and A. Alwan, "LPC AUGMENT: An LPC-Based ASR Data Augmentation Algorithm for Low and Zero-Resource Children’s Dialects," in ICASSP 2022, IEEE, pp. 8577--8581.

[10] Y. Zhu, **R. Fan**, and A. Alwan, "Towards Better Meta-Initialization with Task Augmentation for Kindergarten-aged Speech Recognition," in ICASSP 2022, IEEE, pp. 8582--8586.

[9] G. Yeung, **R. Fan**, and A. Alwan, “Fundamental frequency feature warping for frequency normalization and data augmentation in child automatic speech recognition,” Speech Communication, 2021, doi: https://doi.org/10.1016/j.specom.2021.08.002.

[8] **R. Fan**, W. Chu, P. Chang, J. Xiao and A. Alwan, "An Improved Single Step Non-autoregressive Transformer for Automatic Speech Recognition," Proc. Interspeech 2021, pp. 3715-3719.

[7] J. Wang, Y. Zhu, **R. Fan**, W. Chu and A. Alwan, "Low Resource German ASR with Untranscribed Data Spoken by Non-native Children–INTERSPEECH 2021 Shared Task SPAPL System," Proc. Interspeech 2021, pp. 1279-1283.

[6] **R. Fan**, W. Chu, P. Chang, and J. Xiao, “CASS-NAT: CTC Alignment-based Single Step Non-autoregressive Transformer for Speech Recognition,” in ICASSP 2021, IEEE, pp. 5889–5893.

[5] **R. Fan**, A. Afshan and A. Alwan, "BI-APC: Bidirectional Autoregressive Predictive Coding for Unsupervised Pre-training and its Application to Children’s ASR," in ICASSP 2021, IEEE, pp. 7023–7027.

[4] G. Yeung, **R. Fan**, and A. Alwan, "Fundamental Frequency Feature Normalization and Data Augmentation for Child Speech Recognition," in ICASSP 2021, IEEE, pp. 6993-6997.

[3] V. Ravi, **R. Fan**, A. Afshan, H. Lu, and A. Alwan, “Exploring the use of an unsupervised autoregressive model as a shared encoder for text-dependent speaker verification,” Proc. Interspeech 2020, pp. 766–770, 2020.

[2] **R. Fan**, P. Zhou, W. Chen, J. Jia, and G. Liu, “An Online Attention-Based Model for Speech Recognition,” in Proc. Interspeech 2019, 2019, pp. 4390–4394.

[1] **R. Fan** and G. Liu, “CNN-based audio front end processing on speech recognition,” International Conference on Audio, Language and Image Processing. IEEE, 2018, pp. 349–354.

Skills and Coursework

Computer Languages and Open-source Framework

- Python, C/C++, Shell and Matlab, Pytorch, Tensorflow, Kaldi, Espnet, Fairseq;

Coursework

- ECE: Matrix Analysis, Digital Speech Processing, Advanced Topics in Speech Processing, Linear Programming; CS: Natural Language Generation, Algorithmic Machine Learning
- Overall GPA: 3.93/4.0

Honors and Awards

- UCLA-Amazon Science Hub Fellowship, honored as Amazon Fellow (2021-2022)
- One-year research funding from PAII Inc. (2020-2021)
- A first-year graduate fellowship for the ECE Ph.D. program at UCLA (2019-2020)
- Huawei Enterprise Scholarship (2017)
- Graduate Academic Scholarship (2016-2018)
- Undergraduate National Inspirational Scholarship (2013-2015)