

# Ruchao Fan

SPAPL 63-134 Engineering IV Building, 420 Westwood Plaza, Los Angeles, CA 90095

🌐 Homepage

✉ fanruchao@g.ucla.edu

🐙 Github

📞 phone

## Research Interests

---

Speech processing, End-to-end Speech Recognition, ASR for Children, Unsupervised Representation Learning, Domain Adaptation, Speaker Verification.

## Education

---

**University of California, Los Angeles (UCLA)**

*Ph.D. student in Electrical and Computer Engineering*

**Los Angeles, U.S.A.**

*Sept. 2019 - Present*

**Beijing University of Posts and Telecommunications (BUPT)**

*M.S. in Information and Communication Engineering*

**Beijing, China**

*Sept. 2016 - Jun. 2019*

**Beijing University of Posts and Telecommunications (BUPT)**

*B.Eng. in Communication Engineering*

**Beijing, China**

*Sept. 2012 - Jun. 2016*

## Professional Experience

---

**UCLA-Speech Processing and Auditory Perception Lab**

*Graduate Student Researcher - Automatic speech recognition for children*

**Los Angeles, U.S.A.**

*Sept. 2019 - Present*

- Research on children's ASR as a low-resource task.
- Using data augmentation and model pre-training methods to improve children's ASR
- Advisor: Prof. Abeer Alwan

**PAII Inc. - US Research Lab**

*Research Scientist Intern - Non-autoregressive transformer*

**Palo Alto, USA**

*June 2020 - Sept. 2020*

- Research on non-autoregressive transformer 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 the performance.
- Mentor: Dr. Wei Chu and Dr. Peng Chang

**Sogou Inc. - Voice Interaction Technology Center**

*Research Intern - End-to-end Speech Recognition*

**Beijing, China**

*Apr. 2018 - Aug. 2019*

- Improved training of attention-based encoder decoder (AED) for speech recognition.
- Distributed training of the AED with 50k hours speech data using BMUF.
- Proposed an online AED with 3.5% relative WER degradation compared to an offline AED.
- Improved training of speech-transformer using parallel schedule sampling and relative positional encoding.
- Reduced the latency of the speech-transformer.
- Mentor: Dr. Pan Zhou

**BUPT - Pattern Recognition and Intelligent System Lab**

*Research Assistant - Speech Recognition*

**Beijing, China**

*Sept. 2016 - Jan. 2019*

- Connection temporal classification (CTC) for phoneme recognition [[URL](#)].
- Frequency warping (normalization) for speaker adaptation.
- Advisor: Prof. Gang Liu

## Conference Publications

---

- [6] **Ruchao Fan**, Wei Chu, Peng Chang, and Jing Xiao, "CASS-NAT: CTC Alignment-based Single Step Non-autoregressive Transformer for Speech Recognition," Accepted to ICASSP2021.
- [5] **Ruchao Fan**, Amber Afshan and Abeer Alwan, "BI-APC: Bidirectional Autoregressive Predictive Coding for Unsupervised Pre-training and its Application to Children's ASR," Accepted to ICASSP2021.
- [4] Gary Yeung, **Ruchao Fan**, and Abeer Alwan, "Fundamental Frequency Feature Normalization and Data Augmentation for Child Speech Recognition," Accepted to ICASSP2021.
- [3] Vijay Ravi, **Ruchao Fan**, Amber Afshan, Huanhua Lu, and Abeer 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] **Ruchao Fan**, Pan Zhou, Wei Chen, Jia Jia, and Gang Liu, "An Online Attention-Based Model for Speech Recognition," in Proc. Interspeech 2019, 2019, pp. 4390–4394.
- [1] **Ruchao Fan** and Gang Liu, "CNN-based audio front end processing on speech recognition," in 2018 International Conference on Audio, Language and Image Processing (ICALIP). IEEE, 2018, pp. 349–354.

## Arxiv Papers

---

- [1] Pan Zhou, **Ruchao Fan**, Wei Chen, and Jia Jia, "Improving generalization of transformer for speech recognition with parallel schedule sampling and relative positional embedding," arXiv preprint, arXiv:1911.00203.

## Skills and Coursework

---

### Computer Languages and Open-source Framework

- Python, C/C++, Shell and Matlab
- PyTorch, Kaldi and Tensorflow

### Coursework

- ECE205A Matrix Analysis
- ECE214A Digital Speech Processing: Short-utterance text-independent ASV with style mismatch
- ECE214B Advanced Topics in Speech Processing: Noise Robust ASR
- ECE236A Linear Programming: Handwritten digit recognition with SVM
- CS269 Natural Language Generation: Sentence Controllable DA-based E2E Generation in Restaurant Domain

## Projects

---

### Sogou Micro Smart Recorder C1

*Jan. 2019 - Apr. 2019*

- Improved the training of bidirectional speech-transformer as an offline model for Recorder C1.
- Developed a test pipeline for the model evaluation.

## **A pytorch-book in Chinese [URL]**

*Nov. 2017 - Aug. 2018*

- A book of pytorch tutorials with fun projects including computer vision, natural language processing and speech recognition. I took charge of speech recognition and part of NLP.

## **Honors & Awards**

---

### **Academic Honors**

- A first-year graduate fellowship for ECE Ph.D. program at UCLA(2019)
- Huawei Enterprise Scholarship(2017)
- Graduate Academic Scholarship(2016,2017,2018)
- Undergraduate National Inspirational Scholarship(2013,2014,2015)

### **Competition Awards**

- 3rd Prize in Beijing Electronic Design Contest(Fall 2015)
- Honorable Mention in Mathematical Contest in Modeling(Spring 2015)
- 1st Prize in National Electronic Design Competition TI Invitational Tournament(Fall 2014)
- 1st Prize in Beijing Mathematics Competition(Fall 2013)