Yen-Ju Lu

Emails: yenjulu@cmu.edu, neilyenjulu@gmail.com

RESEARCH INTERESTS

automatic speech recognition, speech processing, generative model, machine learning

EDUCATION

National Taiwan University (NTU), Taiwan — Sept. 2014 – Apr. 2017
M.S. Electrical Engineering and Computer Science

National Taiwan University (NTU), Taiwan — Sept. 2010 – June 2014
B.S. Electrical Engineering

WORK EXPERIENCE

Language Technology Institute, Carnegie Mellon University — May. 2021 – Present
Visiting Researcher

Biomedical Acoustic Signal Processing Lab, Academia Sinica — Mar. 2020 – Present
Research Assistant

MediaTek Inc., Taiwan — Feb.2018 – Feb.2020
Artificial Intelligence Software Engineer

NTU Garage, NTU Entrepreneurship Center, NTU — Mar. 2014 – Aug.2014
Group Leader of Peto

PUBLICATIONS

1. **Yen-Ju Lu**, Zhong-Qiu Wang, Alexander Richard, Yu Tsao, Shinji Watanabe "*Conditional Diffusion Probabilistic Model for Speech Enhancement*," **Submitted to ICASSP 2022**
2. Jing Shi, Xuankai Chang, Tomoki Hayashi, **Yen-Ju Lu**, Shinji Watanabe, Bo Xu "*Discretization and Re-synthesis: An Alternative Method to Solve the Cocktail Party Problem,*" **Submitted to ICASSP 2022**
3. **Yen-Ju Lu**, Chia-Yu Chang, Cheng Yu, Jeih-Weih Hung, Shinji Watanabe, Yu Tsao "*Speech Enhancement Guided by Contextual Articulatory Information,*" **Submitted to TASLP**
4. **Yen-Ju Lu**, Yu Tsao, Shinji Watanabe "*A Study on Speech Enhancement Based on Diffusion Probabilistic Model,*" **Accepted by APSIPA 2021**
5. Xuankai Chang, Takashi Maekaku, Pengcheng Guo, Jing Shi, **Yen-Ju Lu**, Aswin Shanmugam Subramanian, Tianzi Wang, Shu-wen Yang, Yu Tsao, Hung-yi Lee, Shinji Watanabe "*An Exploration of Self-Supervised Pretrained Representations for End-To-End Speech Recognition,*" **Accepted by ASRU 2021**
6. Gang-Xuan Lin, Shih-Wei Hu, **Yen-Ju Lu**, Yu Tsao, Chun-Shien Lu "*QISTA-Net-Audio: Audio Super-resolution via Non-Convex l_Q-Norm Minimization,*" **in Proc. InterSpeech 2021**
7. **Yen-Ju Lu**, Chien-Feng Liao, Xugang Lu, Jeih-Weih Hung, Yu Tsao *"Incorporating Broad Phonetic Information for Speech Enhancement,"* **in Proc. InterSpeech 2020**
8. Szu-Wei Fu, Chien-Feng Liao, Tsun-An Hsieh, Kuo-Hsuan Hung, Syu-Siang Wang, Cheng Yu, HengCheng Kuo, Ryandhimas E. Zezario, You-Jin Li, Shang-Yi Chuang, **Yen-Ju Lu**, Yu-Chen Lin ,Yu Tsao *"Boosting Objective Scores of a Speech Enhancement Model by MetricGAN Post-processing,"* **in Proc. APSIPA 2020**
9. **Yen-Ju, Lu,** advised by Lin-Shan, Lee, *"Enhancing Speech Recognition by Deep Unsupervised Learning."* Master's Thesis, National Taiwan University

RESEARCH EXPERIENCE

WAVLab, Language Technology Institute, Carnegie Mellon University

Visiting Researcher May. 2021 – Present

Supervisor: Associate Professor Shinji Watanabe

Project: Diffusion Probabilistic Model-based Speech Enhancement

- Proposed the first diffusion probabilistic model-based SE (DiffuSE) with supportive reverse process to incorporate the noisy speech signal during sampling to reduce the distortion of enhanced speech.
- Proposed conditional DiffuSE (CDiffuSE) for speech enhancement, defeating state-of-the-art generative time-domain models, and keep good generalization ability in unseen data when discriminative model collapes.
- Collaborated with research scientist in Facebook Reality Lab, Alexander Richard, deriving the formulation of CDiffuSE from the evidence lower bound (ELBO).

Project: Self-supervised Pretrained Representations for E2E-ASR

- Surveyed general applications of seven pretrained speech representations, on end-to-end automatic speech recognition (E2E-ASR) models.
- Designed the self-supervised pretrained representation experiments for CHiME4 data set, outperforming the filter bank input feature for noisy ASR.

Project: ESPNet: End-to-End Speech Processing Toolkit

- Developed speech enhancement (SE) scripts for ESPNet-SE.
- Built E2E ASR and E2E audio visual ASR (AVSR) baslines for MISP Challenges.
- Built E2E SE-beamformer-ASR model for ICASSP 2022 grand challenge, L3DAS22.

Biomedical Acoustic Signal Processing Lab, Academia Sinica

Research Assistant Mar. 2020 – Present

Supervisor: Research Fellow (Professor) Yu Tsao

Project: Speech Enhancement with Broad Phone Speech Recognition

- Designed broad phone classes (BPC), clustering phonmes into several groups through articulatory-based or data-driven methods.
- Extracted broad phonetic posterior-gram to provide more accurate information for SE.
- Joint-trained the SE model with the E2E BPC-ASR system to generate speech with more naturally articulation transitions through contextual information.

Project: Audio super resolution through Image Reconstruction Algorithm

- Drew analogy between audio and image and proposed QISTA-Net-Audio to predict the high resolution speech by solving the sparse signal reconstruction problem.

Speech Processing Laboratory, National Taiwan University

Graduate Student Sept. 2014 – Apr. 2017

Supervisor: Professor Lin-Shan Lee

Project: Deep Unsupervised Learning for Speech Recognition

- Utilized deep unsupervised learning for extracting bottleneck features and combined with unsupervised multi-task learning to achieve accuracy improvement.
- Designed semi-supervised learning to reduce 30% labeled data and achieved better performance than the reference test.

Project: Interactive Question Answering Project with INTEL

- Conducted research on using seq-to-seq, RNN model to generate robots able to learn from asking.
- Improved QA performance through training Question Generator by reinforcement learning.

INDUSTRY EXPERIENCE

MediaTek Inc., Taiwan
Artificial Intelligence Software Engineer Feb.2018 – Feb.2020
Project: Automatic TF-Lite Model Generator
- Designed and developed TF-Lite model generator.
- Won vAward in 2018 and became 2020 Dept. AOP in CAI2 Technology Group.

Project: AI Digital Signal Processor
- Fabricatied new DSP – implemented and optimized Computer Vision/NN algorithms, and constructed the software architecture.
- Took a business trip to introduce the optimized skills on DSP at Arcsoft, China.

NTU Garage, NTU Entrepreneurship Center, National Taiwan University
Group Leader of Peto Mar. 2014 – Aug.2014
Project: IoT Airbag for Cellphone
- Led a team of material, EE, CS major members, designing a protective device that can inflate a mini-airbag on a smartphone when it dropped.
- Selected by NTU Garage and regularly advised by Prof. Chung-Yang Huang and Vice President of Compal Communications, Inc.

TEACHING EXPERIENCE

Biomedical Acoustic Signal Processing Lab, Academia Sinica, Taiwan
Intern Mentor July 2020 – May 2021
- Mentor of summer intern program for M.S. students.

Machine Learning Tutor Mar. 2020 – June 2020
- Designed curriculum for Machine Learning and Programming for M.S. students.

Chen-Li Educational Group, Taiwan
Lecturer of Physics Mar. 2013 – Feb. 2018
- Designed curriculum and taught more than **1000 physics classes with 150-300 students** in a class.
- Taught an individual student who won the **IPHO Gold Medal** in 2019.

HONORS & AWARDS

InterSpeech 2020 Young Scientist Granted 2020
- InterSpeech 2020 travel granted and 1-year ISCA membership.

MediaTek 2020 Dept. AOP 2019
- Project "TF-Lite model generator" was chosen as **one of the three Dept. AOPs** in 2020.

MediaTek vAward 2018
- Designed and implemented the TF-Lite model generator for verification.
- Proqmoted to other departments and projects, reaching beyond the original scope.

MediaTek vAward 2018
- Analyzed and optimized the operations on the DSP, improving its capability.

SKILLS

Programming: Python, Shell Script, C++, Matlab
Machine Learning: Pytorch, Tensorflow, Keras

LANGUAGES

English (Proficient), Chinese (Native), Japanese (Elementary)