Mu Yang

📞 +1 (213)-245-0584
✉ mu.yang@utdallas.edu
🖥 mu-y.github.io
Google Scholar, *GitHub*

Research Interests

Speech Recognition, Speech Synthesis, Natural/Spoken Language Processing.

Education

08/2021– present
University of Texas at Dallas, Dallas, USA.

○ **Ph.D. in Electrical Engineering**. Supervisor: *Dr. John H. L. Hansen*

05/2019 **University of Southern California**, Los Angeles, USA.
○ **M.Sc. in Electrical Engineering**.

06/2017 **Chongqing University**, Chongqing, China.
○ **B.Eng. in Communication Engineering**.

Publications

○ **Mu Yang**, Naoyuki Kanda, Xiaofei Wang, Junkun Chen, Peidong Wang, Jian Xue, Jinyu Li, Takuya Yoshioka, "DiariST: Streaming Speech Translation with Speaker Diarization", *ICASSP*, 2024

○ **Mu Yang**, Ram C. M. C. Shekar, Okim Kang, John H. L. Hansen, "What Can an Accent Identifier Learn? Probing Phonetic and Prosodic Information in a Wav2vec2-based Accent Identification Model", *Interspeech*, 2023.

○ Ram C. M. C Shekar, **Mu Yang**, Kevin Hirschi, Stephen Looney, Okim Kang, John H. L. Hansen, "Assessment of Non-Native Speech Intelligibility using Wav2vec2-based Mispronunciation Detection and Multi-level Goodness of Pronunciation Transformer", *Interspeech*, 2023.

○ **Mu Yang**, Andros Tjandra, Chunxi Liu, David Zhang, Duc Le, Ozlem Kalinli, "Learning ASR Pathways: A Sparse Multilingual ASR Model", *ICASSP*, 2023.

○ **Mu Yang**, Kevin Hirschi, Stephen D. Looney, Okim Kang, John H. L. Hansen, "Improving Mispronunciation Detection with Wav2vec2-based Momentum Pseudo-Labeling for Accentedness and Intelligibility Assessment", *Interspeech*, 2022.

○ **Mu Yang**, Shaojin Ding, Tianlong Chen, Tong Wang, Zhangyang Wang, "Towards Lifelong Learning of Multilingual Text-To-Speech Synthesis", *ICASSP*, 2022.

○ **Mu Yang**, Darpit Dave, Madhav Erraguntla, Gerard L. Cote, Ricardo Gutierrez-Osuna, "Joint Hypoglycemia Prediction and Glucose Forecasting via Deep Multi-task Learning", *ICASSP*, 2022.

○ Mingyu Derek Ma, Jiao Sun, **Mu Yang**, Kung-Hsiang Huang, Nuan Wen, Shikhar Singh, Rujun Han, Nanyun Peng, "EventPlus: A Temporal Event Understanding Pipeline", *NAACL (Demonstrations)*, 2021.

○ **Mu Yang**, Karolina Nurzynska, Ann E. Walts, Arkadiusz Gertych, "A CNN-based Active Learning Framework to Identify Mycobacteria in Digitized Ziehl-Neelsen Stained Human Tissues", *Computerized Medical Imaging and Graphics*, 2020.

- Kung-Hsiang Huang, **Mu Yang**, Nanyun Peng, "Biomedical Event Extraction with Hierarchical Knowledge Graphs", *EMNLP (Findings)*, 2020.
- Prashanth Shivakumar*, **Mu Yang***, Panayiotis Georgiou, "Spoken Language Intent Detection Using Confusion2Vec", *Interspeech*, 2019.
- Rujun Han, I-Hung Hsu, **Mu Yang**, Aram Galstyan, Ralph Weischedel, Nanyun Peng, "Deep Structured Neural Network for Event Temporal Relation Extraction", *CONLL*, 2019.

Work Experiences

05/2023– **Research Intern**, Microsoft, Mentor: *Naoyuki Kanda, Xiaofei Wang*, Manager: *Takuya*
08/2023 *Yoshioka*, Redmond, USA.

- AI research on speech translation.

05/2022– **Research Intern**, Meta AI, Mentor: *Andros Tjandra, Chunxi Liu, David Zhang*, Manager:
08/2022 *Duc Le, Ozlem Kalinli*, New York City, USA.

- AI research on multilingual speech recognition technologies .

08/2019– **Research Assistant**, USC Information Sciences Institute, *Plus Lab*, Supervisor: *Nanyun*
08/2020 *(Violet) Peng*, Los Angeles, USA.

- NLP projects including Event Extraction and Event Temporal Relation Extraction.

Selected Projects

Summaries, demos and codes are available at: `https://mu-y.github.io/#featured`

Mis-pronunciation Detection on Non-native speech.

- Audio Demo: `https://mu-y.github.io/speech_samples/mpd_l2arctic/l2arctic_chinese.html`
- Implemented a text-dependent Mis-pronunciation Detection (MPD) system in PyTorch.
- Explored pre-trained acoustic representations including Wav2vec, Wav2vec 2.0

WaveNet-based Singing Voice Synthesis.

- Audio Demo & Code: `https://mu-y.github.io/speech_samples/synthsing/`
- Collected isolated vocal tracks and obtained time-aligned phonetic transcripts.
- Trained WaveNet-based Timbre model to predict vocoder features conditioning on singer identity, F0 contour, phoneme identity. Used WORLD vocoder to synthesize audio.

Lyrics Dataset Collection, Cleaning and Genre Classification.

- Summary & Code: `https://mu-y.github.io/publication/lyrics_classification/`
- Web crawled ~14k lyrics for 8 music genres based on the metadata returned by iTunes search API.
- Performed classification using models including Naive Bayes, SVM, Bidirectional LSTM.

Equalizer Design for Loudspeaker-Room Correction.

- Audio Demo & Code: `https://mu-y.github.io/speech_samples/roomIR/`
- Implemented second-order filter based equalizers in Matlab with flexible target frequency responses.
- Applied equalizers on multiple Room IRs and asked 21 people to give preferences on un-equalized and equalized audio.

Teaching Experiences

08/2021–12/2021	**Teaching Assistant**, ENGR 3341 Probability Theory and Statistics, UTD.
08/2018–05/2019	**Grader**, EE 483 Introduction to Digital Signal Processing, USC.

Activities & Awards

- ISCA travel award for Interspeech, 2023
- Ok Kyun Kim and Youngmoo Cho Kim Fellowship, UTD, 2023
- Exchange Student, National Sun Yat-sen University, Taiwan, 02/2016–06/2016
- National Scholarship of China (top 1%), 2015
- Outstanding Undergraduate Student Scholarship, consecutive, 2014–2016

Skills

- **Programming Languages**
  Python, Bash/Shell, Matlab, C/C++, Java.
- **Technical Tools**
  Pytorch, Tensorflow, Keras, Kaldi, Vim, Git, Audacity.