ZHEPEI WANG

Amazon Web Services
2100 University Ave C,
East Palo Alto, CA 94303

Phone: +1 (909) 480-6613 (US)
+86 13662609293 (CN)
Email: zhepeiw03@gmail.com
Homepage: http://zhepeiw.github.io

Scholar: LinkedIn: Twitter: GitHub:

Education

Ph.D., Computer Science August 2018 - December 2023
University of Illinois Urbana-Champaign, Champaign, IL, USA

- Division: Artificial Intelligence
- GPA: 3.97/4.00
- Dissertation: ***Data-Efficient Approaches for Audio Classification and Separation*** (pdf) (slides)
  - Proposed the first supervised and self-supervised continual learning approaches to train a sound classifier that can incrementally recognize new sound classes
  - Introduced network designs and training frameworks that enable adaptive computation and better generalization for sound event detection
  - Integrated semi-supervised learning to improve the performance of separation models for music and speech applications
- Advisor: Prof. Paris Smaragdis
- Relevant courses: *Machine Learning for Signal Processing, Signal and Image Analysis, Natural Language Processing, Optimization in Computer Vision, Adversarial Machine Learning*

B.S., Computer Science August 2014 - May 2018
Harvey Mudd College, Claremont, CA, USA

- GPA: 3.93/4.00
- Graduate with High Distinction and Honor in Computer Science

Interests

speech enhancement and separation, sound detection and classification, machine learning, artificial intelligence, deep learning, continual learning, representation learning, multi-modal learning

Project Experience

Amazon Web Services September 2023 - present

- ***Applied Scientist II*** at the Chime SDK Science team
- Enrollment-free Headset and Headset+ VoiceFocus for Premium Voice
- Real-time single-channel speech enhancement using distance-based features without enrollment utterances
- Supervisors: Dr. Ritwik Giri, Dr. Michael M. Goodwin

Amazon Web Services Summer 2022

- ***Scientist intern*** at the Chime SDK Science team
- Integration of VoiceFocus and Personalized VoiceFocus
- Unified framework for real-time personalized and non-personalized speech enhancement
- Supervisors: Dr. Ritwik Giri, Dr. Michael M. Goodwin

Amazon Web Services Summer 2021

- ***Scientist intern*** at the Chime SDK Science team
- Personalized VoiceFocus Pro for offline personalized speech enhancement
- Semi-supervised target speaker extraction using speaker identity cues
- Supervisor: Dr. Ritwik Giri

Amazon Web Services Summer 2020

- ***Scientist intern*** at the AWS AI Audio Signal Processing team
- Semi-supervised singing voice separation and data augmentation
- Supervisor: Dr. Ritwik Giri

Tencent AI Lab Summer 2019

- ***Research intern*** at Audio Group
- Score-to-sound singing synthesis using a neural vocoder based on Text-to-Speech (TTS) pipelines
- Supervisor: Dr. Shiyin Kang

Tencent AI Platform Department Summer 2018

- ***ML Engineering intern*** at Image and Vision Group
- Implemented object detection and image segmentation networks including SSD and Mask-RCNN using Tensorflow
- Supervisor: Dr. Xiaolong Zhu

Amazon Prime Now - HMC Clinic 2017 - 2018

- ***Project leader*** for a 4-member team of the senior capstone project
- Report: ***Image-Text Classification to Correct the Amazon PrimeNow Search Experience*** (pdf) (poster)
- Designed a system that automatically detects mismatches between product images and text descriptions with deep learning
- Designed and implemented a workflow for training a deep neural network to determine the similarity between a pair of images
- Supervisor: Prof. Yekaterina Kharitonova

Engineering Department, Harvey Mudd College 2017 - 2018

- ***Student researcher*** at the Music Information Retrieval Lab
- Live song identification using supervised deep learning and unsupervised machine learning methods
- Supervisor: Prof. Timothy J. Tsai

Teaching Experience

Teaching Assistant, UIUC

- CS 545 (CS 598PS): Machine Learning for Signal Processing — Fall 2020, 2021
- CS 498PS: Audio Computing Lab — Spring 2020, 2021, 2023

Teaching Assistant, HMC

- CS 181B: Advanced Topics in Algorithms — Fall 2018
- CS 140 (MATH 168): Algorithms — Fall 2017, Spring 2018
- CS 70: Data Structures/Program Development Lab — Spring 2017

Honors and Awards

Outstanding Reviewer Recognition, ICASSP 2023
Awarded to the reviewers with outstanding contributions (220/4445)

Saburo Muroga Endowed Fellowship ($ 6,740) 2018 - 2019
Awarded to outstanding graduate students in computer science

Outstanding Clinic Individual Award 2018
Awarded to top-performing students in the senior capstone project (4/100)

Harvey S. Mudd Merit Award ($10,000 per academic year) 2014 - 2018
Awarded to students with superior academic achievement

Reviewer Experience

International Conference of Acoustics, Speech and Signal Processing (ICASSP) 2021 - 2024

International Conference on Learning Representations (ICLR) 2024

Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA) 2021, 2023

Workshop on Detection and Classification of Acoustic Scenes and Events (DCASE) 2022, 2023

IEEE/ACM Transactions on Audio, Speech, and Language Processing (TASLP) 2023

Publications

Z. Wang, Y. C. Sübakan, K. Subramani, J. Wu, T. Tavares, F. Ayres, and P. Smaragdis, "Unsupervised Improvement of Audio-Text Cross-Modal Representations", *In IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA),* Oct. 2023 (pdf) (code) (poster) (talk)

Z. Wang, R. Giri, D. Shah, J.-M. Valin, M. Goodwin, and P. Smaragdis, "A Framework for Unified Real-time Personalized and Non-Personalized Speech Enhancement", *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)* June. 2023 (pdf) (poster) (talk)

Z. Wang, Y. C. Sübakan, X. Jiang, J. Wu, E. Tzinis, M. Ravanelli, and P. Smaragdis, "Learning Representations for New Sound Classes With Continual Self-Supervised Learning", *In IEEE Signal Processing Letters,* vol. 29, pp. 2607-2611, 2022 (pdf) (code) (poster)

E. Tzinis, **Z. Wang**, X. Jiang, and P. Smaragdis, "Compute and Memory Efficient Universal Sound Source Separation", *In Journal of Signal Processing Systems,* vol. 9, no. 2, pp. 245-259, 2022 (pdf)

S. Yuan, **Z. Wang**, U. Isik, R. Giri, J.-M. Valin, M. Goodwin, and A. Krishnaswamy, "Improved Singing Voice Separation with Chromagram-Based Pitch-Aware Remixing", *In IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP),* May. 2022 (pdf)

Z. Wang, J. Casebeer, A. Clemmitt, E. Tzinis, and P. Smaragdis, "Sound Event Detection with Adaptive Frequency Selection", *In IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA),* Oct. 2021, (pdf) (code) (poster) (talk) *Nomination of Best Paper Award*

E. Tzinis, J. Casebeer, **Z. Wang**, and P. Smaragdis, "Separate But Together: Unsupervised Federated Learning for Speech Enhancement from Non-IID Data", *In IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA),* Oct. 2021 (pdf)

Z. Wang, R. Giri, U. Isik, J.-M. Valin, and A. Krishnaswamy, "Semi-supervised Singing Voice Separation with Noisy Self-training", *In IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP),* June. 2021 (pdf) (poster) (talk)

E. Tzinis, **Z. Wang**, and P. Smaragdis, "Sudo rm -rf: Efficient Networks for Universal Audio Source Separation", *In IEEE International Workshop on Machine Learning for Signal Processing (MLSP),* Sept. 2020 (pdf)

E. Tzinis, S. Venkataramani, **Z. Wang**, Y. C. Sübakan, and P. Smaragdis, "Two-Step Sound Source Separation: Training on Learned Latent Targets", *In IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP),* May. 2020 (pdf)

Z. Wang, Y. C. Sübakan, E. Tzinis, P. Smaragdis, and L. Charlin, "Continual Learning of New Sound Classes using Generative Replay", *In IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA),* Oct. 2019 (pdf)

J. Casebeer[♯], **Z. Wang**[♯], and P. Smaragdis, "Multi-view Networks For Multi-Channel Audio Classification," *In IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP),* May. 2019 (pdf)

[♯] Equal Contribution

Preprints

F. Paissan, **Z. Wang**, M. Ravanelli, P. Smaragdis, and Y. C. Sübakan, "Audio Editing with Non-Rigid Text Prompts", Sept. 2023 (demo)

T. Tavares, F. Ayres, **Z. Wang**, and P. Smaragdis, "Effects of Data Leakage in Zero-Shot Learning with Contrastive Audio-Text Pretraining", Sept. 2023

Z. Wang, R. Giri, S. Venkataramani, U. Isik, J.-M. Valin, P. Smaragdis, M. Goodwin, and A. Krishnaswamy, "Semi-supervised Time Domain Target Speaker Extraction with Attention", *arXiv preprint arXiv:2206.09072*, June. 2022 (pdf) (code)