Siwen (Sivan) Ding

+1-646-683-8105 | sivan.d@nyu.edu | | | | | She/They

Brooklyn, NY - 11238, USA

Research Interests

- Machine Perception and Embodiment: Multi-modal Deep Learning, Spatial Audio, Music Informatics
- Audio Representation Learning: Compositionality, Disentanglement, and Generalizability

Education

- **New York University** *Sep 2023 – Jun 2028 (Expected)*
  2nd year Ph.D. in Computer Science at Music and Audio Research Lab, GPA: 4.0/4.0 New York, NY
  ◦ Computer Vision, 3D Audio, Music Information Retrieval, Digital Signal Processing, Machine Learning
- **Columbia University** *Sep 2021 - Dec 2022*
  Master of Science in Data Science, GPA: 3.4/4.0 New York, NY
  ◦ Algorithm, Statistical Inference, Computer System, Deep Learning, Reinforcement Learning, Music Signal Processing
- **Wuhan University** *Sep 2017 - Jun 2021*
  Bachelor of Engineering in Energy and Power Engineering (Track: Thermal Engineering), GPA: 3.8/4.0 Wuhan, CN
  ◦ Advanced Mathematics, Mechanics, Electronics, Dynamics, Automatic Control, Computational Fluid Dynamics

Experience

- **Dolby Laboratories [] | Video to Spatial Audio Generation** *May 2024 – Aug 2024*
  Research Intern with Mark Thomas and Lie Lu San Francisco, CA
  ◦ Engineered a 360° video to spatial audio generation paradigm using diffusion, spatial audio codecs, and contrastive learning
  ◦ Proposed a novel weighted group residual vector quantization cross-entropy loss with ablation analysis
- **Dolby Laboratories [] | Robust User Localization in Dolby Atmos FlexConnect** *Jan 2023 – May 2023*
  Acoustic Mapping Intern with Avery Bruni and Mark Thomas San Francisco, CA
  ◦ Repurposed speech enhancement model to improve DOA prediction robustness in user localization in reverberant rooms
  ◦ Designed and implemented novel PCM covariance matrix mask, leading to a 62% improvement in localization accuracy

Projects

- **Self-Supervised Multi-View Learning for Disentangled Music Representations (ISMIR LBD 2024)** *Aug 2024– Oct 2024*
  Music and Audio Research Lab, NYU (with Julia Wilkins, Juan Bello, and Magdalena Fuentes) []
  ◦ Architected a self-supervised multi-view learning framework to disentangle audio features, focusing on timbre and frequency
  ◦ Demonstrated improved classification of music attributes and separation of subspaces through controlled experiments
- **Soundscape Simulation, Augmentation and Visualization (ICASSP 2024)** *Jul 2023 – Dec 2023*
  Music and Audio Research Lab, NYU (with Iran Roman, Chris Ick, Brian McFee, and Juan Bello) []
  ◦ Developed a Python library for data simulation, augmentation, spatialization, and visualization of spatial audio
  ◦ Conducted ablation studies with DCASE SELD challenge to manifest 37% improvement of augmentation over baseline
- **Voice Anti-Spoofing and Audio Deepfake Detection (ICASSP 2023)** *May 2022 – Nov 2022*
  Audio Information Research Lab, University of Rochester (with You Zhang and Zhiyao Duan) []
  ◦ Innovated a novel loss function for speaker attractor multi-center one-class supervised learning with 120K voice data
  ◦ Refined generalizability of audio spoofing detection to achieve SOTA EER by 38% relative improvement

Publications

C=Conference, S=In Submission

[C.1] **Ding, Siwen**, You Zhang, and Zhiyao Duan. **SAMO: Speaker Attractor Multi-Center One-Class Learning for Voice Anti-Spoofing**. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1-5. IEEE, 2023.

[C.2] Roman, Iran R., Christopher Ick, **Sivan Ding**, Adrian S. Roman, Brian McFee, and Juan P. Bello. **Spatial scaper: a library to simulate and augment soundscapes for sound event localization and detection in realistic rooms**. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1221-1225. IEEE, 2024.

Skills

- **Programming:** Python, R, SQL, Shell (Linux), Slurm, Docker, C, C++, PySpark, Singularity
- **Python Tools:** PyTorch, NumPy, Pandas, Librosa, Scikit-Learn, Scipy, Plotly, Transformers, Gym, Habitat, TensorFlow
- **Media:** Logic Pro, Final Cut Pro, After Effects, Max/MSP, Processing, Sonic Pi, Ableton Live, PureData
- **Interests:** Electric Guitar, Guzheng, Music Production, Song Writing, Video Production