Yongyi Zang

yongyi.zang@rochester.edu | 585-303-5914 | yongyi.dev | Rochester, NY

CAREER OBJECTIVE

I am a highly motivated and skilled audio algorithm engineer with a strong academic background and diverse industry experience both in technology and music. Proficient in audio signal processing, deep learning, and software development, I aim to create innovative solutions to enhance our perception of and interaction with sound.

EDUCATION

University of Rochester (GPA: 3.8/4.0, Cum Laude, Dean's Lists in all available semesters) *09/2019 - 05/2023*

- B.Sc. in Audio & Music Engineering, minor in Computer Science
- Coursework: Digital Signal Processing / Computer Audition / Audio Software Design / Acoustics / Audio Electronics / Intro to Artificial Intelligence / Data Structures and Algorithms / Computer Organization / Design and Analysis of Efficient Algorithms / Database Systems

EXPERIENCES

Audio Information Research Lab (Advisor: Zhiyao Duan) *07/2021 - Present*

Research Assistant

- **Spoofing-Aware Speaker Verification.** Proposed verification technique using phase artifacts; contributed data to the largest challenge in the field - ASVspoof5, and submitted a paper to INTERSPEECH 2023 (first author).
- **SynthTab: Large Scale Synthetic Guitar Transcription Dataset.** Developed pipeline for automatic tablature rendering with string-accurate concatenative synthesis, significantly improving previous state-of-the-art performance; preparing for submission to ISMIR 2023 (first author).
- **GuitarZero: Zero-Shot Guitar FX Transfer.** Designed deep learning network architecture for guitar effects transfer, requiring only five seconds of audio with effect as reference to replicate effect on other DI signals.
- **Euterpe: A Web Framework for Audio Interactive Application.** Designed a web audio application framework to accelerate deployment of audio generation and music information retrieval algorithms to end users; paper submitted to Journal of the Audio Engineering Society (first author); deployed a Baroque-style human-AI real-time counterpoint improvisation system BachDuet to demonstrate its performance.

Neosensory (Supervisor: David Eagleman, Izzy Kohler, Kaan Donbekci) *01/2023 - 05/2023*

Audio Software Development, Algorithm (Intern)

- **Fractive Phoneme Detection with Real-Time TCN.** Designed an algorithm achieving a 19% smaller model size and ~50% less latency on a Cortex-M4 MCU without reducing performance.
- **Remedy High-Frequency Hearing Loss with Motor Vibration.** Conducted research on high-frequency hearing loss compensation, devising a novel method for mapping high frequency content to vibration motors.

Voice Biometrics Group (Supervisor: Peter Soufleris) *05/2022 - 08/2022*

Audio Software Development, Algorithm (Intern)

- **Optimizing Speaker Verification Network Backbone.** Implemented a more efficient network backbone, with 82.8% smaller model size and ~30% less inference time without impacting performance.
- **Building Java DL4J Training Pipeline.** Using DL4J and ND4J, implemented the entire DL training pipeline.
- **Migrating to TorchServe-based Serving Solution.** Proposed and built TorchServe-based solution, identified efficiency bottleneck and set up CentOS CI/CD testing server for faster research-to-deployment turnaround.

Withlight Studio *09/2019 - 07/2022*

Founder, Music Producer

- **Music Production.** Collaborated with Sony Music Asia and Universal Music Greater China to produce pop songs for the Chinese market with Golden Melody Awards and Academy Awards (Oscar) nominees and winners; composed and produced promotional songs for APEX, Lenovo, Baidu and Dell.

COURSE PROJECTS

Enhanced Real-time Acoustic Echo Cancellation (AEC) for ASR Systems

- Developed an AEC algorithm on a Teensy 4.1 (Cortex-M7) system, specifically targeting automatic speech recognition (ASR) systems, with significant improvement in Word Error Rate (WER) for smart speaker deployments, optimizing speech recognition performance.

Advanced Pitch Detection and HMM-based Smoothing

- Implemented YIN algorithm for monophonic pitch detection, then refined estimation results using the Viterbi algorithm and a transition probability matrix learned from human speech, enhancing overall performance.

Onset-based Beat Tracking with Dynamic Programming

- Developed a spectral-based onset detection algorithm and employed dynamic programming techniques for beat tracking on detected onset strength curves, enabling accurate beat analysis.

Source Separation with Non-negative Matrix Factorization (NMF) and CNN-LSTM Architecture

- Implemented a NMF algorithm for source separation, leveraging K-L divergence; designed a model using CNN-LSTM architecture to calculate spectrogram masks for source separation on the DSD100 dataset.

Faust Simulation of "Belton Brick" Spring Reverb

- Successfully simulated the renowned "Belton Brick" sound using multiple filtered delay lines in Faust IDE.

Embedded Hardware and JUCE Plugin for Envelope Follower-modulated Stereo Tremolo

- Implemented an envelope follower on a SHARC development board, and designed a mapping algorithm to modulate a stereo tremolo effect in real-time; utilized JUCE to develop a VST plugin version of the algorithm for seamless integration with various software DAWs.

Least Significant Bit Audio Steganography Toolbox in C

- Engineered a high-performance Least Significant Bit (LSB) toolbox in C, surpassing Python and MATLAB speeds by an average of 3.3x (encoding) and 1.9x (decoding), enabling efficient audio steganography.

PRESENTATION & AWARDS

Euterpe: A Web Framework for Audio Interactive Application (First Author)

Undergraduate Research Expo 2023 (Oral Presentation, one of three Math/Engineering speakers)
University Technology Showcase 2022 (Poster Presentation)
7th Annual Frameless XR Symposium, Frameless Labs (Demo Presentation)

How Do Anti-Spoofing Models Generalize on Unseen Speakers or Attacks? (First Author)

Speech and Audio in the Northeast (SANE) 2022 (Poster Presentation)

MAMA (Composer, Arranger)

Bilibili Tuyiyue Original Music Awards - First Place
Surpassing second place (a Golden Melody Awards team with ten-fold budget) by 31% in weighted score

There's No Place Like Home (Composer, Arranger, Producer)

2021 GZDOC - "Best Documentary" Nominee

An Ode to the Sky (Composer, Arranger, Producer)

2021 AniWoW! - "Best Animation" Nominee, "Best Market Potential" Winner

SKILLS

Programming Languages:	Python, JavaScript, Java, C (proficient); C++, PHP, HTML/CSS, Shell Script
Software Tools:	PyTorch, TensorFlow, Vue.js, React.js, JUCE, Git, Django, Firebase, MySQL
Audio Production Tools:	Logic Pro X, Pro Tools (proficient); REAPER, Audacity, Ableton Live, Studio One
Instruments:	Piano, Synthesizers (proficient); Guitar, Drum Kits, Flute, Accordion