

# JINGYAN (JOY) XU

(+1) 626-688-7539 ◇ jyxu0830@gmail.com ◇ **Portfolio:** <https://yellowwavesr.github.io>◇

## EDUCATION

---

### Georgia Institute of Technology

Aug. 2022 - May. 2024

Master of Science in Electrical and Computer Engineering  
Master of Science in Music Technology

### Shanghai Jiao Tong University

Sep. 2017 - Jun. 2021

Bachelor of Engineering in Information Engineering

**Research interests:** Machine Learning and Deep Learning, Signal Processing in multimedia, Music Information Retrieval, Multi-model

**Coursework:** Mathematical Modeling, Data Structures and Algorithms, Statistical Machine Learning and Deep Learning

**Rewards:** School of Design Dean Scholarship, Georgia Institute of Technology

**Academic Activities:** Reviewer for ICASSP 2024, 2025

## WORK EXPERIENCE

---

### Deep Learning Intern - ClearMotion, Inc.

May. 2024 - Aug. 2024

*Learning-Based Video and Audio Processing for Motion-Enabled 4-D Movies*

- Implemented deep learning models to automatically convert videos into **4-D movies** using **motion-enabled seats**.
- Analyzed the audio information from videos using Sound Event Detection (SED) model **YAMNET** and segmented the audio using audio captioning model.
- Obtained the detailed beat information from music recordings based on **transformer** and **fine-tuning** with irregular music in various genre. Create the labels for more than **1000 hours** of music data.
- Make captions for critical scenes from videos based on **LLM**. Created the labels for more than 200 hrs of video data.

### Data Scientist (Audio) Intern - Shure Inc.

May. 2023 - Aug. 2023

*Speech enhancement for linear microphone arrays*

- Enhanced the performance of **linear microphone arrays** using **deep transfer learning** techniques.
- Constructed **mathematical models** of realistic microphone systems with **pyroomacoustics** package. **Visualized** the response of the model using **polar pattern charts** to verify the accuracy of the model.
- Simulate the microphone to obtain the **dataset** in a virtual conference room. **Transferred** an existing **LSTM-CNN neural network** into the speech enhancement task. Trained and tested the network using the simulated recordings.
- The proposed solution could reduce the **cost** of a microphone system for conference rooms by around **50%** and achieve the best subjective experience in similar tasks.

### Adjunct Research Assistant - New York University Shanghai

Sep.2021 - May.2022

*Symbolic music beat analysis*

- Developed software tools for music analysis using machine learning techniques.
- **Tokenize the symbolic music** for the **BERT model**. Add beat information together with pitch, duration, and onset. Make the pre-trained model obtain the beat information from symbolic contexts in different types of music.
- **Finetune the pre-training language model** to detect music beats and note-relationships, reaching an accuracy of **87%**, work on irregular rhythms in natural transitions.

## RESEARCH EXPERIENCE

---

**3D-reconstruction for human head in pre-surgery diagnosis - Georgia Institute of Technology** Aug. 2024 - present

- Modify **NeRF**(Neural Radiance Fields) on clinical data to reconstruct 3D mesh objects for human head.
- Simulate the photo shots in **OpenGL** and transform to **camera-to-world** matrix.
- Use **LLM** to encode patient information and combine with 3D information for diagnosis estimation.
- Build up baselines for the same original data in 2-D **Convolutional Neural Network(CNN)** and **3DCNN**.
- Tag the **landmarks** for human faces in 2-D and 3-D.

**Cavity Detection Using Image Classification and Annotation - Georgia Institute of Technology** Feb. 2024 - May. 2024

- Developed deep learning models to process **dental radiographs** and determine the presence of cavities.
- Prepared the images using three different **clinical datasets** in different qualities and from patients at different ages. Deployed and **fine-tuned** the **YOLO-v8** model using processed images to **segment** the images into single teeth.
- Built and trained the **U-Net** model to **classify** the severity of cavities, and to **annotate** the locations of the cavities.
- Integrated two models into a single system and provided **API** to process detection requests. The model achieved an accuracy of **93%** in label-level classification and **86%** in pixel-level annotation.

**Multi-channel speech enhancement for mobile phones - Shanghai Jiao Tong University** Dec.2018 - Oct. 2019

- Designed digital signal processing (**DSP**) algorithms for audio **beamforming** and **noise reducing**.
- Performed STFT on recorded audio signals. Constructed Long-short Term Memory Network(**LSTM**) to estimate noise and speech signal **masks**.
- Performed Generalized Eigenvalue Decomposition (**GEV**) on mask covariance matrices and applied Blind Analytic Normalization (**BAN**) to obtain **filter parameters**. Filtered the frequency domain for **noise deduction**.

## TECHNICAL SKILLS

---

- **Software programming Language:** Experienced with Python (6 years), MATLAB (4 years), Familiar with C, C++ (2 years), Bash (3 years), JavaScript, CSS and HTML (1 year)
- **Hardware development:** Verilog, VHDL, assembly language
- **Deep learning tools:** Torch, TensorFlow, Azure Cloud, Amazon AWS
- **Documentation:** Git,  $\text{\LaTeX}$ , Confluence, BitBucket