Project Name Singing Data Labeling Tool

Team Lead:Avinash PersaudTeam Member(s):Avinash Persaud, Nandith Narayan, Carlos CepedaFaculty Advisor(s):Dr. William Shoaff

Project Summary:

We created a tool to assist the user in labeling singing data for vocal synthesis. Our tool combines the functionality of various other currently available tools to provide the user with an all-in-one tool. We used machine learning models to automate highly repetitive tasks.

Major Challenges:

We had to devise a method to automate the classification of phoneme labels. A data structure to store all the user's project data needed to be made. A fluid and easy to navigate Graphical user Interface had to be created. Computationally intensive calculations needed to be performed.

Solution Methods:

We created a hierarchical classification system using a series of convolutional neural networks which took in the MEL spectrogram of the audio sample as input. By using images instead of raw audio, the data size drastically decreased, hence smaller models could be used. We used the Qt graphics library to design a Graphics User Interface and connected it to our intermediate data structure. We used multi-threading to speed up the computation of computationally intensive calculations such as computing the spectrogram using the fast fourier transform.

