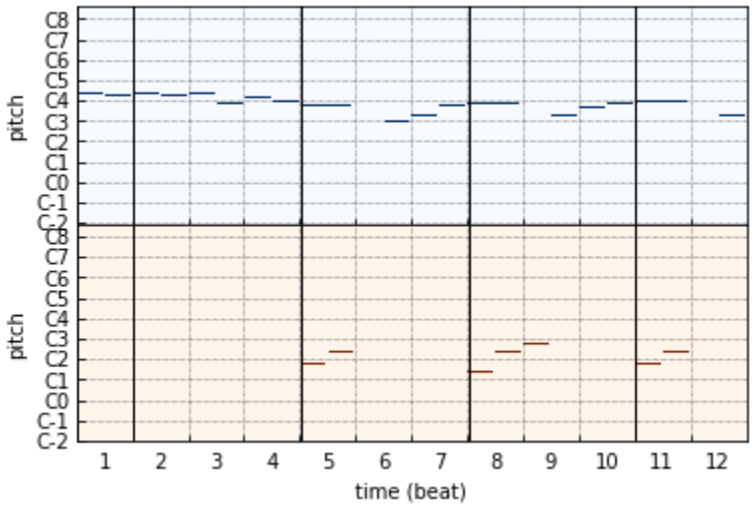
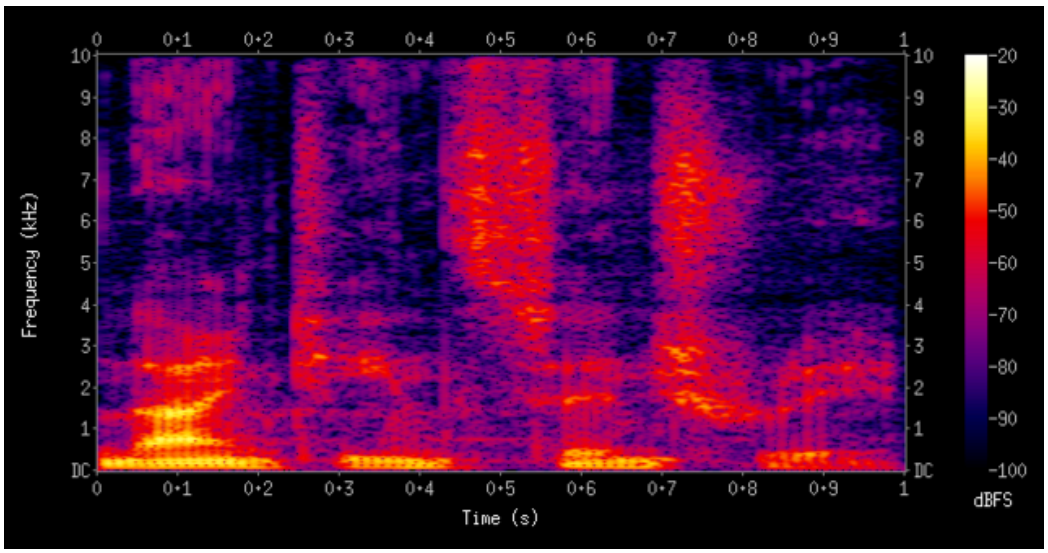
1. Choice of Dataset

https://www.kaggle.com/datasets/jembishop1/classical-music-piano-rolls/code
Our dataset contains 7134 pieces from 40 different composers arranged in piano roll format. Each piece contains rows of 1 by 128 vectors, indicating which notes are playing at a given time, as well as labels for the name of the piece and the composer. The music pieces are converted into piano roll format which displays the pitch of the piece at a given beat, producing a graph like this: (Dong)



However, when doing our research, we noticed that another way to display audio in an image format was using spectrograms, which is used in the Shazam algorithm.



We would need to turn the raw audio file into a waveform and then into the spectrogram. (https://www.audiolabs-erlangen.de/resources/MIR/cross-era)

2. Methodology

a. Data Preprocessing

Since all of the pieces are different lengths, we need to resize our data in order to perform matrix operations. To do this, we could split each piece into smaller segments so that each piece is at least one segment long. For example, a short piece would be processed as one segment, while a longer piece could be split into multiple sub-pieces.

In addition, not all composers are equally represented in our dataset. Some lesser known composers only have a few pieces in the dataset, while other more famous composers have many of their works represented. As a result, our model might be less accurate at predicting these composers, or could skew the results of predicting other composers. If this is the case, we may consider removing a few composers from our dataset.

b. Machine Learning Model

From the given dataset, we want to predict the composer of the piece. Given that we are doing a classification of audio, we propose to use a Convolutional Neural Network (CNN). The music pieces are converted into piano roll format which displays the pitch of the piece at a given beat, giving us a graph to feed the CNN.

We expect that if a certain composer has a unique signature motif, the CNN would be able to detect such a pattern and be able to deduce the composer. The benefit of using CNN is that it would be able to learn such relevant features from the raw input data. We would also have less data preprocessing to do, other than converting the music pieces into piano rolls or spectrograms. The CNN would also be able to classify even if the music file is muddled with background noise. However, the CNN would require a large amount of labeled training to achieve a high accuracy and it can lead to overfitting if the training data is limited. (Tamanna)

c. Evaluation Metric

Since we are trying to classify a piece of music by composer, we could use a confusion matrix to determine how well our model is performing. With this method, we could identify how well our model is performing and which composers it gets wrong. In particular, a confusion matrix would highlight how well our model is predicting smaller composers without much representation in our dataset. Based on our results, we could alter our dataset to either include or omit these composers. With that, we wish to predict the correct composer with at least a 80% accuracy.

3. Application

We can host our model on a simple web application, where the user can input a pre-recorded audio file or record an audio file (like Shazam). The application would turn the raw audio file into a piano roll or spectrogram (based on which format we choose eventually), which is fed into the trained model. The model will produce the composer it thinks created the piece, and we could use the wikipedia API to fetch a synopsis on the composer to present to the user along with the composer name to print out on the website.

Works Cited

Dong, Hao-wen. “Getting Started — Pypianoroll documentation.” *Hao-Wen Dong*, 2018, https://salu133445.github.io/pypianoroll/getting_started.html. Accessed 8 February 2024.

Tamanna. “Exploring Convolutional Neural Networks: Architecture, Steps, Use Cases, and Pros and Cons.” *Medium*, 24 April 2023, https://medium.com/@tam.tamanna18/exploring-convolutional-neural-networks-architecture-steps-use-cases-and-pros-and-cons-b0d3b7d46c71. Accessed 8 February 2024.

Bhandari, Aniruddha. “Understanding & Interpreting Confusion Matrix in Machine Learning (Updated 2024).” Analytics Vidhya, 11 Jan. 2024, www.analyticsvidhya.com/blog/2020/04/confusion-matrix-machine-learning/.