

CSCE 585 Project Proposal - Melomusic

Ronnie Rahenkamp - Loyalumbreon49

Nathaniel Nguyen - NathanielHNguyen

Github Repository

Github Repository: <https://github.com/csce585-mlsystems/Music-Transcription.git>

Introduction

In the music industry, most writers do not directly write sheet music when first composing a song. Most writers start with a beat, a rhythm, or a melody that they start with before expanding that melody into a song. Melomusic aims to aid musicians with creating their sheet music and identifying the placement of repeated melodies throughout the music. We intend to use a CNN based system to take pre-processed audiofiles as spectrograms as a method of obtaining a sense of temporalty throughout the process of transcription.

Problem Statement

Our project, Melomusic, will transcribe sheet music from an audio file and also search for, identify, and highlight each instance of said melody. Taking the audio file, the ML system will be in charge of determining the correct format for sheet music, placing it on the sheet music neatly and accurately, and identifying melodies within said work. We will give the system a dataset of audio files and examples of sheet music of those audio files. From there, we expect to receive precise and accurate sheet music with repeating patterns identified throughout the system.

Technical Approach

To transcribe music into sheet music, we will use Convolutional Neural Networks (CNNs) designed to analyze spectrograms which are a visual format that maps time, frequency, and intensity onto an image-like grid, making it ideal for CNNs to process and interpret audio data effectively. This approach leverages the CNN's ability to identify patterns and structures within the spectrogram, enabling the accurate translation of complex musical compositions into notation. The system will be trained on a diverse dataset of musical pieces to enhance its accuracy and adaptability across different music genres.

Evaluation

We intend to evaluate this model in a quantitative and qualitative manner. Quantitatively we will bar the model on percent accuracy for a given musical output to determine how many notes were accurate to the original audio file. This will be done via handing the music off to a musician to allow them to determine the amount of correct notes transcribed. Furthermore, qualitatively, we will be examining how readable the given sheet music is and how well it follows proper sheet music format.

In terms of our expansion upon what has been done before, our system will also be tracking melody lines and recognizing these patterns throughout the music. This will be measured primarily quantitatively on the percent accuracy of the pattern recognition.

Related Work

Other publishers are working on the same field of work that our research is derived from. We acknowledge that consumer products such as Klangio have taken an approach to them but have not attained the melodic identification system we intend to create with Melomusic.

References

Prosis, Jeff. "Audio Classifications with CNN's". November 8, 2021. Atmosera. Available at: [https://www.atmosera.com/blog/audio-classification-with-cnns/#:~:text=One%20way%20to%20perform%20audio,CNNs\)%20to%20classify%20the%20spectrograms](https://www.atmosera.com/blog/audio-classification-with-cnns/#:~:text=One%20way%20to%20perform%20audio,CNNs)%20to%20classify%20the%20spectrograms).

T. Kwon, D. Jeong, and J. Nam, "Polyphonic piano transcription using autoregressive multi-state note model," in Proceedings of the 19th International Society for Music Information Retrieval Conference, 2018.

Yujia Yan, Zhiyao Duan (2024). "Scoring Time Intervals using Non-Hierarchical Transformer for Automatic Piano Transcription".