Music to Image Generation with Machine Learning Models

Brian Ng, Maximilian Manzhoshov, Tim Nadolsky, and Haichang Li

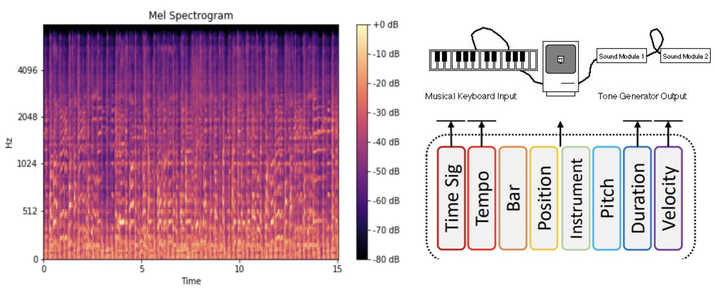
PURDUE UNIVERSITY
College of Engineering

Purpose

- Music-to-Image generation is an unexplored field in deep learning
- With existing music classification models and image generation models, we aim to research a possible synesthetic connection between the two

Research Questions

- How can music be qualitatively characterized into visual descriptions?
- How is music data to be interpreted for characterization and classification?
- What quantifiable descriptions define the synesthesia between art and music?



Acknowledgements

Professors

- Dr. Yung-Hsiang Lu, ECE
- Dr. Kristen Yeon-Ji Yun, Music

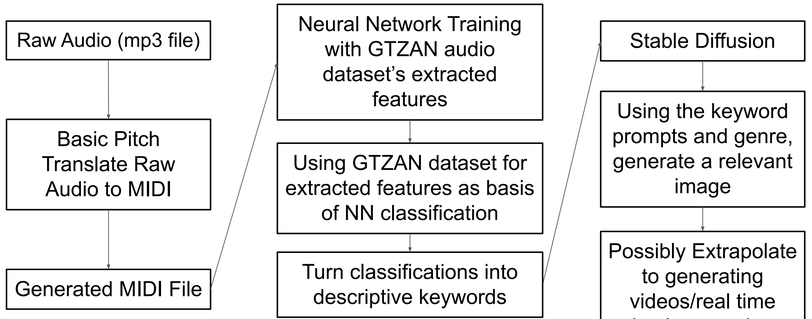
Graduate Student Mentor

- Purvish Jajal

Project Overview

- Through research of existing resources, we have determined the best method to collect music data for processing and classification is MIDI
- Classification is based on a Neural Network training model of the GTZAN audio dataset and its features
- We will base our image generation on the open source Stable Diffusion tool on the basis of keyword prompts

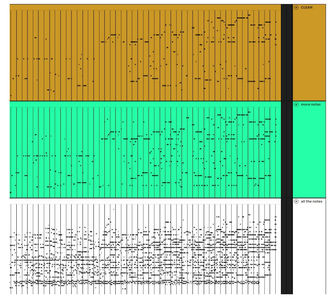
Our Approach

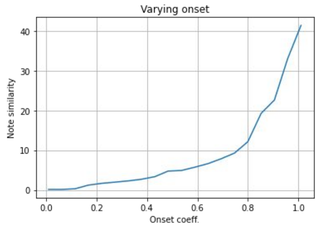


Progress

- The base system structure is implemented for which to map raw audio inputs to relevant images
- More research is to be done with differentiating important keywords and how to extract them
- Applying music theory concepts to analyze chord patterns of MIDI will also assist with sub-genre classification
- Future work will look into emotions as keywords with the VA model, as well as methods of enabling real-time use of this system

MIDI File generation

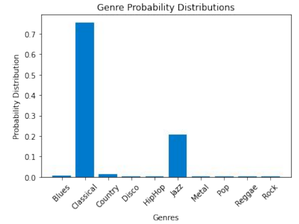




Testing Accuracy on Genre

```
genres = ['Blues', 'Classical', 'Country', 'Disco', 'HipHop',
          'Jazz', 'Metal', 'Pop', 'Reggae', 'Rock']
predictions = model.predict(X_test)

x = np.arange(len(genres))
fig, ax = plt.subplots()
ax.bar(x, predictions[2], color='#007acc')
ax.set_xticks(x)
ax.set_xticklabels(genres, rotation=45)
ax.set_xlabel('Genres')
ax.set_ylabel('Probability Distribution')
ax.set_title('Genre Probability Distributions')
plt.show()
```



Diffusion generated Images

References

1. Stable Diffusion Open AI
2. Classification of Music Genre: A Machine Learning Approach
3. Estimating Resemblance of MIDI Documents
4. OP-Z Music Synthesizer and Stability AI text to art generator
5. Spotify Basic Pitch
6. GTZAN Dataset