Music Genre Classification

Univ.AI

About the project

- Understand techniques used in the audio domain
- Build a machine learning model which classifies music into its respective genres.

Data:

- GTZAN Dataset
- 10 genres
- 100 songs per genre
- 30 seconds per song

Audio Preprocessing:

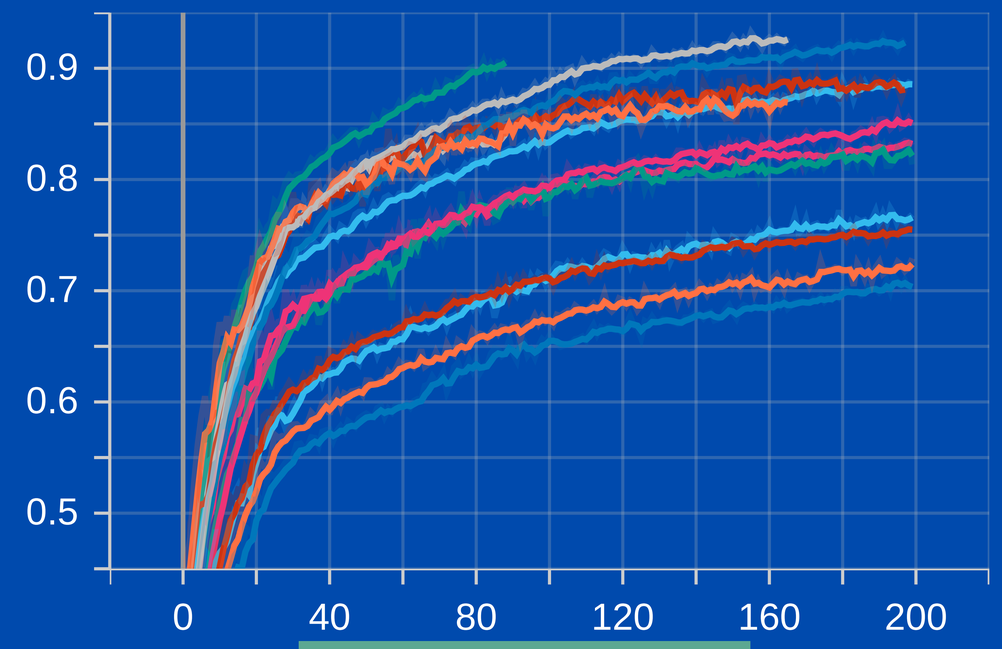
- Analog to Digital
- Digital to FFT
- FFT to STFT
- STFT to Mel Frequency Cepstral Coefficients (MFCC)
- Used Librosa Library

Conclusion

- CNNs performed better than FFNNs
- More filters in the initial layers is useful
- Using STFT for training is memory intensive

Future scope

- Use Recurrent Neural Networks
- Classify using Auto-Encoders
- Use Skip-connections in the architecture
- Transfer Learning
- Use more MFCC coefficients



Epochs vs Accuracy

(Using Tensorboard)

Filters per Conv2D layer	Accuracy	Loss
[128, 64, 64, 32]	82.78	0.516
[256, 128, 64, 64]	85.53	0.518
[256, 128, 64, 32]	86.63	0.477
[32, 64, 128, 256]	79.98	0.640
[64, 128, 256, 256]	85.08	0.574

References:

The dataset: https://www.kaggle.com/datasets/andradaolteanu/gtzan-dataset-music-genre-classification
Librosa - https://librosa.org/doc/latest/index.html

By Flipping -A -Coin:
Shambhavi Aggarwal, Lakshay Chawla, Devansh Shrestha, Rommel Jalasutram, Varnika Vatsyayan