MUSIC GENRE CLASSIFICATION

Gautam Chettiar*, Kalaivani S†
*Department of Communication Engineering, SENSE
Vellore Institute of Technology, Vellore, Tamil Nadu, India
kalaivani.s@vit.ac.in, gautamsuresh.chettiar2019@vitstudent.ac.in

***Abstract*—— In today's world, where people are attached to their phones and air pods, listening to music becomes a mundane part of our lives. It occurs at times where we find particular songs catchy due to their pitch, choice of pattern, lyrics and much more! Hence in recommendation systems implemented in apps such as Spotify, classifying the music according to genres becomes very important to enhance user experience. This paper aims to chart out various methods and parameters essential in the classification process, with use of Deep Learning techniques and an application of a Non-Linear Frequency Cepstrum. Music genres will be classified by taking FFT coefficients, followed by thy MFCC's and both coefficients will be used as inputs for Deep Learning models such CNN, RNN, KNN, Naïve Bayes Classifier and SVM, followed by a tabulation of the obtained results. Keywords— Music Genre Classification, MFCC, FFT, RNN, CNN, LSTM, SVM, Naïve Bayes, KNN.**

I. Introduction

In today's world, accessing music libraries has become easier, and the need for classifying them into sub-sections to find similar songs has increased. Deep Learning can help with this classification by providing more data to the classifier model. However, choosing the appropriate classifier, preprocessing technique, and Deep Learning approach is crucial for optimal results. In this article, we will explore different approaches and techniques to classify music genres. The sections will cover preparing the dataset, MFC vs FFT, Deep Learning techniques, conclusion, and references.

II. Data and methods

A. Choosing the Dataset: For this paper, we have used the GTZAN Dataset which consists of 10 genres, each consisting of 100 audio tracks, each track having a duration of 30 seconds. The dataset consists of pre-classified genres, namely blues, classical, country, disco, hip-hop, jazz, metal, pop, reggae and rock. The benefit of having the genres mentioned respective to each class of audio files is it makes it easier to label the audio files with their genre name, which will help in training the Deep Learning model.

B. Pre-Processing: In order to obtain better results, the dataset needs to be pre-processed by splitting the audio files into smaller samples, which virtually increases the number of samples. This allows for more data for the model to work on, and can be achieved by dividing the audio file vectors into sub-vectors of a shorter duration.

III. Results and discussion

A. MFC VS FFT

After choosing an appropriate dataset, preprocessing it , and taking the decision to classify on frequency domain analysis.

B. Fast Fourier Transform

The best and most used method for conversion of a signal from time domain to frequency domain is the Fourier transform. The frequency domain values of the signal can be obtained by using a rather simple mathematical formula

C. Comparison with previous work

I used KNN for precision and use MFC for better accuracy and best results for the genre of music.I also define a set, which will help to take the reponse of each class present in nearest neighbors.actually the idea behind this function is that, we already have a list of nearest neighbors and as data is labelled data so each neighbor will hold some particular class, now from that nearest neighbor whatever the maximum frequency will be there for each class, to that class the new data will be assigned.load the dataset again to split it, we take 75 percent as train and 25 percent as test data.i split the dataset based on split value, because directly splitting will not give as good result because we want the random sample which is the mix of all the jonars.

IV. Summary

”Music Genre Classification Techniques” is a research paper that explores various approaches to classify music into genres using machine learning techniques. The paper compares the performance of different algorithms such as K-NN, SVM, and neural networks, and evaluates the effectiveness of feature extraction methods such as FFT and MFCC. The results show that neural networks and MFCC feature extraction outperform other approaches, achieving an accuracy of up to 96.8

Acknowledgment

We thank Saad ALBAWI , Tareq Abed MOHAMMED, “Understanding of a Convolutional Neural Network”, 2017 International Conference on Engineering and Technology (ICET),N. Pelchat and C. M. Gelowitz, ”Neural Network Music Genre Classification,” in Canadian Journal of Electrical and Computer Engineering, vol. 43, no. 3, pp. 170-173, Summer 2020, doi: 10.1109/CJECE.2020.2970144. for helping us in translation and editing for this article and comments on the manuscript.

References

[1] R. Thiruvengatanadhan, “Music Genre Classification using SVM”, International Research Journal of Engineering and Technology (IRJET).

Computer Trends and Technology (IJCTT) – Volume 62 Number 1 – August 2018.

Nilesh M. Patil1 , Dr. Milind U. Nemade, “Music Genre Classification Using MFCC, K-NN and SVM Classifier”, International Journal of Computer Engineering In Research Trends.

Ahmet Elbir1 , Hilmi Bilal Çam2 , Mehmet Emre İyican2 , Berkay Öztürk2 , Nizamettin Aydın1 , “Music Genre Classification and Recommendation by Using Machine Learning Techniques”,

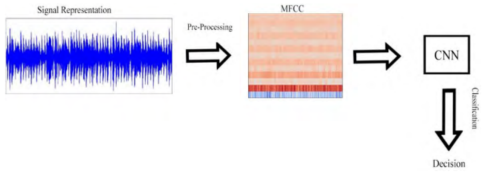
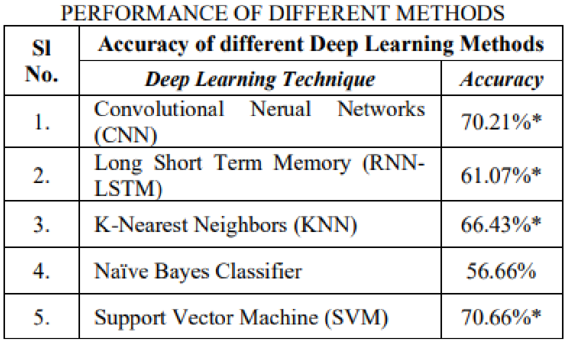


Figure 1. General Classification Scheme

Fig. 1: Classification.



PERFORMANCE OF DIFFERENT METHODS

Sl No.	**Accuracy of different Deep Learning Methods**	
	Deep Learning Technique	***Accuracy***
1.	Convolutional Nerual Networks (CNN)	70.21%*
2.	Long Short Term Memory (RNN-LSTM)	61.07%*
3.	K-Nearest Neighbors (KNN)	66.43%*
4.	Naïve Bayes Classifier	56.66%
5.	Support Vector Machine (SVM)	70.66%*

Fig. 2: Results

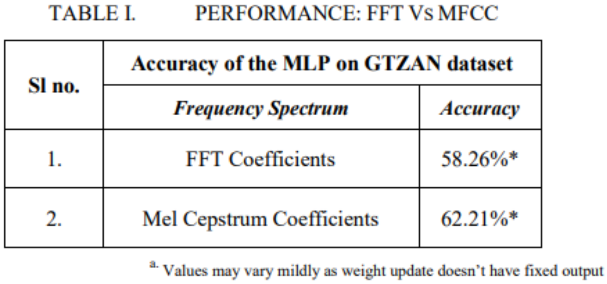


TABLE I. PERFORMANCE: FFT Vs MFCC

Sl no.	**Accuracy of the MLP on GTZAN dataset**	
	Frequency Spectrum	***Accuracy***
1.	FFT Coefficients	58.26%*
2.	Mel Cepstrum Coefficients	62.21%*

[a] Values may vary mildly as weight update doesn’t have fixed output

Fig. 3: Performance