WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Neural Network Music Genre Classification

Bhavana R Maale[1], Samreen Ifath[2]

[1]*Assistant Professor, Dept. Of Computer Science And Engineering, VTU Centre for PG Studies Kalaburagi, India*

[2] *Student, Dept. Of Computer Science And Engineering, VTU Centre for PG Studies Kalaburagi, India*

Abstract

Music is an area that is constantly changing or evolving. Every day, a large number of songs, instrumentals, symphonies, and other types of music are published. This music encompasses a wide range of genres, moods, and themes. To enhance the user experience, music streaming services must identify these songs (UX). A song's various features (such as vocals and instruments) can be used to categorize it into various forms. However, recognizing these characteristics is a difficult job. We use deep learning techniques to define the features of a song that can accurately classify them into their respective genre with minimal human interaction. In recent years, neural network-based music genre classification has had some modest success. Song libraries, machine learning algorithms, input formats, and the types of neural networks used have had varying degrees of performance. In this field, machine learning techniques are used. It also includes some preliminary studies on the classification of musical genres. The study feeds images of spectrograms created from time-slices of songs into a neural network, which then classifies the songs into their respective musical genres.

Keywords: Music, Machine Learning, genre, moods, themes

1.0. INTRODUCTION

One of humanity's most fascinating and vital inventions is music. Music can be characterized as a rhythmic combination of vocals and/or sound created by instruments in a broad sense.Music has been an integral part of people's lives since the beginning. It is a portal into people's hearts. Theoretically, music is a type of data that has developed into a complex pattern. The type of music we listen to is determined by our mood, i.e., the type of music and the listener's mood are inextricably linked. We call it mood for feelings, but we call it genre for music. We have various types of music all over the world with vastly different styles, written with a variety of musical instruments such as percussions and string instruments, as well as special singing patterns. In the world of music, there are approximately 250 distinct genres. Not every genre is diametrically opposed to the others. There are genres that belong to the same family, but they can be distinguished by a number of characteristics. The rock genre, for example, is a wide category that can be further divided into hard rock, progressive rock, soft rock, punk rock, pop rock, and folk rock. Each musical genre has a distinct pattern that can be accurately defined by the model using a mel spectrogram representation. The mel spectrogram representation's pixel intensity level is the deciding factor in determining the input audio file's genre..

2. Related Work

[1] S. Gollapudi, "Practial Machine Learning," Packt Publishing Ltd., 2016. You'll learn about supervised and unsupervised machine learning techniques,

including decision trees, Naive Bayes classifiers, and linear and clustering approaches, as well as tactics for a really advanced approach to statistical data analysis. The book also delves into the latest cutting-edge advances in machine learning, including worked examples and instruction on deep learning and reinforcement learning, as well as practical demonstrations and samples to help you understand even the most complex machine learning approaches. [2] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning.(Report)," Nature, vol. 521, no. 7553, pp. 436, May 2015, 2015. Artificial neural networks (ANNs) are a class of models in cognitive science that are inspired by biological neural networks (e.g., animal central nervous systems) [1]. The method used to develop a model to recognize the semantic libraries of items within small photos is discussed in this study. Two datasets (bird and bottle) were used to train and validate the pictures, as well as to assess the model's accuracy by comparing the output to the testing data. The rest of the paper will go through the methodology in greater depth, as well as the results of the data sets used in contrast to the training datasets. [3] J. Shah, "Neural Networks for the Beginners. Popular types and applications" https://blog.statsbot.co/neural-networks-for-beginnersd99f2235efca, [January 12, 2019, 2017]. In the world of computer science, the terms "neural network" have recently generated a lot of buzz, and it has piqued the interest of many people. But what is all of this about, how do these things function, and are they truly beneficial?Neural networks are made up of layers of processing units called neurons, each with its own set of connections. Data is transformed by these networks until it can be classified as an output. Each neuron multiplies an initial value by a weight, adds the results to additional values flowing into the same neuron, modifies the result by the bias of the neuron, and finally normalizes the output with an activation function.

3. PROPOSED SYSTEM

- Genre Recognition is interesting feature in multimedia applications like music player, streaming applications etc.
- Creating a deep learning project to classify different musical genres from audio data automatically. We'll categorize these audio files based on their low-level frequency and time domain characteristics. And also the machine learning techniques utilized in this area.
- System displays what type of song it is i.e. pop, classical…etc.
- Proposing high accuracy.

4. SYSTEM ARCHITECTURE

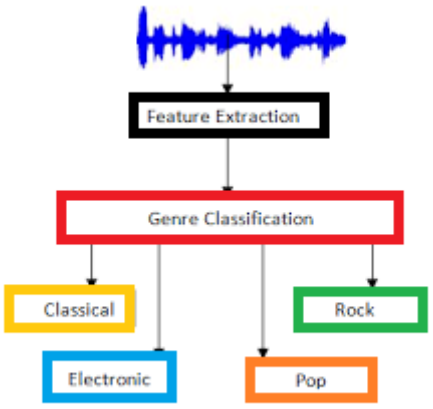


Figure 1: System Architecture

5. Methodology:

Dataset: The GTZAN genre collection dataset contains 1000 audio files, each with a duration of 30 seconds. There are ten classes (ten music genres) with 100 audio tracks each. The tracks are all in.wav format. It includes audio tracks from the ten genres listed below:

- Blues
- Classical

- Country
- Disco
- Hiphop
- Jazz
- Metal
- Pop
- Reggae
- Rock

Feature Extraction: Features and components are extracted from audio recordings. It entails determining the linguistic substance and filtering out noise. Cepstral Coefficients for Mel Frequency: These are cutting-edge features that are employed in autonomous speech and speech recognition research.

These features are created by a series of steps:

- We break the audio signals into smaller frames since they are always changing.
- We then try to detect distinct frequencies present in each frame, which takes about 20-40 ms.
- Separate the language frequencies from the background noise now.
- It then performs a discrete cosine transform (DCT) on these frequencies to remove the noise. We only keep a specific sequence of frequencies with a high likelihood of information using DCT.

Classification: This dataset can be classified using a variety of methods. These are some of the approaches:

Support vector machines with several classes

Clustering K-nearest neighbors with K-means

Convolution neural networks (CNNs) are a type of neural network The K-nearest-neighbors technique was used.

6. Implementation:

Figure 2: Home Screen

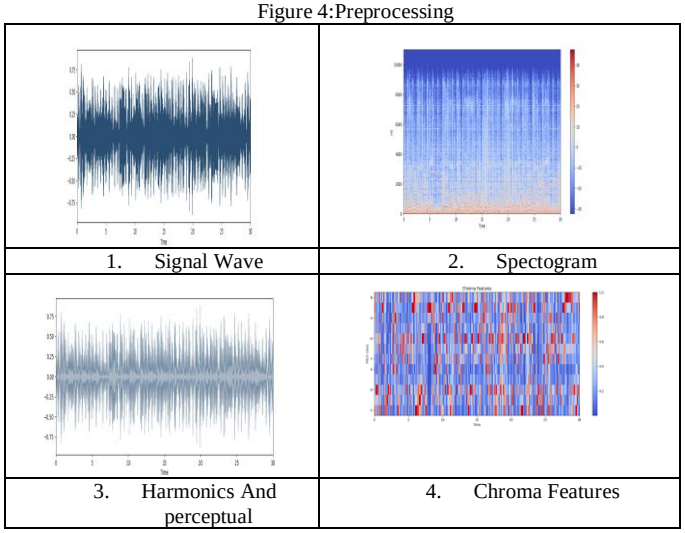
```
           filename  length  chroma_stft_mean  ...  mfcc20_mean  mfcc20_var  label
0    blues.00000.wav  661794          0.350088  ...     1.221291   46.936035  blues
1    blues.00001.wav  661794          0.340914  ...     0.531217   45.786282  blues
2    blues.00002.wav  661794          0.363637  ...    -2.231258   30.573025  blues
3    blues.00003.wav  661794          0.404785  ...    -3.407448   31.949339  blues
4    blues.00004.wav  661794          0.308526  ...   -11.703234   55.195160  blues
..               ...     ...               ...  ...          ...         ...    ...
995   rock.00095.wav  661794          0.352063  ...    -1.193787   49.950665   rock
996   rock.00096.wav  661794          0.398687  ...    -2.795338   31.773624   rock
997   rock.00097.wav  661794          0.432142  ...    -2.106337   29.865515   rock
998   rock.00098.wav  661794          0.362485  ...    -3.590644   41.299088   rock
999   rock.00099.wav  661794          0.358401  ...     1.155239   49.662510   rock

[1000 rows x 60 columns]
(1000, 60)
label                 0
mfcc20_var            0
mfcc5_mean            0
mfcc4_var             0
mfcc4_mean            0
mfcc3_var             0
```

Figure 3: Dataset

```
mfcc9_mean              1000 non-null float64
mfcc9_var               1000 non-null float64
mfcc10_mean             1000 non-null float64
mfcc10_var              1000 non-null float64
mfcc11_mean             1000 non-null float64
mfcc11_var              1000 non-null float64
mfcc12_mean             1000 non-null float64
mfcc12_var              1000 non-null float64
mfcc13_mean             1000 non-null float64
mfcc13_var              1000 non-null float64
mfcc14_mean             1000 non-null float64
mfcc14_var              1000 non-null float64
mfcc15_mean             1000 non-null float64
mfcc15_var              1000 non-null float64
mfcc16_mean             1000 non-null float64
mfcc16_var              1000 non-null float64
mfcc17_mean             1000 non-null float64
mfcc17_var              1000 non-null float64
mfcc18_mean             1000 non-null float64
mfcc18_var              1000 non-null float64
mfcc19_mean             1000 non-null float64
mfcc19_var              1000 non-null float64
mfcc20_mean             1000 non-null float64
mfcc20_var              1000 non-null float64
label                   1000 non-null object
dtypes: float64(57), int64(1), object(2)
memory usage: 468.9+ KB
```

Figure 4:Preprocessing



1. Signal Wave	2. Spectogram
3. Harmonics And perceptual	4. Chroma Features

Figure 5: Feature Extraction

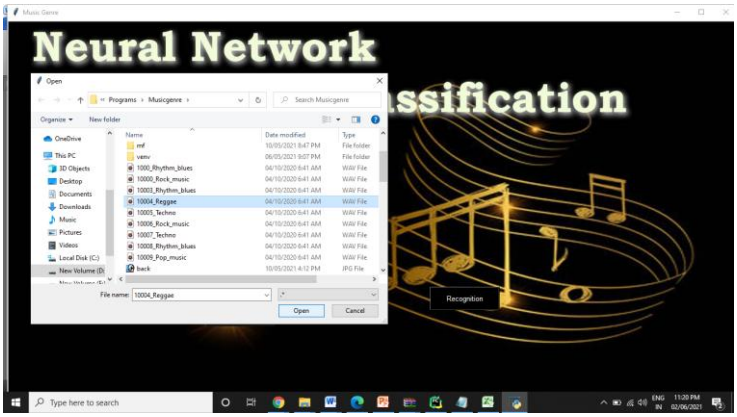


Figure 6: Select wav file for prediction

Figure 7: prediction

7. Conclusion

We propose an effective Music Genre Classification System in this system, which can classify the uploaded audio file into different genres. To train our system, we used the GTZAN dataset, which contains 1000 songs. Librosa, a Python module, is used to extract features, and then spectrograms are created. The CNN Algorithm is being used for classification. The CNN classifier's performance, i.e. categorised songs, is then separated into their own folder. Finally, we can conclude that when we compared CNN to other algorithms such as K Nearest Neighbor, FFNN, and others, we found that CNN was more accurate in classifying the genre. Also, instead of using MFCC (which is less accurate), we use Mel Spectrums to transform audio signals into.png format files, which are then fed into our CNN model.

REFERENCES

S. Gollapudi, "Practial Machine Learning," Packt Publishing Ltd., 2016.

Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning.(Report),” Nature, vol. 521, no. 7553, pp. 436, May 2015, 2015.

J. Shah, "Neural Networks for Beginners: Popular Types and Applications," https://blog.statsbot.co/neural-networks-for-beginnersd99f2235efca, [January 12, 2019, 2017].

O. Mogren, “C-RNN-GAN: Continuous recurrent neural networks with adversarial training,” CoRR, vol. abs/1611.09904, 2016.

Y. Chen, “Towards Explaining Neural Networks,” Faculty of Science, Utrecht University, Utrecht, Netherlands, 2017.

Tomas Mikolov, Martin Karafiat, Lukas Burget et al., “Recurrent neural network based language model,” INTERSPEECH, vol. 2, pp. 4, 2010.

Y. M. G. Costa, L. S. Oliveria, and C. N. S. Jr., "An Evaluation of Convolutional Neural Networks for Music Classification Using Spectrograms," 52, https://www.sciencedirect.com/science/article/pii/S1568494616306421 , [December 16, 2018, 2017].

T. O'Brien, "Learning to understand music from Shazam," https://blog.shazam.com/learning-to-understand-music-from-shazam56a60788b62f, [December 19, 2018, 2017].

Y. M. G. Costa, L. S. Oliveria, A. L. Koerich et al., “Music genre recognition using spectrograms,” 2011 18th International Conference on Systems, Signals and Image Processing, pp. 1-4, 2011, 2011.

M. Lopes, F. Gouyon, A. L. Koerich et al., Selection of Training Instances for Music Genre Classification, p.^pp. 4569-4572, 2010.