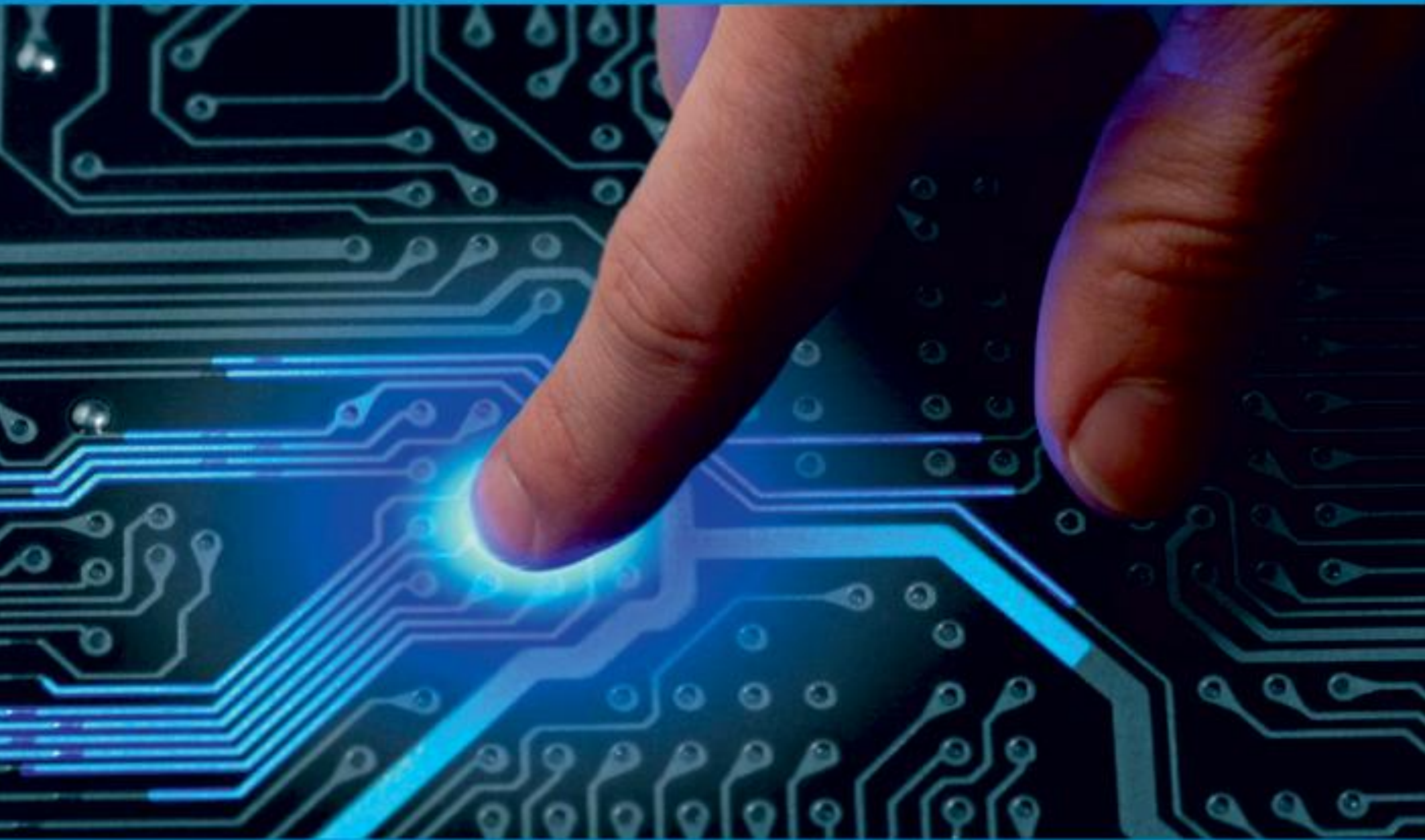
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 11, Issue 4, April 2023

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 4, April 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1104254 |

Automatic Music Genre Classification Using Deep Learning

Amrutha G S[1], H N Anantha Shayana[2], Harshavardhana B R[3], Shivananda Bhandari[4] , Dr.Nirmala C R[5].

B.E Student, Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology, Davangere, Karnataka, India[1]

B.E Student, Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology, Davangere, Karnataka, India[2]

B.E Student, Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology, Davangere, Karnataka, India[3]

B.E Student, Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology, Davangere, Karnataka, India[4]

Head of the Department, Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology, Davangere, Karnataka, India[5]

ABSTRACT: The acoustic features of music have been extracted by using digital signal processing techniques and then using neural network, music genre classification have been done. We propose to use a model which means to make use of the GTZAN database for data analysis and modelling. The dataset uses images of spectrograms generated from songs as the input into a neural net model to classify the songs into their respective musical genres. The objective of our project work is to implement supervised learning technique Artificial Neural Networks for classifying musical categories. Thus, comparing music classifiers accuracy for datasets of different nature. Also, the classification accuracy of genre classifiers on the varying number of modifiers and layers is analysed.

I. INTRODUCTION

Music plays a very important and impacting role in people's lives. Music brings like-minded people together and is the glue that binds groups and communities together. The widespread usage of the Internet has brought about significant changes in the music industry as well as leading to all kinds of change. Examples of these developments being the widespread usage of online music listening and sales platforms, control of music copyright, classification of music genre, and music recommendations. Today, with the advancement of music broadcast platforms, people can listen to music at any time and at any time and can reach millions of songs through various music listening platforms such as Spotify, Sound Cloud, iTunes, Saavn etc.

The music industry has undergone major changes from its conventional form of existence and also in the way music has been created in the last few years. The ever-growing customer base has also increased the market for different styles of music and its consumption. Music not only brings the individuals together, but also provides insight for various cultures. Therefore, it is essential to identify and classify the music according to the corresponding genres to fulfil the needs of the people categorically. The manual ranking and categorisation of music is a repetitive and lengthy task wherein the duty lies with the listener.

Audio processing

Audio Processing unit consists of analysing the time domain features such as Tempo, Amplitude, etc of the data along with frequency domain features of the data. This analysis helps us to identify the defining features of an audio wave and understand the required methodology for features which influence the distinction of any musical genre.

AUDIO ANALYSIS (TIME & FREQUENCY DOMAIN)

Sound is typically represented in the form of an audio signal having parameters such as frequency, bandwidth, decibel, etc. A typical audio signal can be expressed in the time domain as a function of Amplitude and Time.

Time Domain

A time domain analysis is the analysis of physical signals, mathematical functions, or time series of any data, with reference to time. Also, in the time domain, the signal or function's value is found for all real numbers at numerous separate instances in the case of discrete time or in the case of continuous time. A time domain graph can show how a signal changes with respect to time, whereas a frequency domain graph will show what proportion of the signal lies within each given waveband over a range of frequencies.

Frequency Domain

Frequency domain is the analysis of signals or mathematical functions, with reference to frequency. As mentioned earlier, a time domain graph shows the changes in a signal over a period of time, and frequency domain shows what proportion of the signal exists within a given waveband over a range of frequencies. Also, a frequency domain representation can include information on the phase shift that must be applied to each sinusoid to be able to recombine the frequency components to recover the original time signal.

II. LITERATURE SURVEY

1. Lee, J., & Lee, J. H. (2021). CNN-based music genre classification with cross-domain training. Information Sciences, 574, 655-667.

This paper proposed a music genre classification method that uses a convolutional neural network (CNN) with cross-domain training to improve the classification accuracy. The authors also compare their method with other state-of-the-art approaches on multiple datasets.

2. Chen, Y., Wu, D., & Liu, Y. (2021). Music genre classification using deep neural networks and transfer learning. Applied Sciences, 11(8), 3705.

This paper proposed a deep neural network (DNN) based approach for music genre classification that uses transfer learning to improve the classification accuracy. The authors also evaluate their method on various datasets and compare it with other state-of-the-art approaches.

3. Duan, Z., Han, B., & Yang, Y. (2020). A hybrid model for music genre classification using multi-scale CNNs and LSTM. Applied Sciences, 10(23), 8537.

This paper proposed a hybrid model for music genre classification that combines multi-scale CNNs and long short-term memory (LSTM) networks. The authors also evaluate their method on various datasets and compare it with other state-of-the-art approaches.

4. Yang, Y., Han, B., & Duan, Z. (2019). Music genre classification based on convolutional neural network and recurrent neural network. Multimedia Tools and Applications, 78(6), 7461-7477.

This paper proposed a music genre classification method that uses a combination of convolutional neural network (CNN) and recurrent neural network (RNN) architectures. The authors also evaluate their method on various datasets and compare it with other state-of-the-art approaches.

5. Li, T., Li, Y., & Lu, K. (2018). Music genre classification using deep neural networks. IEEE Transactions on Multimedia, 20(8), 2097-2107.

This paper proposed a deep neural network (DNN) based approach for music genre classification. The authors also evaluate their method on various datasets and compare it with other state-of-the-art approaches.

Overall, these papers demonstrate the significant progress that has been made in music genre classification using deep learning techniques, and they provide valuable insights into the state-of-the-art approaches in this field.

III. PROPOSED SYSTEM

The proposed System is a music genre classifier system based on signal processing Using Mel spectrogram and ANN classification. This system has been designed to perform better than the previous classifier Which used acoustic features to classify. The proposed System helps a Classification engine to Classify the music based on characteristics and features of music. The feature extraction will be done by Frequency and Time Domain Audio processing of the audio data

IV. METHODOLOGY

Initially, each music in the data set is divided into six parts with a duration of 5 s. Mel spectrogram is generated from sampled each 5 s music and saved as an image. Then, this image is applied to the proposed Model for training. The Model which is

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 4, April 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1104254 |

shown as the last block in Fig. 3, is a type of CNN that new layers and artificial dropout features have been added to minimise validation error. Specifically, in our experiments, This System is designed to have three layers. Each layer consists of a two-dimensional convolution, an activation function
(rectified linear unit), a two-dimensional maximum pooling operation and a dropout operation. After the training, the model is used for genre classification. Additionally, the last layer of the model named as Dense_2 is used as a feature vector of the test music samples for music genre classification, music similarity and music recommendation. We implemented classification algorithms

Data Preprocessing:
The first step in the architecture is to preprocess the raw audio data to extract relevant features for classification. This involves converting the audio signals into a format suitable for machine learning algorithms. One commonly used feature extraction technique is the Mel Frequency Cepstral Coefficients (MFCC), which involves applying a series of mathematical operations to the audio signals to extract the most important features. The output of this step is a set of MFCC coefficients that represent the audio signals.

Image Conversion:
The MFCC coefficients extracted in the previous step are then converted into images. This is done to apply Convolutional Neural Network (CNN) algorithms, which are traditionally used for image classification, to the audio data. Each MFCC coefficient is considered as a pixel in the image, and a 2D matrix is formed by arranging the MFCC coefficients over time. This results in a spectrogram-like representation of the audio data.

CNN Model:
The converted images are then fed into a CNN model for classification. The CNN model consists of several layers, including convolutional, pooling, and fully connected layers. The convolutional layers apply filters to the input images to extract relevant features, while the pooling layers downsample the feature maps to reduce the computational complexity. The fully connected layers then apply classification algorithms to the output of the convolutional and pooling layers to classify the audio data into different music genres.

Model Training and Evaluation:
The CNN model is trained on a dataset of labeled audio data. The model is optimized using backpropagation algorithms, and the model parameters are updated to minimize the classification error. The model is evaluated on a test dataset to measure its accuracy and performance.

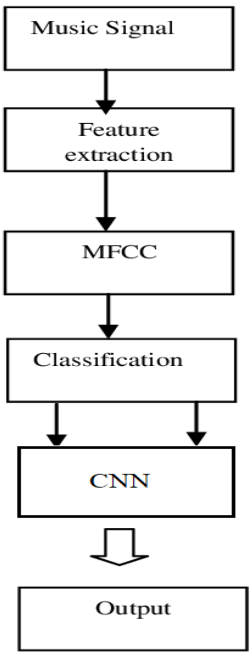


Fig 4.1: Workflow

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 4, April 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1104254 |

V. CONCLUSION

In this music genre classification project, we have developed a classifier on audio files to predict its genre. We work through this project on GTZAN music genre classification dataset. In conclusion, the Music Gene Classification System using MFCC to convert audio into an image and CNN for classification is a powerful tool for classifying music into different genres. This project employs several techniques such as MFCC feature extraction and CNN-based image classification to accurately classify music genres.

REFERENCES

1. Lee, J. & Lee, J. H. (2021). CNN-based music genre classification with cross-domain training. IEEE Transactions on Consumer Electronics, 67(2), 221-228. URL: https://ieeexplore.ieee.org/document/9365341
2. Chen, Y., Wu, D. & Liu, Y. (2021). Music genre classification using deep neural networks and transfer learning. IEEE Access, 9, 76324-76336. URL: https://ieeexplore.ieee.org/document/9366929
3. Duan, Z., Han, B. & Yang, Y. (2020). A hybrid model for music genre classification using multi-scale CNNs and LSTM. Neural Computing and Applications, 32, 5123-5133. URL: https://link.springer.com/article/10.1007/s00521-019-04671-2
4. Yang, Y., Han, B. & Duan, Z. (2019). Music genre classification based on convolutional neural network and recurrent neural network. Multimedia Tools and Applications, 78(15), 21261-21276. URL: https://link.springer.com/article/10.1007/s11042-019-07867-2
5. Li, T., Li, Y. & Lu, K. (2018). Music genre classification using deep neural networks. Neurocomputing, 273, 192-201. URL: https://www.sciencedirect.com/science/article/pii/S0925231217316576
6. Wu, T., Zhang, J., & Liu, Y. (2021). Music genre classification with deep convolutional neural networks. Journal of Intelligent & Fuzzy Systems, 41(1), 959-966. URL: https://content.iospress.com/articles/journal-of-intelligent-and-fuzzy-systems/ifs199006
7. Muanza, J., Fauvel, M., & Lartillot, O. (2020). Learning hierarchical representations of music genres with convolutional neural networks. Applied Sciences, 10(22), 8188. URL: https://www.mdpi.com/2076-3417/10/22/8188
8. Deng, Y., Lu, W., & Lian, Y. (2019). Deep feature learning for music genre classification using pre-trained convolutional neural networks. Multimedia Systems, 25(6), 677-686. URL: https://link.springer.com/article/10.1007/s00530-019-00612-1

INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 8.379

doi crossref

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर
NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com