CONTENT-BASED ANALYSIS OF MUSIC SIGNALS CLASSIFICATION USING DEEP LEARNING

A Project Report submitted in the partial fulfillment of the Requirements for the award of the degree of

BACHELOR OF TECHNOLOGY

IN

COMPUTER SCIENCE ENGINEERING

Submitted by

M. LAKSHMI MANASA	**(Regd.no: 21NE1A05C0)**
K. SNEHA	**(Regd.no: 21NE1A0566)**
K. L. V. SAI SUBHASH	**(Regd.no: 21NE1A0572)**
N. VASU DEVA REDDY	**(Regd.no: 21NE1A05D0)**

Under the Esteemed Guidance of

Dr. R. LALU NAIK

Professor

Department of Computer Science and Engineering

Department of Computer Science and Engineering

TIRUMALA ENGINEERING COLLEGE

(Approved by AICTE & Affiliated to JNTU, KAKINADA, Accredited by NAAC & NBA) Jonnalagadda, Narasaraopet, GUNTUR (Dt), A.P.

2021 – 2025

ABSTRACT

This research paper aims at classification of musical genres which is an example of content-based analysis of musical signals. Here, we propose a new method that combines knowledge from the study of human perception in the classification of musical genres and the neurophysiology of the auditory system. This method works by training a complex neural network to classify a short segment of a music signal. The deep learning method, in which the CNN model is trained end-to-end, to predict the genre label of an audio signal, using only its spectrogram. The approach uses manual features, both in the time domain and in the frequency domain. Tests are conducted on the Audio dataset and we report an accuracy of 76 percent for the classifier. This paper also aims on the characteristics that contribute the most to this classification task are identified. In this paper the acoustic properties of music were extracted using digital signal processing techniques and then using neural networks, music genre classification was performed. We use the GTZAN dataset for data analysis and modeling.

Keywords: CNN, KNN, MFCC, Spectrogram, Mel-Spectrograms

Signature of the Project Guide **Signature of the project Coordinator**

I. INTRODUCTION

With the rapid development of digital technology, the amount of digital music content is increasing dramatically every day. In order to give better music recommendations to users, it is essential to have an algorithm that can automatically characterize music. This process is known as Music Information Retrieval (MIR) [1] and one particular example is the classification of musical genres. However, the classification of musical genres is a very difficult problem because the boundaries between different genres can be blurred. In addition, the amount of labeled data is often much smaller than the size of the data. For example, the GTZAN dataset used in the current work contains only 1,000 audio tracks, but each audio track is 30 seconds long with a sample rate of 22,050 Hz. Traditionally, using MFCC-designed features (Mel frequency factorial), textures, rhythms, etc., achieved 61accuracy in the classification task. Music plays a very important role and affects people's lives. Music binds like-minded people together and is the glue that holds groups and communities together. The widespread use of the Internet has brought about significant changes in the music industry along with all kinds of changes. Examples of these developments are the widespread use of online music listening and selling platforms, music copyright control, music genre classification, and music recommendation. Today, with the great development of music streaming platforms, people can listen to music at any time and can access millions of songs through various music platforms such as Spotify, Sound Cloud, iTunes, Saavan, etc. The music industry has undergone major changes from the conventional form of existence and the way music is made in recent years. The growing customer base has also increased the market for different musical styles and their consumption. This study explores the application of Deep Learning (DL) algorithms for genre identification and classification of certain uses of a complex neural network trained end to- end on the MEL spectrum of audio signals. With the growth of online music databases and easy access to music content, people are finding it increasingly difficult to manage the songs they listen to. One way to categorize and organize songs is based on genre, which is defined by certain musical characteristics such as rhythmic microphone structure, harmonic content, and instrumentation. Being able to automatically categorize and tag music in a user's library, based on genre, would benefit audio streaming services like Spotify and iTunes. In the second part of the study, we extract features both in the time domain and in the frequency domain of the audio signal. Traditionally, these features are then fed into conventional machine learning models, namely logistic regression, random forest, gradient boosting, and support vectors. The machines are trained to classify certain audio files. The model is evaluated on the Audio dataset. We study the relative importance of different characteristics of an audio file.

II. LITERATURE SURVEY

Background :

The first to categorize musical genres as a pattern recognition challenge was Tzanetakis and Cook (2002) [3]. These researchers classified music using a number of musical characteristics, including rhythm, timbre, and pitch. In addition, the authors produced the GTZAN dataset, which has 100 recordings from each of 10 distinct musical genres. One of the most significant databases for categorising musical genres is this one. It is a widely used dataset that is still relevant today. Studies on the categorization of musical genres have benefited from the contributions of GTZAN and the other datasets produced afterwards. Deep learning techniques are now accessible in the music genre categorization process, which was previously carried out using machine learning techniques. This is possible because of the advancement of technology and the usage of graphics processing units (GPUs). CNN, which we implemented in our study and which, based on its outcomes, has a high success rate, was used for this process for the first time in 2018. Humphrey and Bello (2018) claimed that by employing KNN, which they created to address the issues raised in earlier research and provide an alternative strategy, they outperformed the models used to carry out the chord identification procedure. Two distinct deep learning models were created by Zhang et al. (2016) to improve the performance rate of music genre categorization. In contrast to the previous models, the first one included maximum pooling with average pooling. Using shortcut links, they added features from a separate layer to the network to generate the second model in addition to the first (He et al., 2016). Comparing their results to those of earlier research utilising the GTZAN dataset, they were successful. Deep learning, auditory characteristics, and visual features were all investigated by Nanni et al. (2018) for the categorization of musical genres. The tests led them to the conclusion that CNN's features and those obtained using traditional techniques contain information that is distinct. They produced better outcomes than earlier research by combining the qualities learned from these two scenarios. Supervised machine learning techniques, including the Gaussian Mixture model and k-nearest neighbour classifiers, can solve this issue. For this objective, they proposed three sets of characteristics classed as pitch content, rhythmic content, and timbral structure. Music genre categorization has been investigated using Hidden Markov Models (HMMs), which have been widely employed for speech recognition applications.

Literature Survey :

Labels for musical genres are helpful for classifying songs, albums, and performers into larger groupings that have common musical traits. People are finding it harder and harder to keep track of the songs they listen to with the expansion of online music databases and the ease of access to music information. The genre, which is distinguished by some distinctive elements of the music, is one method of classifying and organising songs. Since the early days of the Internet, categorising music by genre has been a subject of intensive investigation. Since they result from a complex interplay between the general audience, marketing, historical, and cultural variables, musical genres lack specific definitions and bounds. An alternative way to focus musical examples from sound music using convolutional neural framework methods is suggested by Li, Chan, and Chun. Their experiments showed that convolution neural networks (CNN) had a strong capacity to detect supporting elements from musical samples that deviate from the norm when conveying irrelevant earlier information. They developed a technique to effectively separate musical examples from background music. Their element extractors utilise the CNN migrated from the image data retrieval field, requiring less prior learning to construct. They showed through their assessments that CNN is a workable choice for programmed highlight mining. This information confirmed their claim that the properties of the various melodic data are similar to those of visual data. Their CNN approach is remarkably adaptable. Numerous research uses deep neural network approaches to analyse speech and other types of audio data in light of the recent success of these systems (AbdelHamid et al., 2014; Gemmeke et al., 2017). Due to the high sampling rate of audio signals, it is not easy to represent audio in the time domain for neural network input. For audio generating jobs, it has been handled by Van Den Oord et al. (2016). The spectrogram of a signal, which records both time and frequency information, is a typical alternative representation. Spectrograms may be used to train convolutional neural networks (CNNs) and are comparable to pictures (Wyse, 2017). The raw MFCC matrix was used as the input for a CNN that was created to predict the musical genre.

III. CONCLUSION

This study compared deep learning techniques in their suitedness to the task of music genre classification. In this project, in this music genre classification project, we have developed a classifier on audio files to predict its genre. After the classification task, implement a recommendation system using cosine similarity based on features extracted from the song. The experimental result shows that the proposed system is better performance on the GTZAN dataset than other methods. The future work is to improve the performance of the system and also uses Transfer learning as a model.