ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24APR2162

Music Genre Classification

Y. Swathi[1]
Assistant Professor
Department of CSE
Tirumala Engineering College
Narasaraopet Andhra Pradesh

N. Snigdha[2]
Student
Department of CSE
Tirumala Engineering College
Narasaraopet Andhra Pradesh

I.Akhila[3]
Student
Department of CSE
Tirumala Engineering College
Narasaraopet Andhra Pradesh

M. Sowmya[4]
Student
Department of CSE
Tirumala Engineering College
Narasaraopet Andhra Pradesh

M. Balaji[5]
Student
Department of CSE
Tirumala Engineering College
Narasaraopet Andhra Pradesh

Abstract:- The Music Genre Classification model automatically divides music into different genres using a small number of audio files and a range of musical attributes. This topic is highly relevant to the field of music information retrieval since it provides a way to organize and analyze large amounts of music files. For MGC, standard machine learning techniques such as SVM, KNN, Decision trees, and neural networks can be applied. These algorithms are trained to recognize different musical qualities and traits, which allows them to categorize the audio files into different genres. Numerous applications show that deep learning algorithms—such as CNN, ANN, and others—perform better than conventional machine learning algorithms. Consequently, the CNN method is adjusted to perform the categorization of music files. This classifies musical genres using deep learning methods from CNN. To evaluate the effectiveness of the MGC algorithms, accuracy is used. Moreover, the impact of different algorithms on MGC performance can be compared and studied. It can be applied to automated music recommendation systems, music production, and music education.

Keywords:- *Music Genre Classification, Deep Learning, Convolution Neural Network, Transfer Learning, Artificial Neural Network.*

I. INTRODUCTION

"Music genre categorization," a flexible technique, is the division of musical content into various groups according to musical characteristics. This is the primary objective of music information retrieval, along with numerous additional uses such as playlist creation, audio file analysis, and music recommendation (MIR).This work explains the many methodologies and procedures used in the classification of musical genres. Under consideration are a number of musical characteristics that are commonly employed in genre classification, including melody, rhythm, harmony, pace, and timbre. The difficulties and disadvantages of classifying musical genres are also covered, including the need to create complex algorithms that can accurately classify a variety of musical genres even in the presence of noise, distortion, and other signal-degrading factors. You'll have an excellent understanding of music genre classification by the end, and you'll be able to apply it to your own music analysis and classification. It is a kind of self-expression that has always been a part of all societies and organizations. A range of techniques and equipment, including as wind-based, computerized, and stringed instruments, can be used to create it. Music is classified into genres based on characteristics including melody, timbre, speed, rhythm, and harmony. Pop, hip-hop, jazz, electronic, country, blues, pop, rock, and hip-hop are some popular musical genres. Every genre has distinct characteristics that are typically linked to particular historical, social, and cultural contexts. The audio waveform is shown here.

II. PROBLEM STATEMENT AND OBJECTIVE

➢ *Problem Statement*

Music is an integral part of everyone's life. Everybody has various musical preferences. It's really arbitrary and subjective by nature. It is imperative to classify music according to its genre; deep learning algorithms can be used to automatically identify music genres. Identifying distinct genres could be difficult due to potential overlaps. Genre classification by hand is a time-consuming procedure. Therefore, developing a trustworthy and accurate music genre classification model requires in-depth subject expertise. CNN is therefore employed in the automated MGC process.

➢ *Objectives*

- The ability to understand various audio formats. By analyzing the audio files, the Mel-Frequency Cepstral Coefficients can be extracted.
- To effectively classify the music data into multiple genres using CNN.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24APR2162

III. PROPOSED SYSTEM

One of the core tasks in the field of music information retrieval is music genre categorization, which involves teaching deep learning algorithms to predict the genre or kind of a given music track. This field of study, sometimes known as "music genre classification," is always changing as a result of scholars experimenting with different approaches to improve classification precision. Convolutional Neural Networks (CNNs), a sort of deep learning methodology, have been applied in recent studies and have shown promise in meeting the goals of music genre categorization.

CNNs are used to classify music genres by taking use of the network's capacity to automatically extract pertinent elements from unprocessed audio input. CNNs are capable of autonomously extracting hierarchical representations, which allows them to capture subtleties and nuanced patterns contained in music songs, in contrast to older approaches that rely on handmade characteristics. The capacity to handle the intricate and varied nature of musical genres—many of which have minute variances and complex that are difficult to measure through manual feature engineering—is especially beneficial.

With CNNs, researchers can use deep learning to identify complex correlations and patterns in music signals, leading to more precise genre predictions. Because CNN architectures are hierarchical, it is possible to extract both higher-level semantic representations indicative of genre-specific properties and lower-level features like rhythmic patterns and timbral characteristics.

In addition, CNNs can handle huge and different music libraries due to their scalability and versatility. Iteratively adjusting their parameters during training allows CNN models to maximize their performance and versatility across many musical genres. CNN-based music genre classification systems can continuously increase their accuracy and robustness over time because to this iterative learning process.

To sum up, the utilization of CNN signifies a noteworthy progression in the realm of music genre classification. CNN-based methods present a viable way to improve classification accuracy by utilizing the built-in capabilities of deep learning. This will allow for more efficient categorization and retrieval of music according to genre attributes. Further developments in CNN-based techniques are anticipated to aid in the improvement and optimization of music genre classification systems as this field of study develops

- *Modularization*

Analyzing the Situation The suggested method is divided into four components. They are data processing, audio visualisation, categorization, and assessment. This section discusses the procedure that is followed, the section that provides the result, and the processed data that is provided to each section.

- *Visualization*

Audio visualisation is the presentation of audio or its visual representation. Using this method, auditory signals are converted into visual elements.

Converting an audio file into a waveform representation: An audio file's amplitude is displayed as a function of time when it is plotted on a waveform. The waveform is shown in Figure 1.

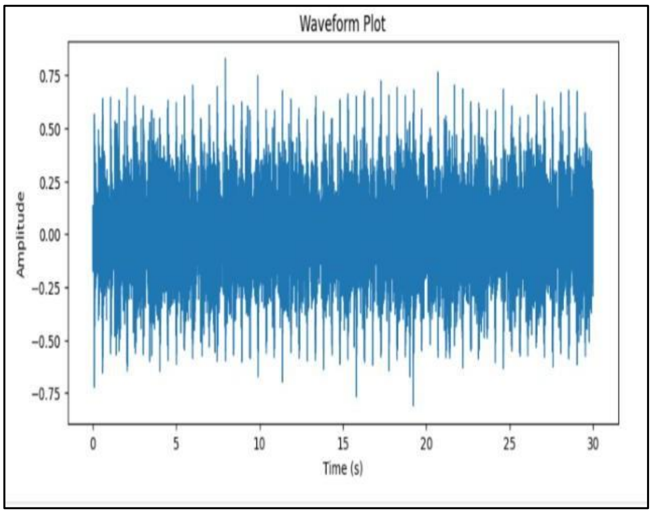

Fig 1 Visualization

Using an audio file to visualise a spectrogram: On the x and y axes of this two dimensional graph, respectively, is the value of time. Figure displays the plot.

Mel-spectrogram visualization of audio files: In conventional spectrograms, the frequency scale is linear. We may view the Mel-spectrogram plot with matplotlib. Figure 2 displays the graph.

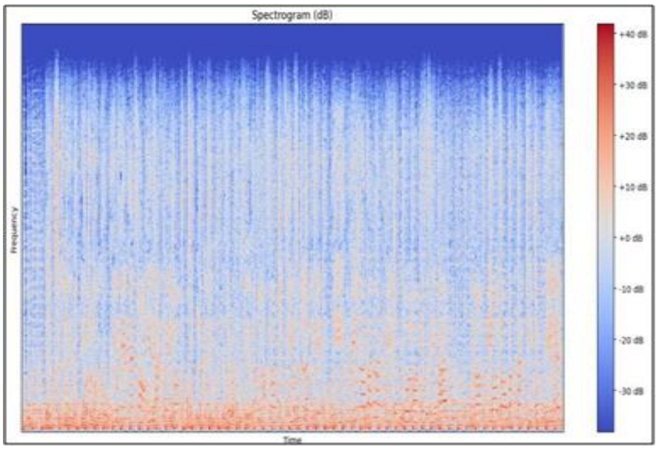

Fig 2 Mel-Spectrogram Visualization

- *Data Preprocessing*

MFCC stands for Mel-Frequency Cepstral Coefficients.The spectral envelope of the signal is described by these coefficients. In Figure 3, the y-axis shows the Mel-Frequency Cepstral Coefficients index, while the x-axis shows the coefficients with time.The color designates the coefficient's magnitude.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24APR2162

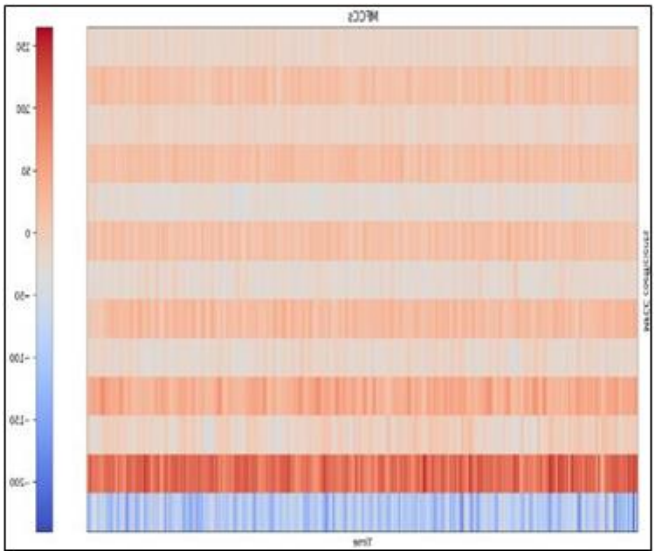

Fig 3 Data Preprocessing

➢ *Classification*

CNNs are utilized in the classification process of musical genres. Convolutional neural networks are a helpful technique for classifying music genres, according to a number of research. Training and test sets of the MFCC from the JSON file are created before the convolutional neural network model is constructed. For testing and training, the split ratios are 30% and 70%, respectively. After this process is completed, the convolutional neural network is defined. Figure 4 shows the CNN model used for genre classification.

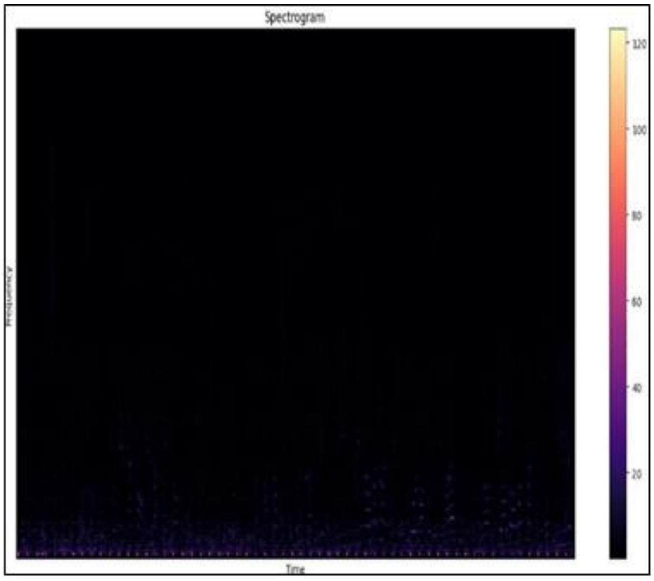

Fig 4 Classification

➢ *Evaluation*

To ascertain the efficiency of the trained model, a thorough examination is necessary. Performance measures are often used to assess the model's ability to forecast different music genres. Among these measurements are recall, F1 score, accuracy, precision, and the use of a confusion matrix.

The assessment procedure includes analyzing the model's performance in respect to different indicators in order to give a comprehensive picture of the model's predictive power. Referred to as sensitivity or recall, recall is the percentage of true positives that the model properly detected. The F1 score offers an objective assessment of the model's performance since it is a harmonic mean of accuracy and recall.

Accuracy, a fundamental metric that is the ratio of successfully predicted occurrences to the total number of anticipated instances, is used to assess how accurate the model's predictions are overall.

The confusion matrix, which provides a comprehensive study of the model's predictions for different classes or genres, is a crucial tool for evaluating performance. Each diagonal member in the confusion matrix represents the number of accurate guesses for a certain class, whereas the off-diagonal elements represent inaccurate predictions. This visual representation allows for a detailed analysis of the model's benefits and drawbacks, allowing for targeted improvements and modifications.

In summary, a thorough analysis of numerous performance indicators, including recall, F1 score, accuracy, precision, and confusion matrix interpretation, is part of the trained model assessment process. By rigorously assessing the model's performance across a variety of parameters, researchers can find areas for future optimization and development as well as helpful insights into its efficacy in music genre categorization.

IV. RESULTS

The GTZAN dataset contains 1000 audio samples (30 seconds each) from publicly available audio snippets. The GTZAN dataset comprises the following ten musical genres: pop, reggae, jazz, blues, pop, country, disco, hip-hop, and reggae. The GTZAN database contains one hundred audio samples for each genre. The GTZAN dataset contains audio clips in the .wav format at a sample rate of 22.05 kHz and a resolution of 16 bits.

The mel-frequency cepstral coefficient (MFCC) is used to train the model and get meaningful information from it. The popular GTZAN dataset is used. MGC works with multi-label categorization. There are several metrics that can be evaluated for these types of positions. The model's accuracy is shown against the quantity of training epochs in a graphic, the accuracy with which the multi-label classification model forecasts the presence or absence of labels.

Confusion matrix is a fundamental tool for evaluating the performance of a classification model. It presents a comprehensive summary of the model's predictions compared to the actual genre labels of the music samples. The matrix consists of four main components: true positives (correctly classified instances), true negatives (correctly rejected instances), false positives (misclassified instances

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24APR2162

as belonging to a genre when they do not), and false negatives (misclassified instances as not belonging to a genre when they do). By analyzing these components, one can derive various performance metrics like accuracy, precision, recall, and F1-score, providing insights into the model's strengths and weaknesses across different music genres. This understanding guides further model refinement and dataset adjustments to enhance classification accuracy and effectiveness.

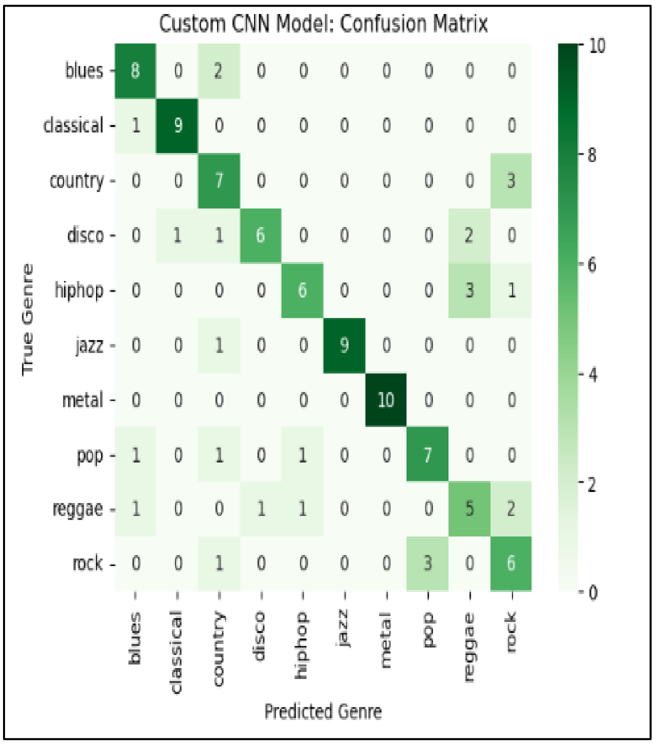

Fig 5 Confusion Matrix

V. CONCLUSION AND FUTURE ENHANCEMENTS

The CNN method is used to accomplish the automated MGC objectives. All of the audio files contain shorter sections. MFCCs are extracted from audio segments and saved as JSON files. Each segment has thirteen MFCCs, which are used to uniquely identify each audio recording. By encoding audio signals in a small feature space, MFCC lowers the dimensionality of the signals. Consequently, Deep learning methods become more straightforward and computationally economical when used to audio processing tasks such as automated music genre classification. This is why it is effective at capturing characteristics unique to a genre that remain consistent across a variety of musical pitches and timbres. Our accuracy rate demonstrates the model's ability to classify genres more precisely. The confusion matrix provides a summary of the forecasts for every genre. If the audio is 30 seconds or less, it is processed now. It takes more research to handle audio files of any length. One could investigate how various audio formats are categorized based on the musical genres in which they are found. The built-in model works wonderfully with WAV files. There are other formats as well, such as FLAC, MP3, and others. Data augmentation can further improve the model's performance. It helps to expand the amount and variety of training data. It is simpler to choose the best combination of hyperparameter values to achieve the best performance with the aid of hyperparameter tuning.

REFERENCES

[1]. Music Genre Classification with Python https://towardsdatascience.com/music-genre-classification-with-python-c714d032f0d8
[2]. T. Shaikh and A. Jadhav, Music Genre Classification Using Neural Network, in Proceedings of the International Conference on Automation, Computing and Communication, ICACC, 05 May 2022, Mumbai, India (2022)
[3]. A. A. Khamees, H. D. Hejazi, M. Alshurideh, S. A. Salloum, Classifying Audio Music Genres Using CNN and RNN, in Proceedings of the Advances in Intelligent Systems and Computing, 5 May 2021, Cairo, Egypt (2021)
[4]. B. Dave, V. Chavan, M. Khan, Dr. V. Shah, J. Emerg. Tech. Innov. Res **8**, 4 (2021).
[5]. N. Ndou, R. Ajoodha and A. Jadhav, Music Genre Classification*: A Review of* Deep-Learning and Traditional Machine-Learning Approaches, in Proceedings of the IEEE Aimtronics, 5 May 202, Toronto, Canada (2021).
[6]. M. Vaibhavi, R. K. Pisipati, J. Inno. Sci. Sust. Tech **1**, 1 (2021).
[7]. Seethal V, Dr. A. Vijayakumar, Intl. J. Tren. Scien. Res. Devp **5**, 4 (2021).
[8]. N. Parab, Shikta Das, G. Goda, A. Naik, Intl. Res. J. Engg. Tech **8**, 10 (2021).
[9]. S. Garg, A. Varshney, Intl. J. Adv. Res. Comp. Com. Engg **11**, 5 (2022).
[10]. Dias, Jessica, Pillai, Vaishak, Deshmukh, Hrutvik, Shah, Ashok, *Music* Genre Classification & Recommendation System using CNN, SSRN, 4111849 (2022).
[11]. J. Mehta, D. Gandhi, G. Thakur and P. Kanani, Music Genre Classification using Transfer Learning on log-based MEL Spectrogram, in Proceedings of the International Conference on Computing Methodologies and Communication, ICCMC, 6 May 2021, Erode, India (2021)
[12]. A. Kamala, H. Hassani, Kurdish, Music Genre Recognition Using a CNN and DNN, in Proceedings of the International Electronic Conference on Applied Sciences, IECAS, 21–25 November 2022, Universidad San Ignacio de Loyola, Peru (2022)
[13]. Zhang, Jingwen, Sci. Prog **21**, 2 (2021)
[14]. D. S. Lau, R. Ajoodha, Music Genre Classification: A Comparative Study Between Deep Learning and Traditional Machine Learning Approaches, in Proceedings of the International Congress on Information and Communication Technology, ICICT, 21-24 February, London (2022)
[15]. R. S. Bhat, B. R. Rohit, K. R. Mamatha, Int. J. Comm. Sci **7**, 1 (2020).
[16]. P. R. S. Sharma, P. Naik, M. Sugna, Intl. Res. J. Moderni. Engg .Tech. Scie **3**, 7 (2021)

[17]. R. P. Ram Kumar, P. Sanjeeva, S. F. Lazarus, D. V. Krishna, Intl. J. Inno. Tech. Explor. Engg **8**, 11S2 (2019)
[18]. M. Thejaswee, V. Srilakshmi, K. Anuradha, G. Karuna, *Performance Analysis of Machine Learning Algorithms for Text Classification*, in Proceedings of the Advanced Informatics for Computing Research (ICAICR 2020), A. K. Luhach, D. S. Jat, K. H. Bin Ghazali, Gao, P. Lingras, (eds), Comm. Comp. Inform. Sci. Springer, Singapore 1393 (2021)
[19]. G. Kalpana, A. Kanaka Durga, G. Karuna, Intl. J. Recent Tech Engg **8**, 2S11 (2019)