Journal of Applied Sciences, Information and Computing
Volume 5, Issue 1, June 2024

ISSN: 1813-3509 https://doi.org/10.59568/JASIC-2024-5-1-04

Application of Deep Learning Model for Music Genre Classification

[1]Olanloye D. Odunayo, [2]Adeniyi A. Emmanuel, [3]Adeyemo A., [4]Emuoyinbofarhe O., [5]Adeleke Isreal, [6]Umezuruike C., [7]Amare Akinfenwa

[1-7]Department of Computer Science, Bowen University, Iwo, Osun State.
Corresponding author email: abidemisteven@gmail.com

Abstract

A musical genre is a term used to classify specific musical styles in accordance with a sense of their significance or a set of guidelines. There are many techniques to classify music into different genres, but attempts to automate this process using machine learning have received much less attention from the media. Blues, Classical, Country, Disco, Hiphop, Jazz, Metal, Pop, Reggae and Rock are the most common musical genres. **Objective**: This study develops and creates a model that can classify music into their respective genre. The dataset, which contains 1000 songs divided into 10 sections with 100 songs each, was obtained from the open-source service "Kaggle" and utilized for the classification. To train and test the model, the dataset was then converted into a JSON format, the new dataset was then splitted into 30% for testing the model and 70% for the training of the model. This model was created using the Convolutional Neural Network (CNN) machine learning technique. The result shows that the model was able to correctly classify 7 audios out of 10 audios and correctly classify them into their original genre. The classifier achieved an accuracy score of about 73%. In conclusion, the study provided a model that will aid in the classification of various types of music into their appropriate genres. This will be highly useful for record companies and other music retailers to group songs for search and retrieval.

Keywords: Deep learning, Artificial intelligent, Music classification, Artist, Prediction.

1. Introduction:

Music has a significant impact on people's lives. Music brings individuals together who share similar interests and is the glue that holds communities together (Clarke, DeNora, and Vuoskoski, 2015). Communities can be identified by the type of music listen they listen to. Various societies and groups listen to various types of music. One important characteristic that distinguishes one style of music from another is its genre.

A variety of applications, including Drinkify (a tool that creates cocktails to fit the music), Pandora, and dynamically producing images that go with the music, make music classification an interesting subject (Nam et al. 2018). However, in the world of music information retrieval, categorizing musical genres has proven to be a difficult issue (MIR). Music genres are highly subjective, making them difficult to categorize and define in a

systematic and consistent manner. Until now, genre classification had been done manually by concatenating it with information from audio files.

Humans are capable of classifying music genres using their hearing, the auditory processing system in their ears, and higher-level cognitive processes in the brain. Humans employ musical genres as a concise description that makes information transfer easier. For example, the phrases "I like jazz more" and "I don't like rock music!" are widely used to communicate opinions and depend on a common understanding of the respective genres and how they relate to culture, history, and musical form. In addition, genre is a category used by record companies and other music retailers to group songs for search and retrieval.

Today, there are over 1300 different forms of music, but this research will focus on just a handful, such as Blues, Classical, Country, Disco, Hip-Hop, Jazz, Metal, Pop, Reggae, and Rock. The CNN model will be created in such a way that it will be able to classify music according to its genre. Many people listen to many types of music without knowing what they are listening to, i.e. they have no idea what genre of music they are listening to, such as Jazz, Hip Hop, Rock, Blues, Reggae, and so on. One of the most prevalent elements used to differentiate musical works is genre (Rustamov, Mamatqulov and Boymatov, 2021). Although there are many broad definitions of genre, human responses to it can be biased. As more music is listened to on digital platforms, it seems obvious that automating the process of music classification would be beneficial to all parties. As a result, this study presents an algorithm that will aid in the classification of diverse musical genres. The novelty of using deep learning model is to predict the genre of the music track by analysing the patterns and features within a given audio signal.

2. Review of Literature:

Kumar, and Alraisi (2022) in their work "Music Genre Classification Using CNN" aimed at constructing a computer model that divides musical samples into many genres. It uses an audio stream as its input and seeks to guess the genre. To produce accurate results, the methodology employed in this study was to pre-process the data once it was collected before separating it into training and testing sets. The model was trained using the CNN algorithm. In order to train the CNN model, Adam Optimizer was employed. After the findings, the result showed that the application is working perfectly fine. Also, the test set's accuracy was 92.93 percent, which is a respectable result. Neural networks are therefore particularly useful in machine learning models. Additionally, this research demonstrates that Convolutional Neural Network (CNN) implementation using Tensor flow is quite beneficial for the classification process.

Hasan et al. (2021) the goal of their project "Music Genre Classification Using Machine Learning" is to develop the best machine learning algorithm for predicting music genres using k-nearest neighbor (k-NN) and Support Vector Machine (SVM). Machine learning methods were constructed using k-nearest neighbor and Support vector machine after data was acquired and pre-processed to classify music into its appropriate genres. The study discovered that the K-NN classifier produced more precise findings than the support vector machine classifier. Despite the fact that k-NN had a 75% success rate, the blues genre was misinterpreted as rock. The k-NN performed poorly when it came to perceiving blues, with a recognition rate of 66%. The SVM misidentified classical music as jazz or hip-hop, whereas rock music was directly associated to a 94% advancement rate. This study demonstrates that SVM can only recognize a limited collection of patterns. The KNN classifier is more effective at categorizing music genres.

Prajwal et al. (2021) in their research work "The goal of "Music Genre Classification Using Deep Learning with KNN" was to create a system for classifying music genres using deep learning and the KNN algorithm that can predict the class and certainty level of Western music from varied genres such as traditional, rock, jazz, rap, and so on. Following the creation of the data set to be used, data preparation was performed, followed by a brief comprehension of how the model works, followed by the creation of a KNN algorithm, and the algorithm was trained using the train dataset and the test dataset to provide the study's output. This study describes an application that conducts music genre classification using deep learning algorithms. To predict its genre, they created a Classifier using audio recordings. According to this research, the program does classification using a K-Nearest Neighbours method. This is accomplished through the use of the libROSA library package for music and audio analysis.

Raval, Dave and Dattani (2021) in their research work "Music genre classification using neural networks" aimed at using CNN's audio benefits and capabilities to construct a music genre categorization model that saves consumers time when looking for different sorts of music. After the datasets were obtained and data pre-processing was completed, the dataset was separated into two parts: the train dataset and the test dataset. In the train dataset, feature extraction was performed, which included pre-processing and feature vectors, followed by the creation of a feature vector database, which was then translated into a deep neural network model, followed by classification and the genre output. Then, for the test data set, feature extraction was done, followed by classification, and finally, the genre output. In this study, they preserved the musical genre allocation consistent in each fold, with 80% of each genre in the train split and 20% of each genre in the verification break. They also reveal that when paired with music transfer learning elements, the multilayer perceptron model performs best. This paper demonstrates a Convolution Neural Network-based automatic music genre classification system. As an outcome, this approach shows

 https://doi.org/10.59568/JASIC-2024-5-1-04

promise for classifying a big collection of music into the correct genre.

Chillara et al. (2019) in their work “The research project "Music Genre classification using Machine Learning Algorithms: A comparison" aimed to create a machine learning model that categorizes music into the correct genre and compare the accuracy of the model. The objective of this study is to obtain a high level of precision, enabling the model to categorize new music correctly according to its genre. Without the use of manually created features, a deep learning algorithm was used to meet the goal of a music genre classification system after pre-processing the data. The CNN network design is used by the deep learning method. CNN's goal in this scenario was to use the Spectrogram, which was used to depict sound waves, to predict the genre designation. This study found that the CNN approach which employs only the Spectrogram as a parameter to forecast a music genre performs best in regards to accuracy, with a test accuracy of 88.5%. With an accuracy of 88.5%, the CNN model was shown to be the best spectrogram-based approach among CNN, CRNN, and CNN-RNN concurrent models in this investigation.

Hareesh Bahuleyan, (2018) in the work “Music Genre Classification using Machine Learning Techniques” compared the effectiveness of two groups of models. The first method uses deep learning to train a CNN model to identify an audio signal's genre only from its spectrogram. In the second strategy, features are extracted from both the time domain and the frequency domain. The adopted methodology was to perform data pre-processing after obtaining the data set. Music genre classification can also be accomplished without the need of hand-crafted features due to the usage of deep learning. A Spectrogram was used to represent the sound wave in this research. CNN's job is to use the Spectrogram to forecast the genre label. An ensemble classifier that incorporates the two suggested approaches achieved an AUC value of 0.894 in the experiments on the Audio set data set. The task of categorizing music genres is investigated in the study using audio set data. The study also demonstrates the value of replicating the CNN and XGBoot models.

3. Research Methodology:

The design of the proposed system has been made to solve key problems that are posed by existing systems. The system is intended to categorize different musical genres into their appropriate categories if they come under this category (Blues, Classical, Country, Disco, Hiphop, Jazz, Metal, Pop, Reggae and Rock). The system receives an audio, processes it once it has passed the algorithm, and then forecasts the genre to which the music or audio belongs. This is a good algorithm that has been put in place to relieve individuals of the worry of categorizing music into different genres. The proposed system will be implemented using python programming language along with Convolutional Neural Network (CNN) machine learning model and libraries such as Numpy, librosa, tensor flow, math, among others. Figure 1 shows the architecture view of the proposed music genre classification system such that a user enters an audio file, the audio file is analysed and visualized, the audio file is converted into a JSON format, the audio file is sent to the model for prediction, the model predicts the output's genre using the trained model, and the user is then shown the output.

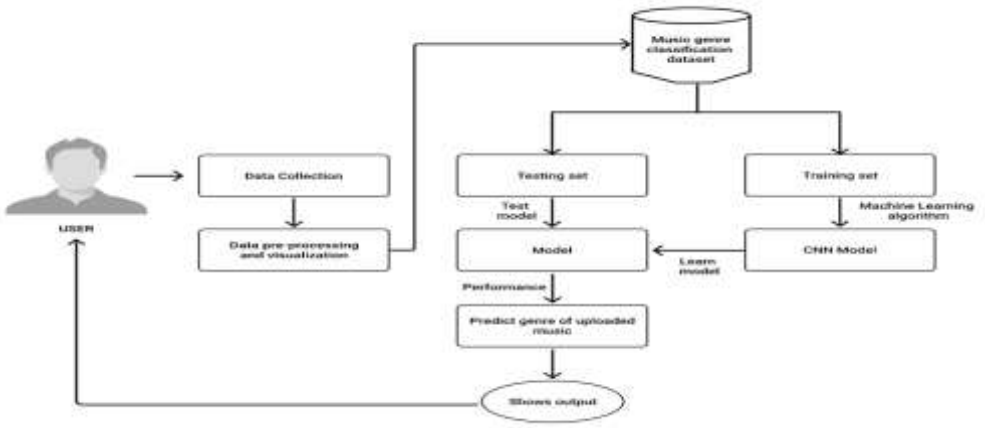

Figure 1. Architectural Design of the Proposed System

 https://doi.org/10.59568/JASIC-2024-5-1-04

3.1 Model Building:

There are several approaches to model building for music genres classification:

Deep Feedforward Neural Networks: Use extracted features as input to a deep feedforward neural network.

Convolutional Neural Networks (CNNs): Treat spectrogram of tracks as images and use CNNs for classification.

Recurrent Neural Networks (RNNs): Given the sequential nature of audio tracks, RNNs or LSTMs can be used.

Hybrid Models: Combine CNN layers (to analyse local patterns) followed by RNN layers (to analyse the sequence of patterns).

This study utilized CNN model for music genre classification due to it novel aspects in the domain of audio signal processing, especially when compared to traditional methods that relied on handcrafted features and simpler machine learning models.

4. Results & Discussion:

4.1 Results:

The algorithm for classifying music into their respective genre was built using a python programming language with a machine learning model which is the Convolutional Neural Network (CNN) algorithm. The dataset used for the classification was sourced and gotten form Kaggle, the dataset (GTZAN) contains 1000 audio files which were classified into 10 genres (Blues, Classical, Country, Disco, Hip-Hop, Jazz, Metal, Pop, Reggae, and Rock).

An audio was taken from the dataset in order to perform analysis and visualization on it, the image as shown in figure 2 shows the waveform plot for the audio that was selected from the dataset. In figure 3 shows a plot for the power Spectrum for the audio that was selected from the dataset. Also, figure 4 and figure 5 shows a plot which shows the Spectrogram and Spectrogram (dB) for the audio which was selected from the dataset. Figure 6 shows a plot for the MFCCs for the audio that was selected from the dataset.

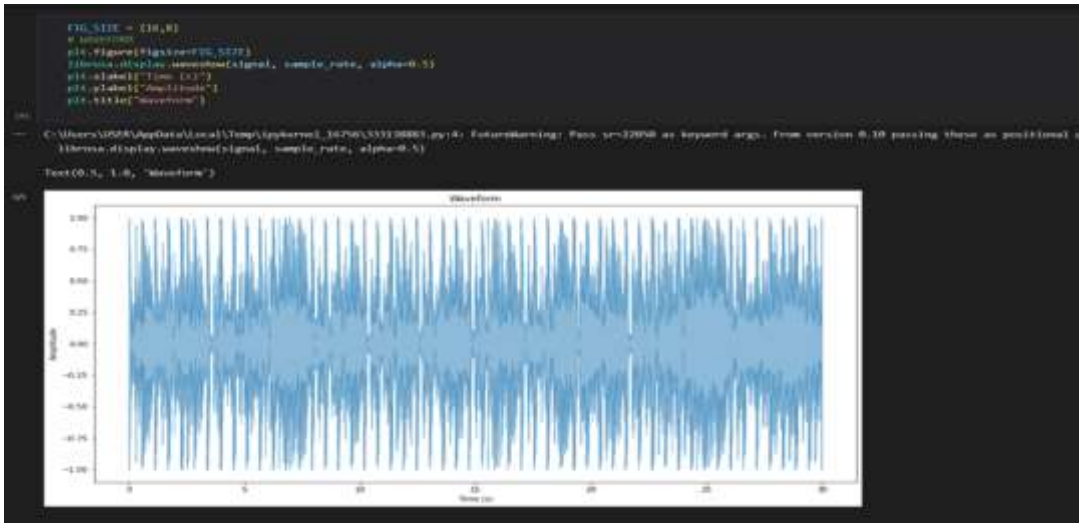

Figure 2 Waveform Plot for the selected audio from the dataset

Figure 3 Power Spectrum Plot for the selected audio from the dataset.

 https://doi.org/10.59568/JASIC-2024-5-1-04

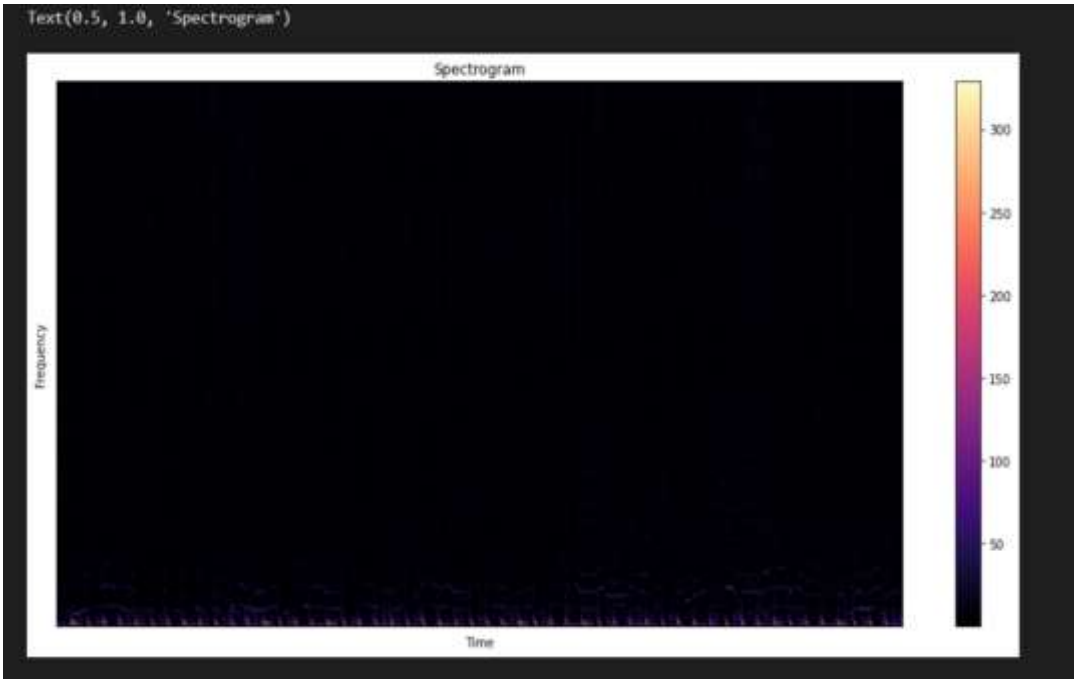

Figure 4 Spectrogram Plot for the selected audio from the dataset

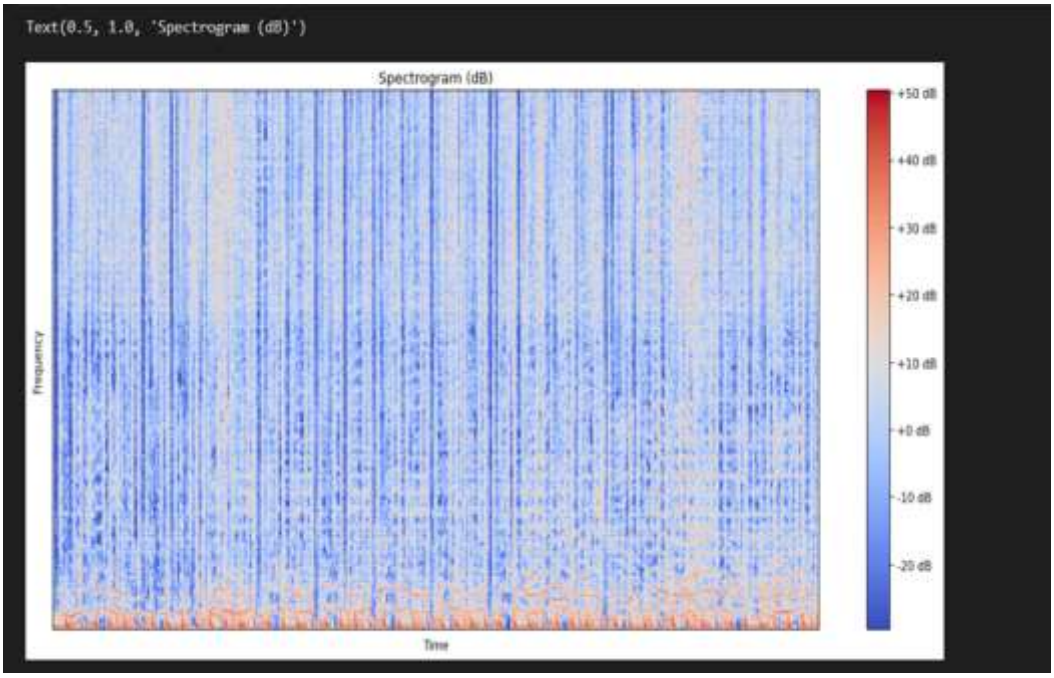

Figure 6 MFCCs Plot for the selected audio from the dataset

The GTZAN dataset was converted to JSON format in order to train the model. JSON (JavaScript Object Notation) is an open standard file format for data sharing that stores and transmits data as human-readable text. The JOSN extension is used to identify JSON files. JSON is an excellent alternative to XML because it requires less formatting. JSON is a language-independent data format that is derived from JavaScript. Many current programming languages allow JSON production and parsing. Figure 7 depicts the original dataset's JSON format.

 https://doi.org/10.59568/JASIC-2024-5-1-04

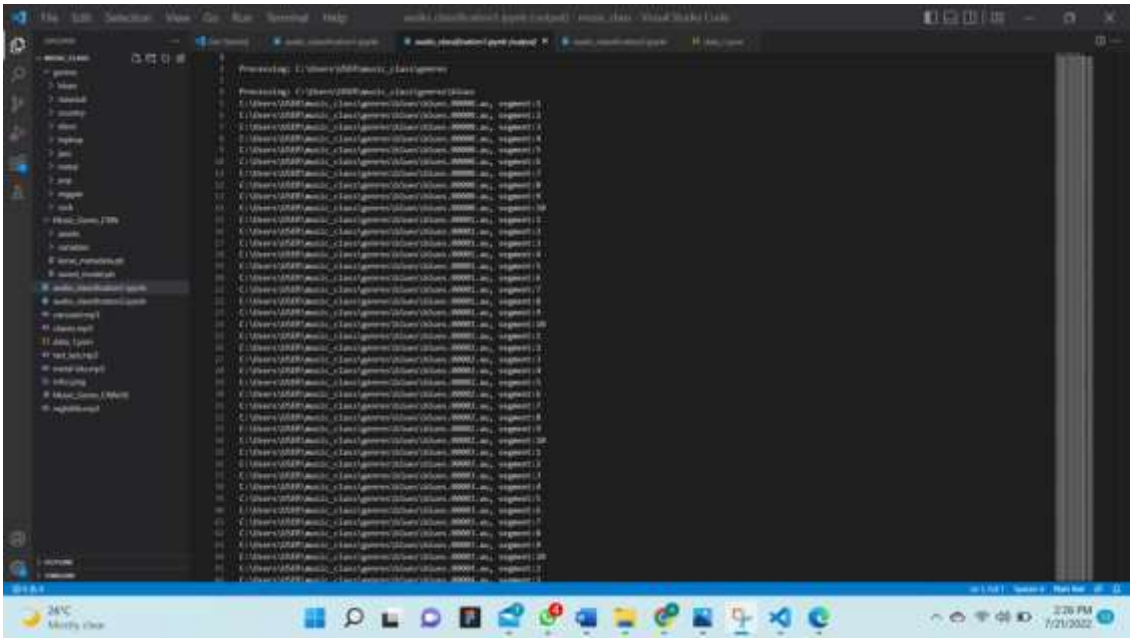

Figure 7 JSON format for the GTZAN dataset

Convolutional Neural Network (CNN) machine learning was used to create and train the model for classifying music into its relevant genre. The training dataset and the testing dataset were separated from the overall dataset. A total of 0.7 percent of the dataset was used to train the model, and the remaining 0.3 percent was used to test the model. A score of 0.738 percent was obtained for accuracy after the model had been successfully trained. The accuracy score for the trained model and the test accuracy are plotted as Shown in figure 8.

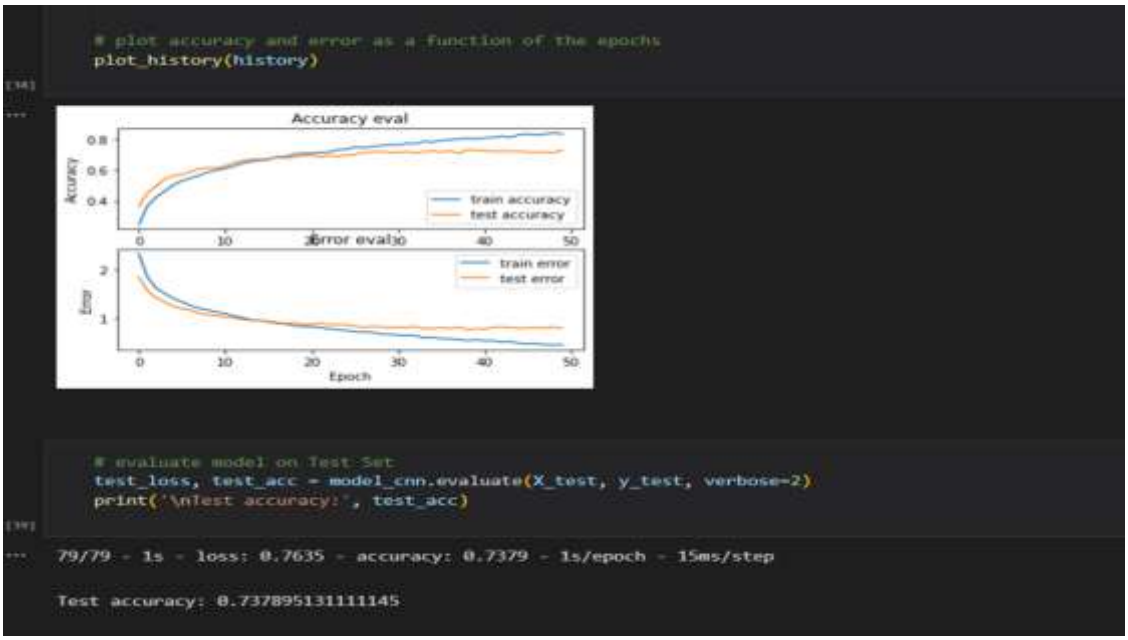

Figure 8. A plot showing the accuracy score from training the model.

4.2 Discussion:

An additional audio file that was downloaded from the internet and not included in the original dataset was used to further verify the model's correctness. In order to categorize the audio into its genre, the audio path (the system location where it is stored) was entered into the model after the music had been downloaded. The algorithm effectively classified the downloaded music into its appropriate genre, according to online sources, the first audio was pop, and when an algorithm was used to determine its genre, it correctly identified this music as pop.

 https://doi.org/10.59568/JASIC-2024-5-1-04

5. Conclusion:

This study was able to categorize different types of music into their respective genres such as (Blues, Classical, Country, Disco, Hip-Hop, Jazz, Metal, Pop, Reggae, and Rock.). This study also shows the different ways in which researchers have helped to solve music genre classification problems. Hence, the proposed system of this project worked with a Convolutional Neural Network (CNN) model, which helps to predict music into their respective genre. The model had an accuracy of 0.73% which is 73% after the model was trained. Additionally, the model was tested on brand-new music that wasn't part of the dataset, and it was successful in predicting the genre of the music. This study provides a new and easy way to help classify music into their respective genre (only the music genres as seen in the scope of study). It also helps to classify music into genres in a faster way.

6. References:

[1] Clarke, E., DeNora, T., & Vuoskoski, J. (2015). Music, empathy and cultural understanding. Physics of life reviews, 15, 61-88.

[2] Nam, J., Choi, K., Lee, J., Chou, S. Y., & Yang, Y. H. (2018). Deep learning for audio-based music classification and tagging: Teaching computers to distinguish rock from bach. IEEE signal processing magazine, 36(1), 41-51.

[3] Rustamov, I. T., Mamatqulov, J. S., & Boymatov, A. A. (2021). The Place of Small Genre Texts in Genre Features. Annals of the Romanian Society for Cell Biology, 5325-5331.

[4] Kumar, B., & Alraisi, S. R. (2022, May). Deep fake's audio detection techniques using deep convolutional neural network. In 2022 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COM-IT-CON) (Vol. 1, pp. 463-468). IEEE.

[5] Hasan, R., Hossain, S., Alam, F. I., & Barua, M. (2021, December). Bangla Music Genre Classification Using Fast and Scalable Integrated Ensemble Boosting Framework. In 2021 3rd International Conference on Sustainable Technologies for Industry 4.0 (STI) (pp. 1-6). IEEE.

6] Prajwal, R., Sharma, S., Naik, P., & Mk, S. (2021). Music genre classification using machine learning. Int. Res. J. Mod. Eng. Technol. Sci, 3(7), 953-957.

[7] Raval, M., Dave, P., & Dattani, R. (2021). MUSIC GENRE CLASSIFICATION USING NEURAL NETWORKS. International Journal of Advanced Research in Computer Science, 12(5).

[8] Chillara, S., Kavitha, A. S., Neginhal, S. A., Haldia, S., & Vidyullatha, K. S. (2019). Music genre classification using machine learning algorithms: a comparison. Int Res J Eng Technol, 6(5), 851-858.

[9] Bahuleyan, H. (2018). Music genre classification using machine learning techniques. arXiv preprint arXiv:1804.01149.

 https://doi.org/10.59568/JASIC-2024-5-1-04