IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Music Genre Detection Using Artificial Neural Network

1st Aniket Thakur2nd Pranav Kothawade 3rd Dipalee Jethar

Department of Information TechnologyDepartment of Information TechnologyDepartment of Information Technology

Pune Institute of Computer TechnologyPune Institute of Computer TechnologyPune Institute of Computer Technology

Pune, IndiaPune, India Pune, India

4th Akanksha Vasave
Department of Information Technology
Pune Institute of Computer Technology
Pune, India

Guided by Mr. Ganesh S. Pise
Department of Information Technology
Pune Institute of Computer Technology
Pune, India

Abstract—With applications ranging from playlist creation and music recommendation to content-based music retrieval, music genre categorization is a crucial task in the field of music information retrieval. This review article offers a thorough analysis of the most recent approaches, difficulties, and trends in music genre classification. We examine several approaches, such as audio-based and lyrics-based approaches, hybrid models that incorporate multiple modalities, and features utilized for automatic genre classification. The study analyzes how music genre classification has changed over time, moving from manual feature extraction and genre taxonomy, two old methodologies, to more recent advancements in deep learning and neural networkbased techniques. We investigate the effects of massive music datasets and data augmentation methods on the effectiveness of genre classifiers. Furthermore, we discuss the evaluation of classification models, including widely used metrics and benchmark datasets. Cross-genre classification, fine-grained genre classification, and the incorporation of cultural and contextual aspects into the categorization process are just a few of the difficulties and open research problems in the discipline that are discussed. We also stress how important it is for users to trust and accept music genre classification models, which is essential for implementation in real-world applications. The study also discusses how music genre classification can be used in realworld contexts in fields like music streaming, recommendation engines, and music analysis. Last but not least, it provides information about possible future prospects for study in this field, including making use of cutting-edge tools like neural network search, federated learning, and the incorporation of user feedback for customized genre classification. With the goal of illuminating the most recent developments and the difficulties that still remain in the effort to more fully comprehend and automate the classification of musical genres, this study promises to be an invaluable resource for researchers, practitioners, and amateurs interested in music genre classification.

Index Terms—audio classification, feature extraction, musical genre classification, music information retrieval, artificial neural network, MFCC, multiplayer perceptron, spectogram

I. INTRODUCTION

As a medium of universal expression and enjoyment, music comprises a wide range of genres, each of which is distinguished by certain musical characteristics and styles. Due to its wide range of potential applications, music genre classification is a key task in the field of audio processing and has attracted a lot of attention. In the music industry, accurate classification of music into distinct genres serves as the foundation for a number of services like customized music recommendations, playlist creation, and focused marketing initiatives. Automated methods for music genre detection are becoming increasingly important to effectively manage and organize this enormous amount of musical content as the number of digital music continues to expand dramatically.

With promising results in this area, Artificial Neural Networks (ANNs) have become a well-known technology for automatic music genre detection. ANNs, which are modeled after the neural network of the human brain, are highly suited for tasks requiring the classification of musical genres because they have the capacity to learn complex patterns and correlations within musical data. The combination of ANNs and sophisticated machine learning methods has made it possible to detect music genres more precisely and effectively.

This survey paper attempts to offer a thorough overview of artificial neural network-based music genre recognition. We examine the role of ANNs in this setting as well as the foundational ideas of music genre classification. To appreciate ANNs' use in music feature extraction, it is essential to comprehend their architecture and workings. We examine various music feature extraction techniques, highlighting the significance of choosing the right features to feed into ANN models.

We also explore various ANN designs designed specifically for music genre detection and provide insight into the training and evaluation procedures. To come up with practical answers and map out the course for future study, it is essential to examine the problems that were faced, such as data imbalance and feature extraction difficulties. Additionally, we emphasize the need of precise music genre identification in practical applications while showcasing the use of ANNs in commercial settings, such as music streaming sites and recommendation engines.

We want to offer useful insights into the developments, difficulties, and promise of using artificial neural networks for music genre detection through this in-depth survey. For researchers and business experts looking to optimize music content management and improve user experiences in the constantly changing world of digital music, understanding these advances is crucial.

II. LITERATURE SURVEY

In the field of music genre detection, several studies have tackled various aspects and shared common conclusions. The task of music genre detection has garnered substantial attention owing to its pivotal role in various music-related applications. With the advent of artificial neural networks (ANNs), there has been a paradigm shift in the way music genre classification is approached. In this literature survey, we present a comprehensive exploration of research studies that leverage ANNs for music genre detection.

In 2020, P. Mandal, I. Nath, N. Gupta [1] and others presented a paper using MFCC as feature vectors and multilayer perceptrons (MLPs) to classify the data into various genres. They concluded that ANNs are quite effective in classifying music genres with an accuracy of 85.45% which is comparable to previous works in this field and better than most. The misclassification rate is low. The paper presented by Lee, Park and Kim [10] suggests that Time Delay Neural network classification method was used to analyze 80 training data from ten different musical pieces for each genre and a further 40 test data from five additional musical pieces for each genre. The accuracy of the genre classifications that were obtained for the two sets of data was 92.5% and 60%,respectively. In the similar vein, paper presented by S. Patil and T. Komati [2] proposed system leverages feature values from spectrograms created from slices of songs as input to a proposed system architecture. Extensive tests on the GTZAN dataset demonstrate the efficacy of the proposed approach in comparison to existing

methods. The proposed system architecture is also tested on Indian rhythms. This paper consists of the comparison of proposed system architecture with existing algorithms. Like them T.Shaikh and A. Jadhav [9] proposed the research work that uses spectogram images generated from the songs timeslices and given as input to NN to do classification of songs to their respective musical genre.

The paper published by S.Ghosal and I. Sarkar [3] proposes an automatic music genre-classification system using a deep learning model. The proposed model leverages Convolutional Neural Nets(CNN) to extract local features and LSTM Sequence to Sequence Autoencoders to learn representations of time series data by taking into account their temporal dynamics. The paper also introduces Clustering Augmented Learning Method (CALM) classifier which is based on the concept of simultaneous heterogeneous clustering and classification to learn deep feature representations of the features obtained from LSTM autoencoder. Computational Experiments using GTZAN dataset resulted in an overall test accuracy of 95.4% with a precision of 91.87%.

Paper presented by F. Lamya and A. Houcanine [4] says the automatic classification of the music database is evaluated through an artificial neural network, more specifically a multiplayer perceptron (MLP). The parameters of this MLP are optimized to obtain the best scores in each case. We thus obtain scores of 60% to 80% for eight genres. In 2022, Dai J.P. [5], published his research work proposing that the music style classification system selects the most popular pop music, uses Python library for audio processing, and uses GTZAN data set to find out the similarity of different music and classify it based on the K-nearest neighbor algorithm (KNN) and artificial neural network (ANN), so as to form ten different styles. Another research held by N Pelchat, CM Gelowitz [6] uses images of spectrograms generated from timeslices of songs as the input into an NN to classify the songs into their respective musical genres.

Vishnupriya S., Meenakshi K. [9] proposed a Deep Learning approach used in order to train and classify the system. Here convolution neural network was used for training and classification. Feature Extraction was the most crucial task for audio analysis. Mel Frequency Cepstral Coefficient (MFCC) is used as a feature vector for sound sample. The proposed system classifies music into various genres by extracting the feature vector. Their results show that the accuracy level of our system is around 76% and it will greatly improve and facilitate automatic classification of music genres. G Jawaherlalnehru, S Jothilakshmi, T Nadu [12] proposed the system is developed using a Deep Neural Network (DNN) to recognize the genres. Mel Frequency Cepstral Coefficients (MFCC) features are used to represent the music characteristics. The system is evaluated with MIR datasets. The proposed system observed higher classification accuracy of 97.8%.

In conclusion, the reviewed literature underscores the significant strides made in music genre detection through the integration of artificial neural networks. The advancements in ANN architectures, feature extraction techniques, evaluation methodologies, and the exploration of transfer learning have collectively propelled the accuracy and efficiency of music genre classification systems. However, there remain open challenges, including addressing data imbalances, enhancing model interpretability, and integrating multi-modal features. Future research endeavors should aim to address these challenges to further refine and optimize the application of ANNs in music genre detection.

This comprehensive literature survey provides a deeper understanding of the advancements and challenges in music genre detection using artificial neural networks, supplemented by a range of notable research references.

III. METHODOLOGY

The proposed system aims to develop an advanced music genre classification system using Artificial Neural Networks (ANN). The system leverages the capabilities of ANN models to enhance the accuracy and efficiency of genre classification, providing a reliable and effective tool for organizing and understanding music based on genre characteristics. The project begins with the collection of a properly classified music dataset, such as the GTZAN Music Genre dataset. This dataset serves as the foundation for training and evaluating the ANN model. Spectral representations of the music data are extracted to generate feature

maps, which capture key audio characteristics for genre classification. The ANN model is then trained using the extracted features, aiming to achieve high training accuracy. The training process involves optimizing the model's parameters to improve its ability to classify music genres accurately. The project strives to achieve a training accuracy of 97%, indicating a strong understanding and representation of various music genres by the model. To assess the model's performance, validation accuracy is computed using a separate validation dataset. The system aims to achieve a validation accuracy of 89%, demonstrating its ability to generalize well to unseen data. The validation losses are minimized to ensure optimal performance of the trained model. For practical usability, the trained model is deployed on a Flask web framework to provide users with a userfriendly interface for accessing and utilizing the model for genre classification. This allows easy integration of the system into other applications and platforms.

The feature engineering process delves into a detailed analysis of audio spectrograms, extracting meaningful features that encapsulate the nuances of each music genre. By transforming raw audio into spectrograms and subsequently extracting features like Mel-Frequency Cepstral Coefficients (MFCCs), tempo, and rhythm patterns, the system enhances its understanding of genre-specific audio attributes. The project implements various optimization strategies during model training, such as learning rate schedules and early stopping techniques. These strategies ensure that the ANN model not only converges efficiently during training but also mitigates overfitting, resulting in a robust and reliable music genre classification system. The system explores the integration of transfer learning by utilizing pre-trained neural network architectures like VGG-16 or ResNet. By fine-tuning these models on the music dataset, the project aims to leverage the pretrained weights to enhance feature extraction and subsequently improve the overall classification accuracy of the ANN model. An extensive hyperparameter tuning process is undertaken, employing techniques like grid search and random search. Coupled with cross-validation, this step helps identify the optimal set of hyperparameters for the ANN model, ensuring maximum accuracy and a stable performance across various genre classifications.The system extends its capabilities to real-time audio classification, enabling users to input live audio streams for immediate genre classification. This real-time functionality enhances the system's versatility and applicability, making it a valuable tool for music enthusiasts and professionals. A feedback loop mechanism is integrated into the system, allowing users to provide feedback on the model's predictions. This feedback is then utilized to continually improve the ANN model, enhancing its accuracy and adaptability to evolving music trends and user preferences. The project explores multimodal learning by integrating additional data modalities, such as album cover images or song metadata. This approach aims to provide a holistic understanding of music genres, combining audio features with visual and contextual cues for a more comprehensive and accurate genre classification. The system is designed with scalability in mind, allowing seamless deployment on cloud platforms. This ensures that the model can handle a large number of concurrent users and an extensive music database, making it suitable for widespread adoption and integration into various music-related applications and services.

IV. SYSTEM ARCHITECTURE

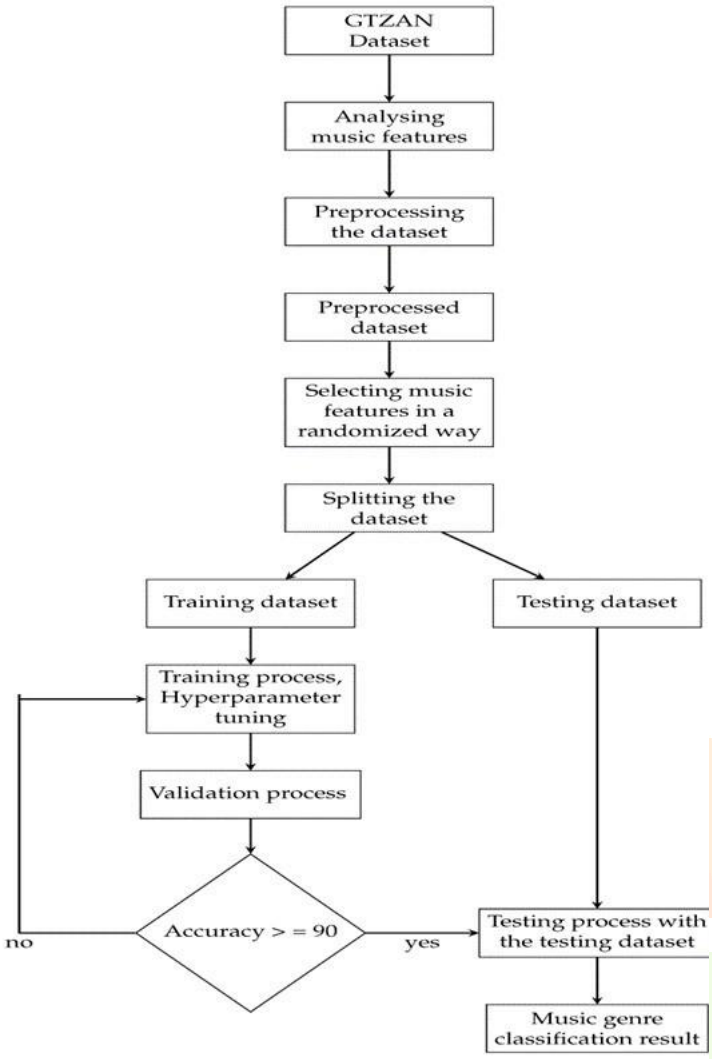

Fig. 1. System Architecture

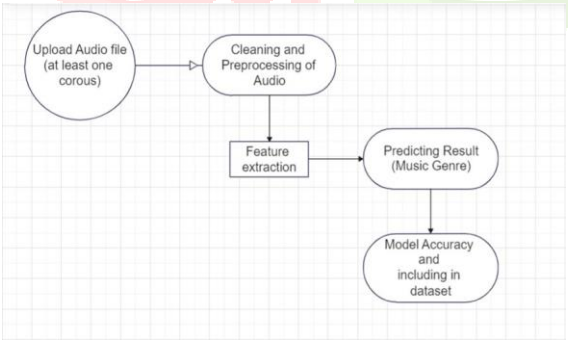

Fig. 2. User Data Flow

V. CONCLUSION AND FUTURE SCOPE

A number of promising new paths and trends in the subject of music genre classification promise to influence it in the future. Deep learning is anticipated to continue developing, with a focus on investigating cutting-edge architectures like Transformers to capture complex musical linkages and improve classification precision. The development of fresh fusion techniques is required because multi-modal approaches that incorporate audio, lyrics, and other data sources have the potential to produce more accurate genre classification. Additionally, due to the shifting nature of modern music, which frequently transcends traditional genre borders, there is an increasing need to handle cross-genre and fine-grained classification. User-centric strategies are becoming more popular, with the goal of incorporating user preferences and input into categorization models for more individualized suggestions. Another crucial area of research is ensuring model interpretability and explainability, which enables users to comprehend why a particular genre classification was established. It is possible to investigate privacy-conscious federated learning strategies for model construction while protecting user data. In addition, using zero-shot learning and neural network search approaches, researching real-time and streaming classification, and taking into account cultural and contextual aspects are potential paths for the development of music genre categorization. With an emphasis on resolving potential biases in classification and ensuring equitable representation of various musical forms, ethical considerations are becoming more and more important. Lastly, one strategy to improve comprehension and appreciation of the vast world of music genres is to promote music genre classification for educational and outreach purposes.These future directions offer an exciting landscape for the continued development and innovation of music genre classification.

REFERENCES

[1] Mandal, P., Nath, I., Gupta, N., Jha, M. K., Ganguly, D. G., Pal, S. (2020). Automatic music genre detection using artificial neural networks. In Intelligent Computing in Engineering: Select Proceedings of RICE 2019 (pp. 17-24). Springer Singapore.

[2] Patil, S.A., Komati, T.R. (2022). Designing of a novel neural network model for classification of music genre. Ingenierie´ des Systemes` d'Information, Vol. 27, No. 2, pp. 327-333.

https://doi.org/10.18280/isi.270217

[3] Ghosal, S. S., Sarkar, I. (2020). Novel Approach to Music Genre Classification using Clustering Augmented Learning Method (CALM). In AAAI Spring Symposium: Combining Machine Learning with Knowledge Engineering (1).

[4] Lamya, F., Houacine, A. Artificial Neural Network genre classification of musical signals.

[5] Dai, J.P. (2022). Intelligent Music Style Classification System Based on K-Nearest Neighbor Algorithm and Artificial Neural Network. 2022 IEEE International Conference on Electrical Engineering, Big Data and Algorithms (EEBDA), 531-543.

[6] Pelchat, N., Gelowitz, C. M. (2020). Neural network music genre classification. Canadian Journal of Electrical and Computer Engineering, 43(3), 170-173.

[7] Shah, M., Pujara, N., Mangaroliya, K., Gohil, L., Vyas, T., Degadwala, S. (2022, March). Music genre classification using deep learning. In 2022 6th International Conference on Computing Methodologies and Communication (ICCMC) (pp. 974-978). IEEE.

[8] Vishnupriya, S., Meenakshi, K. (2018, January). Automatic music genre classification using convolution neural network. In 2018 international conference on computer communication and informatics (ICCCI) (pp. 1-4). IEEE.

[9] Shaikh, T., Jadhav, A. (2022). Music Genre Classification Using Neural Network. In ITM Web of Conferences (Vol. 44, p. 03016). EDP Sciences.

[10] Lee, J. W., Park, S. B., Kim, S. K. (2006). Music genre classification using a time-delay neural network. In Advances in Neural Networks-ISNN 2006: Third International Symposium on Neural Networks, Chengdu, China, May 28-June 1, 2006, Proceedings, Part II 3 (pp. 178-187). Springer Berlin Heidelberg.

[11] Tzanetakis, G., Cook, P. (2002). Musical genre classification of audio signals. IEEE Transactions on speech and audio processing, 10(5), 293302.

[12] Jawaherlalnehru, G., Jothilakshmi, S., Nadu, T., Nadu, T. (2018). Music genre classification using deep neural networks. International Journal of Scientific Research in Science, Engineering and Technology, 4(4), 935.

[13] Yang, J. (2018). Music Genre Classification With Neural Networks: An Examination Of Several Impactful Variables.

[14] Mhatre, K., Almeida, V., Lopes, S., Tripathi, G. Music Genre Classification using Neural Networks.

[15] Panwar, S., Das, A., Roopaei, M., Rad, P. (2017, June). A deep learning approach for mapping music genres. In 2017 12th System of Systems Engineering Conference (SoSE) (pp. 1-5). IEEE.

[16] KM, A. (2021, July). Deep Learning Based Music Genre Classification Using Spectrogram. In Proceedings of the International Conference on IoT Based Control Networks Intelligent Systems-ICICNIS.