WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Feature Extraction and Classification Techniques in Music Genre Analysis

Gembali Saikiran

Student, Rajam, Vizianagaram, 535127, Andhra Pradesh ,India.

ABSTRACT

Music Information Retrieval(MIR) is the task of extracting high-level information from music such as genre, artist or instrumental from music. Music Genre Classification(MGC) is an important and rapidly evolving research area MIR. With rapid growth ,MGC can be used in lot of ways to organize and manage music recommendation systems, advertising, and streaming services. The task typically involves two stages: Audio feature extraction and machine learning modelling. The study aims to compare and analyze the feasibility, performance and understandability of features used to describe music by predicting the genre using Machine learning algorithms like Support Vector Machine , K-Nearest Neighbor, Random Forest. The application is very important and requires automation to reduce the manual error and time because if music is classified manually then one have to listen out each file for complete duration. Spotify and Sound Cloud apply genre classification to recommend music to their users.

Keywords: Music Genre Classification, feature Extraction, Support vector Machine, K-Nearest Neighbor, Random Forest.

1. Introduction

Music serves as a source of pleasure and entertainment, offering relaxation all while exerting significant effects on social and psychological aspects of our lives. It is a powerful medium that convey meaning to the listeners by combining a variety of musical elements synchronously and sequentially. Music genre classification involves training model to automatically categorize songs into predefined genre based on audio features. Musical classifications are made up of two fundamental consistent whiles been input with raw audio input and process with classification tags based on characteristics of known database. Since most of the music in the current music streaming media only has the title and author, and most of this music does not carry specific tags. This makes it a challenging task to be able to identify the hidden tags in songs and to classify songs according to genre. In this research GTZAN dataset used it consist of 1000 audio tracks each 30 seconds long .It contain 10 genre each represented by 100 tracks. The 10 genres are blues classical, country, disco, hip-hop jazz ,metal , pop, reggae ,rock. These audio clips are typically in WAV format .To train the machine learning model you need to extract relevant features from audio some of the features include Mel frequency cepstral coefficient ,chroma feature, spectral contrast, tempo, zero crossing rate, spectral bandwidth. Train the selected model on training dataset using extracted audio features as input and as genre labels as target output.

2. Literature Survey

In paper [1]. This paper is the first to explore the task of automatic music genre classification for traditional Nigerian music. To introduce the ORIN dataset, containing Nigerian songs, as a resource for music genre classification research. For genre classification, the ORIN datasets was trained on 4 different classifiers- k-Nearest Neighbor, Support Vector Machine, extreme Gradient Boosting(XGBoost) and Random Forest- with 85–15 train-test splits. The ORIN dataset may be limited in size, potentially impacting the generalizability of the findings. Use of multiple machine learning models for comprehensive evaluation. XGBoost as the best-performing model with an accuracy of 81.94% and a recall of 84.57%. here is a need for further research to expand the dataset to include more diverse Nigerian-English contemporary songs, potentially improving model generalizability.

In paper [2]. This paper presents a web application that retrieves songs from YouTube and classifies them into music genres. For this purpose, they have used classifiers from distinct Machine Learning paradigms: Feed-forward and Recurrent Neural Networks and Support Vector Machines (SVMs), Decision trees, fully connected neural networks (FCNNS),convolutional neural networks (CNNS),recurrent neural networks (RNNS). The lack of a standardized and universally accepted definition of music genres poses challenges in validating genre predictions, especially when comparing them with categorizations from other sources. The application is also a first step towards an eventual user centered MGC tool, in which the users can submit feedback about the correctness of the predictions. The deep learning models produce results comparable with the Audio set baseline generic classifier study, which achieves a mean average precision (AP) of 0.314 and an average AUC score of 0.959. While the application introduces the concept of user-generated feedback for genre classification, its validity and effectiveness as a validation tool require further investigation.

In paper [3]. The study aims to experimentally verify the usability of the trajectory of fifths as a source of knowledge in automated music classification processes. The study experiments were limited to only two genre groups, which may not fully represent the complexity of real-world music genre classification scenarios. Outcomes of the study encourage further research into the use of the trajectory of fifths in music classification, particularly in more complex genre classification scenarios. Mean balanced accuracies exceeding 0.9 for all machine learning algorithms. While the basic coefficients yielded promising results, there is a potential gap in exploring more complex and advanced features related to the trajectory of fifths for enhanced music classification accuracy.

In paper [4]. In this paper, the author presents a novel approach that combines multiple types of information related to music using cross-modal contrastive learning, allowing us to learn an audio feature from heterogeneous data simultaneously. The study mentions that it did not utilize all available types of information in the dataset, such as title, playlist tags, authors, and other metadata. This limitation may leave untapped potential for improving the model. Using contrastive learning allows us to reach higher performance than using the models trained directly to predict the genre. Playlist-level information which will require an additional level of abstraction.

In paper [5]. The study aims to implement a machine learning system for musical genre classification using Convolutional Neural Networks (CNNs). The current implementation tests accuracy on only a short segment (2.56 s) of songs, which may not fully represent the diversity of music genres within longer songs. The author discusses future work related to modifications and feature engineering, indicating a commitment to improving the model's performance. The reported accuracy of 85% on the test data suggests that the implemented CNN-based model is effective in classifying music genres. While the study mentions the issue of false positives, it does not propose specific methods or strategies for addressing this challenge.

3. Methodology

Here's a methodology flow for feature extraction and classification in music genre analysis:

1. Data Collection:

- Gather a dataset of audio files representing diverse genres.
- Ensure quality and consistent format (e.g., WAV, MP3).
- Label each file with its accurate genre.

2. Preprocessing:

- Segmentation: Divide long audio files into manageable segments for analysis.
- Normalization: Adjust audio levels to a consistent range.
- Resampling: Convert to a uniform sampling rate (e.g., 44.1 kHz).
- Mono conversion: Transform stereo tracks to mono if necessary.

3. Feature Extraction:

- Time-domain features:
    - Zero-crossing rate (ZCR)
    - Spectral centroid
    - Spectral roll-off
    - Mel-frequency cepstral coefficients (MFCCs)
    - Rhythmic features (tempo, beat strength)
- Frequency-domain features:
    - Spectral flux
    - Spectral density

4. Feature Selection (optional):

- Identify the most relevant features for genre classification.
- Reduce dimensionality and computational cost.
- Use techniques like correlation analysis, principal component analysis (PCA), or feature importance scores.

5. Classification:

- Choose a classification algorithm:
    - Support vector machines (SVMs)
    - Decision trees
    - Random forests
    - Neural networks (deep learning)
- Split the dataset: Divide data into training and testing sets.
- Train the model: Feed features and genre labels to the algorithm.
- Evaluate performance: Assess accuracy, precision, recall, and F1-score on the testing set.

6. Model Deployment:

- Integrate the trained model into applications for genre classification.
- Use it for music recommendation, organization, or content-based retrieval.

7. Refinement:

- Explore different feature combinations and classification techniques.
- Optimize hyperparameters for improved accuracy.
- Continuously evaluate and refine the model based on new data and performance metrics.

4. Results and Discussion

Algorithm	**Accuracy**
XGBoost	82.0%
ATMGCM	85.3%
CNN	84%
Neural Networks	82%
CNN	91%

The primary task at hand is coming up with an organized system for categorizing different musical genres. It can be challenging to distinguish between various musical genres at this point because they can blend. The beat, the instruments used, and the mood a song conveys are just a few examples of the unique sounds and components that are discussed by genre. There is disagreement over the most effective techniques for teaching computers to identify these genres. Some argue that sophisticated new techniques such as deep learning which is inspired by the way our brains learn are the way to go, while others favor more straightforward, conventional approaches for teaching computers to identify patterns in music. Furthermore, the conversation goes beyond technology to include how these labels impact our ability to discover and enjoy new music as well as how they may represent societal influences and trends. The accuracy value of the XGBoost classifier for the ORIN datasets was 81.94% for classifying Nigerian songs, and 84.57% for recall, the BMNET-5 classifier had an accuracy value of 90.32% for classifying Bangla songs. Therefore, Using the XGBoost classifier may get more accuracy in classifying the music genre.

5. Conclusion

The model helps to categorize music according to its genre. A fascinating and constantly developing field of study, machine learning approaches for music genre classification have major practical applications and implications for the music industry. The accuracy value of the XGBoost classifier for the ORIN datasets was 81.94% for classifying Nigerian songs, and 84.57% for recall; the BMNET-5 classifier had an accuracy value of 90.32% for classifying Bangla songs. Research into machine learning methods for classifying music genres has revealed a variety of approaches, such as deep learning models and feature-based techniques. Spotify and Sound Cloud use genre categorization to suggest songs to their users. The subjective and dynamic nature of genres means that classifying music remains a difficult task despite advancements. By employing sophisticated models and expanding the size of the dataset, genre classification can be made more effective. Furthermore, there is a strong need for additional research in the field of music genre classification, especially in light of the increased diversity of genre labels and the consideration of multiple languages.

REFERENCES

[1] Folorunso, S. O., Afolabi, S. A., & Owodeyi, A. B. (2022). Dissecting the genre of Nigerian music with machine learning models. Journal of King Saud University-Computer and Information Sciences, 34(8), 6266- 6279

[2] Castillo, J. R., & Flores, M. J. (2021). Web-based music genre classification for timeline song visualization and analysis. *IEEE Access*, *9*, 18801-18816

[3] Łukaszewicz, T., & Kania, D. (2022). A Music Classification Approach Based on the Trajectory of Fifths. *IEEE Access*, *10*, 73494-73502.

[4] Ferraro, A., Favory, X., Drossos, K., Kim, Y., & Bogdanov, D. (2021). Enriched music representations with multiple cross-modal contrastive learning. *IEEE Signal Processing Letters*, *28*, 733-737.

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