BANGLAGITI: A NOVEL DATASET FOR BANGLA MUSIC GENRE CLASSIFICATION WITH A COMPREHENSIVE ANALYSIS

Anonymous authors
Paper under double-blind review

ABSTRACT

This paper presents a comprehensive exploration into the classification of Bengali music genres, utilizing a novel dataset, **'BanglaGITI: Bangla Genre-wise Indexed Tracks and Interpretations'**, specifically curated to capture the rich diversity of Bengali musical heritage. Our study is structured around a comparative analysis of traditional Machine Learning (ML) techniques, advanced Deep Learning (DL) methodologies, and innovative ensemble approaches that integrate the strengths of both ML and DL through Transfer Learning. Our dataset includes a total of 1410 audio files across 6 different genres. For the ML segment, features such as Mel-frequency cepstral coefficients (MFCCs), zero-crossing rate (ZCR), root mean square(RMS), chroma, tempo and spectral bandwidth were leveraged to encapsulate the unique characteristics of Bengali music. These features serve as a foundation for employing classic ML classifiers that demonstrate robust performance in genre classification tasks. Our methodology includes Decision Tree, Random Forest, Gradient Boost and KNN. Conversely, our DL models are designed around the extraction and analysis of Log-Mel spectrograms, capitalizing on their ability to represent complex musical structures in a manner that is both comprehensive and conducive to DL techniques. This approach allows for the deep neural networks to learn from a richer representation of audio data, potentially uncovering nuanced patterns inherent in Bengali music genres. DL techniques feature pre-trained CNN-based models such as DenseNet, ResNet and VGGNet. Furthermore, our paper innovates by proposing ensemble models that combine the predictive capabilities of ML and DL methods respectively, aiming to harness their complementary strengths for enhanced classification accuracy. The ensemble models resulted in achieving almost 80% accuracy in ML and state of the art **96%** accuracy in DL methods while the precision recall and F1-score of **96.09%**, **96.05%** and **96.04%** respectively. Our findings not only shed light on the efficacy of different computational approaches in the realm of music genre classification but also contribute to the understanding of Bengali music through the lens of machine intelligence. The use of our self-made dataset, which is among the first of its kind for Bengali music, adds a significant value to the study, offering a new benchmark for future research in this area. Through this comprehensive study, our aim is to provide insights that will guide the development of more sophisticated and culturally nuanced music classification systems.