Paper Title:
Music Genre Classification using Deep Learning

Paper Link:
https://ieeexplore.ieee.org/document/9753953

1 Summary:

1.1 Motivation

This paper addresses the need to develop personalized recommendation systems for music which categorize music into genres on the basis of factors like the tempo, harmony, as well as instruments. The method involves using the GTZAN data set to build machines to classify music as well as improve the customer experience.

1.2 Contribution

The authors have created an original Convolutional Neural Network (CNN) model that can be used for classifying music by genre. It has features like "relu" activation that improves efficient computation, Max-Pooling2D layer to reduce complexity, and a Dropout layer that prevents overfitting. The model also adds more filters in order to take advantage of more intricate capabilities.

1.3 Methodology

The process includes feature extraction using the Librosa Python package. It also trains models employing the GTZAN dataset. The GTZAN dataset comprises 1000 audio tracks that are split into ten categories. The features extracted are derived from both frequency and time domains.

1.4 Conclusion

The paper ends with a brief recap of its design and brief discussions about related work on the classification of music genres, including the application of different machine learning methods and the role of CNN for improving the accuracy of tests.

2 Limitations

2.1 First Limitation

The deep learning models used to classify music genres depend heavily upon the high quality recorded audio files that are used as input. The precision at which these models recognize specific features of genres is diminished in the event of audio that is noisy or damaged that could lead to inaccurate classifications. To counteract these issues using preprocessing methods, such as noise reduction are able to cleanse the sound prior to it entering the model. Another approach is for the model to be exposed to diverse audio qualities throughout the training process, including ones with distortion and noise and thereby preparing the model with the ability to deal with the various scenarios that require audio quality that cannot be assured.

2.2 Second Limitation

A different issue arises when deep-learning models try to discern between genres featuring similar musical traits, for example, metal and rock and metal, that have a lot in common with musical instruments and styles of singing. To be able to recognize the subtle differences between genres the model needs a wide, large and comprehensive dataset that reveals the distinct characteristics of every genre. Data augmentation methods can increase the variety of training data by introducing different variations to the sound that allow the model to adapt to changing conditions. In addition, implementing the multi task approach to learning lets the model learn from multiple perspectives, and not simply analyzing genres, but making sense of the instruments performed, leading to an improved knowledge and the separation of related genres.

3 Synthesis

The paper overall helps in the area of music genre classification through formulating the CNN model to enhance the accuracy in classifying music genres. The work is based on current machine learning methods and provides enhancements to the CNN structure to improve results in music recommendation systems. A complete synopsis that includes limitations is not feasible currently because the section on limitations could not be found in the text.