INTRODUCING R-CRNN FOR MUSIC GENRE CLASSIFICATION

Rationale/ Introduction

With the rapid advancement of technology in today's digital era, music genre classification has fallen behind its music recognition counterpart in terms of speed and accuracy. Moreover, these existing approaches often rely on a limited database, restricting the capabilities of current systems (Tudor et al., 2020). Additionally, the evolution of music and its genres poses a significant challenge to technology. In the past, music genres represented distinct differences between songs. However, these strict boundaries have blurred in today's age due to the influence of technology and other factors (Neu, 2024). Consequently, it becomes increasingly difficult to classify music genres accurately and precisely. Another problem is the lack of exploration on the potential of the architecture of Convolutional Recurrent Neural Network (CRNN), if explored, the results came behind the classical models used to identify music genres. In a study conducted in 2021, a comparative study was conducted on deep learning models for music genre classification. The deep 2 learning models were CNN, RNN, and CRNN. The researchers used different types of architecture in these deep learning models, and the results showed that CNN and RNN had outperformed CRNN in accuracy and precision (Rafi et al., 2021). Although there are studies regarding CRNN, mostly in other areas but not in music, there is little to no research or resources about CRNN.

To address this gap, this research proposes a novel approach to music genre classification that incorporates residual connections to enhance the retention of musical features. By doing so, the study contributes significantly to the field by introducing a fresh perspective on a standard Convolutional Recurrent Neural Network (CRNN) architecture. Approaching CRNN with residual connection layers will help maintain the musical features of a song and will be helpful for the RNN layers to classify the music genres more clearly. Another gap that this research proposal aims to address is to contribute in this field by conducting this study.

Conducting this study will produce a new type of architecture of CRNN, namely Residual Convolutional Recurrent Neural Network (R-CRNN). The R-CRNN will perform better in music genre classification in terms of accuracy and precision. Additionally, this architecture will benefit deep learning in the areas of music and other similar areas that

can benefit the most by using residual networks. It can also be beneficial in the allied health field for better machine interpretation of X-Rays, Magnetic Resonance Imaging (MRI).

This study proposes a novel approach of Convolutional Recurrent Neural Network architecture. This novel approach aims to solve the research gap found in a study conducted in 2021 that CRNN fell behind classical models in music genre classification. Thus, this study proposes to develop a new architecture to solve this gap. The expected output of this study will see an increase of performance in music genre classification. Another expected output of this study is to introduce R-CRNN that will be the main contribution in this CS field.

Significance of the Study

The rise of technology has been the primary driving force behind the evolution of artistry and musicality, leading contemporary music to transcend genre boundaries that once defined songs. This convergence has prompted researchers and developers to seek innovative solutions to accurately differentiate songs and classify them as the songwriters intended.

This study is driven by the urgent need for more efficient and consistent classification models for automated tagging of music genres uploaded to various Digital Streaming Platforms (DSPs). While numerous music genre classification models have been extensively studied, existing methods often fall short in terms of accuracy and precision, making them unsuitable for real-world applications.

This study introduces an innovative approach to Convolutional Recurrent Neural Network (CRNN) model architecture by incorporating residual connections. This method not only enhances the robustness and performance of a standard CRNN but also offers a novel way to potentially preserve musical features for a longer duration while processing layers of CRNN. Compared to conventional neural network models, Residual CRNN leverages the powerful features of Residual Neural Network (ResNet) by incorporating residual blocks to retain musical features. It also capitalizes on the strengths of Convolutional Neural Network (CNN) in feature extraction. Moreover, it adopts the sequential nature of Recurrent Neural Network (RNN) as temporal features play a crucial role in music genre classification. This represents a significant advancement over conventional neural network classification

models and paves the way for new directions in research.

Scope and Limitations of the Study

The scope of this study focuses on developing a novel neural network architecture called Residual Convolutional Recurrent Neural Network (R-CRNN) for music genre classification by integrating Residual Neural Networks (ResNet) with Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) to enhance feature retention and improve classification accuracy. The research will utilize existing datasets such as GTZAN and Free Music Archive (FMA) while also curating a smaller dataset to evaluate performance on limited data. Performance evaluation will be conducted through accuracy, precision, confusion matrices, and epoch analysis, comparing R-CRNN with classical models like CNN and RNN. Additionally, a platform will be developed to deploy the neural network models for real-world testing. While this study primarily focuses on music genre classification, it will also explore the potential applicability of residual networks in medical imaging and other domains. However, this research has limitations, including dataset constraints, as the newly created dataset will be smaller than conventional datasets, which may impact generalization.

Limited computational resources may also restrict hyperparameter tuning and slow down training times. Additionally, the overlap of modern music genres introduces subjectivity in classification, potentially affecting model accuracy. The study will benchmark R-CRNN against CNN and RNN but will not include other advanced architectures such as Transformers or Attention-based models, which may offer additional insights. Lastly, while a platform will be developed for deployment, full integration with commercial streaming services (e.g., Spotify, Apple Music) is beyond the scope due to API and licensing constraints. Despite these limitations, this research aims to advance the field by improving music genre classification and setting the foundation for further exploration of residual networks in deep learning applications.

Objectives of the Study

This study aims to analyze and compare three different neural network model architecture for music genre classification. It specifically aims to:

1. Develop a novel neural network architecture that combines Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) while applying Residual

Neural Network approach.

2. Create a new dataset that features a smaller amount than conventional datasets such as GTZAN and Free Music Archive (FMA) for measuring the performance of the newly developed neural network model at a smaller, limited dataset.
3. Evaluate and analyze the accuracy of the music genre classification models architectures through epochs, confusion matrices, and tabular data using GTZAN, FMA, and the newly created dataset.
4. Compare the newly developed R-CRNN with the likes of classical, conventional models such as CNN and RNN.
5. Develop a platform for music genre classification.

Expected Outputs

The expected output of this research proposal is to have a fully developed neural network model namely R-CRNN. This novel algorithm is expected to have an increased accuracy and precision performance by at least 3%. There will also be a platform developed to deploy the neural network models in order to see the performance of the neural networks. After conducting the study, the novel algorithm will be the main contribution to the CS field for future researchers to explore its possibilities beyond music genre classification since the residual layer block can be helpful in maintaining data during training.

References

Choi, K., Fazekas, G., Sandler, M., & Cho, K. (2016). Convolutional recurrent neural networks for music classification. arXiv (Cornell University). https://doi.org/10.48550/arxiv.1609.04243

How AI is transforming the creative economy and music industry. (2024, November 5). OHIO Today. https://www.ohio.edu/news/2024/04/how-ai-transforming-creative-economy-music-industry

Katz, M. (2022). Music and Technology: A Very Short Introduction. Oxford University Press.

Rafi, Q. G., Noman, M., Prodhan, S. Z., Alam, S., & Nandi, D. (2021). Comparative analysis of three improved deep learning architectures for music genre Classification. International Journal of Information Technology and Computer Science, 13(2), 1–14. https://doi.org/10.5815/ijitcs.2021.02.01

The machine learning behind Hum to search. (n.d.). https://research.google/blog/the-machine-learning-behind-hum-to-search/